Archaea — timeline of the third domain

Ricardo Cavicchioli

Abstract | The Archaea evolved as one of the three primary lineages several billion years ago, but the first archaea to be discovered were described in the scientific literature about 130 years ago. Moreover, the Archaea were formally proposed as the third domain of life only 20 years ago. Over this very short period of investigative history, the scientific community has learned many remarkable things about the Archaea — their unique cellular components and pathways, their abundance and critical function in diverse natural environments, and their quintessential role in shaping the evolutionary path of life on Earth. This Review charts the 'archaea movement', from its genesis through to key findings that, when viewed together, illustrate just how strongly the field has built on new knowledge to advance our understanding not only of the Archaea, but of biology as a whole.

Domain
The highest level of taxonomic division; the three domains are the Archaea, Bacteria and Eukarya. In descending order, the other levels include: kingdom, phylum, class, order, family, genus and species.

Extremophile
An organism that requires extreme environments for growth, such as extremes of temperature, salinity or pH, or a combination of these.

Methanogen
An anaerobic organism that generates methane by the reduction of carbon dioxide, acetic acid, or various one-carbon compounds such as methylamines or methanol.

Halophile
An organism that requires high concentrations of salt (typically greater than 1 M NaCl) for growth.

School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, NSW 2052, Australia.
e-mail: r.cavicchioli@unsw.edu.au

doi:10.1038/nrmicro2482

The Archaea are a unique and interesting group of organisms that are clearly distinguishable from the domains of the Bacteria and the Eukarya (originally proposed as 'Eucarya')[1,2]. First and foremost, the Archaea are critical for evolutionary studies probing the origins of life, with the discovery of the third domain prompting robust debate about the course of evolution on Earth. The Archaea are also immediately recognizable as 'extremophiles'. Although not all extremophiles are members of the Archaea[3] (and not all archaea are extremophiles), many archaeal species are prevalent in extreme environments and they hold a disproportionately high number of the records for growth and survival under many ecological extremes[4]. As a result of their extremophilic nature and their role in the evolution of life, the properties of archaea have stimulated the astrobiology community to expand their horizons in the search for extraterrestrial life (BOX 1).

The Archaea are characterized by several distinguishing traits, including the incorporation of ether-linked isoprenoid lipids with a glycerol-1-phosphate backbone, and they are the only life forms known to carry out methanogenesis to produce biological methane (TABLE 1). However, the Archaea share traits with both the Bacteria and the Eukarya: they have information-processing systems (for example, for DNA replication, transcription and translation) in common with the Eukarya and metabolism in common with the Bacteria. The unique and shared traits of the Archaea have stimulated a broad range of mechanistic and evolutionary studies that have attempted to rationalize the occurrence of these traits[5,6]. By providing stable homologues and simplified cellular systems (such as the archaeal transcription apparatus), the study of archaeal proteins has accelerated our understanding of eukaryotic cell biology (BOX 1).

The Archaea have traditionally been grouped into methanogens, extreme halophiles (haloarchaea) and thermoacidophiles. The methanogens are found in anaerobic marine and freshwater environments and in the gastrointestinal tracts of animals, where they participate in the conversion of organic matter by utilizing the metabolic products of bacteria (for example, CO_2, H_2, acetate and formate) and other simple methyl compounds that are available in the environment (for example, methylamines and methanol) and converting them into methane (CH_4). Haloarchaea reside in hypersaline environments (such as salterns, lakes and the Dead Sea), where they grow as heterotrophs, often in association with phototrophic algae. Thermoacidophiles (including hyperthermophiles, which grow fastest at temperatures above 80 °C) colonize volcanic terrestrial environments and deep-sea hydrothermal vents, growing aerobically or anaerobically as heterotrophs or autotrophs, and often deriving energy by sulphur oxidation or reduction. In addition to these traditional groups, archaea have been found to carry out previously unknown functions (for example, anaerobic methane oxidation and aerobic ammonia oxidation) and to colonize vast reaches of the planet (for example, throughout the world's oceans). As a result of their diverse functional capacities, including unique specialities, and their ubiquitous abundance on Earth, the critical roles that the Archaea have in driving global biogeochemical cycles and, hence, maintaining the health of the planet are now clearly recognizable.

Box 1 | Research motivated by archaeal extremes

- Fundamental research assessing molecular, physiological and evolutionary mechanisms of adaptation. This represents a core theme in archaeal studies that continues today. Studies were driven by the intrigue of novel biochemical pathways (for example, methanogenesis) and the ability of cellular components to function effectively under extreme conditions (for example, hyperthermophiles at temperatures ≥ 100 °C, acidophiles at pH ≈ 0, and haloarchaeal proteins in 3–5 M intracellular KCl)[4].
- Archaea as surrogates of eukaryotes, using archaeal proteins (particularly from hyperthermophiles) for the purpose of determining the structural properties or function of proteins while avoiding the relative instability of eukaryotic proteins and complexity of eukaryotic systems. A good example of the success of this approach is the solving of the crystal structure of the archaeal RNA polymerase, which is structurally and functionally similar to RNA polymerase II from the Eukarya[149].
- Biotechnological efforts aimed at exploiting cellular components with intrinsic properties that are reflective of extreme conditions. This is well illustrated by the genome-sequencing programme for the hyperthermophilic methanogen *Methanocaldococcus jannaschii* (formerly known as *Methanococcus jannaschii*)[58], which was anticipated to be useful for generating renewable energy and synthetic chemicals, as well as a vehicle for the concentration and clean-up of toxic wastes. The value of learning about the mechanisms of adaptation and the novel properties of archaea for defining biotechnological exploitation has been described for a range of archaea and extremes[150–153].
- Astrobiological studies that use knowledge of extremophiles to guide exploration for past or present extraterrestrial life and possibly for use in terraforming[154–157].
- Origin-of-life studies that reflect on the fact that hyperthermophiles have deep branching points and short branch lengths in the universal tree of life, consistent with early lifeforms having a hyperthermophilic, chemolithotrophic lifestyle (see, for example, REF. 158). Studies extend to considerations about the emergence of the three domains from prebiotic conditions, including the hypothesis that life evolved in microcompartments within structured hydrothermal vents in which FeS precipitates at temperatures suitable for biological life[159]. The carbonate chimneys of the Lost City Hydrothermal Field serve as a useful model for this view of life's evolution[160]. This is a contentious topic, with archaea often taking centre stage in discussions (see, for example, REF. 161).

Thermoacidophile
An organism that requires high temperatures (typically greater than 60 °C) and a low pH (typically less than pH 3) for growth.

Heterotroph
An organism that uses organic compounds as nutrients to produce energy for growth.

Phototrophic
Pertaining to the growth of an organism: able to use sunlight to generate energy for growth.

Hyperthermophile
An organism that requires extremely high temperatures (typically greater than 80 °C) for growth.

Autotroph
An organism that can grow on carbon dioxide as a sole source of carbon.

By studying the Archaea, numerous oddities have also been discovered. For example, the first square-shaped free-living microorganism discovered was a haloarchaeon that was found in 1980 (REF. 7). However, the conditions that are suitable for isolating and cultivating this square haloarchaeon, *Haloquadratum walsbyi*, were not identified until 2004 (REF. 8). This in itself is somewhat ironic, given that the ability to grow haloarchaea aerobically on solid medium makes them the easiest phenotypic group of archaea to grow and manipulate[9]. Perhaps the most striking conundrum is the lack of any known archaeal pathogen for any animal or plant[10]. No single reason has been described that can account for this observation. Methanogens exist in the human gastrointestinal tract at concentrations of up to 10^{10} per gram of stool sample, and they have been associated with colorectal and periodontal disease[10]. However, other than an association with a disease state[11], no member of the Archaea has been shown (for example, by Koch's postulates) to be the primary cause of an animal or plant disease. Perhaps by providing an improved understanding of archaeal diversity and of the archaeal interactions with other microorganisms (including viruses and phages) and their hosts, metagenome programmes such as the US National Institutes of Health Metagenomics Analyses of the Oral Microbiome and the human gut[12] and bovine rumen[13] metagenome projects may in the future provide clues about the capacity of archaea to be pathogens.

This Review reflects on how far and how rapidly the field of biology has grown through studies of the Archaea. There is a wealth of discovery that, when connected across the years and decades through a timeline of the third domain, reveals important scientific themes, insightful individuals and a strong culture of acceptance of work outside the box (TIMELINE). The limited examples chosen here are an attempt to illustrate the essential foundation studies that defined the existence and subsequent phylogeny of the Archaea, while also covering select seminal advances from studies of the methanogens, haloarchaea and thermoacidophiles that helped to put the Archaea on the biological map. This Review attempts to cover a balance of historical precedent and contemporary advancement, to highlight the part that technological development has played in advancing the field and to cite emerging fields of importance (such as archaea in the cold biosphere).

Identification of the Archaea as the third domain

> *"The time will come I believe, though I shall not live to see it, when we shall have very fairly true genealogical trees of each great kingdom of nature."*
>
> Charles Darwin, letter 2143 to Thomas Huxley (26 Sep 1857)[14].

The discovery of the Archaea as a distinct domain of life arose through studies led by Carl Woese in the late 1960s, which aimed to determine how cellular life had evolved to the forms presently existing on Earth[15,16]. At this time in biology, evolutionary studies tended to focus on metazoan life, and the macromolecules used for phylogenetic evaluation were proteins, such as haemoglobins[17,18]. Instead, Woese investigated whether nucleic-acid markers (more specifically, ribosomal RNA (rRNA) sequences) could be used to define phylogenetic relationships, with the aim of constructing a universal tree of life — one that incorporated all cellular life[19,20].

Although he wanted to examine the evolution of bacteria, rather than attempt to define taxonomic subgroups based on a preconception that bacteria were phylogenetically cohesive (as they were classified at the time), Woese focused on learning about the evolution of life as a whole[21]. By choosing an analytical approach to evaluate evolution across all species, a new view of life emerged[20] that recognized that all known cellular life contained functionally and structurally conserved small-subunit rRNA (SSU rRNA; 16S rRNA in the Bacteria and the Archaea) and was therefore likely to have evolved from a common ancestor. Strikingly, the resulting universal tree of life described three distinct taxonomic divisions ('archaebacteria', 'eubacteria' and 'urkaryote')[20], and in 1990 these were formally proposed as the three domains, 'Archaea', 'Bacteria' and 'Eucarya'[22]. Until

Table 1 | **Selected traits of the Bacteria, Archaea and Eukarya***

Trait	Bacteria	Archaea	Eukarya
Carbon linkage of lipids	Ester	Ether	Ester
Phosphate backbone of lipids	Glycerol-3-phosphate	Glycerol-1-phosphate	Glycerol-3-phosphate
Metabolism	Bacterial	Bacterial-like	Eukaryotic
Core transcription apparatus	Bacterial	Eukaryotic-like	Eukaryotic
Translation elongation factors	Bacterial	Eukaryotic-like	Eukaryotic
Nucleus	No	No	Yes
Organelles	No	No	Yes
Methanogenesis	No	Yes	No
Pathogens	Yes	No	Yes

*The occurrence of a trait is for the majority of cases (not necessarily absolute) — for example, haloarchaea often possess organelles in the form of gas vesicles, but most archaea possess no organelles.

Woese determined the first SSU rRNA sequence data (for a methanogen), archaea — which, like bacteria, lack a nucleus — were 'simply' considered to be bacteria; the discovery was so surprising, it even stunned Woese[23].

The implications of these findings stimulated fevered debate concordant with the paradigm shift that they represented for all biology — the revelation being that a fundamental taxonomic division of life (archaeabacteria) existed that had evolved as a distinct lineage from all other cellular life and that was recognizable by molecular signature (SSU rRNA) alone, independently of the ecology, cellular morphology or growth characteristics of the life form. However, one notable feature of the discovery of the Archaea as a separate domain is that although this view was challenged (for example, see REFS 24,25), debate soon shifted to the nature of the relationships between domains (such as whether the Archaea are more related to the Bacteria or the Eukarya), to the subtleties of intradomain branching and to the coherence of the proposed subdivision kingdoms. In effect, even the strongest doubters came to terms with the fact that it was only prejudice that supported their view.

Nonetheless, on reading testaments about the challenges of the time[23,26–28], it is clear that interpreting the universal tree of life based on SSU rRNA sequences was contentious. The field moved forwards in part due to phylogenetic studies that used conserved protein sequences. In particular, analyses of paralogous pairs of genes (for example, those encoding translation elongation factors)[29,30] not only described the topology of the universal tree of life, but also provided a means of rationalizing the branch points between the domains at the root of the tree. Such phylogenetic analyses presented a similar view of the universal tree of life to that obtained by the analysis of SSU rRNA, but the advent of whole-genome sequencing in the mid 1990s provided a powerful new way to consider evolution. In essence, genomics provided clarification about the evolutionary path and lineage coherence of the Archaea, as well as supplying a gene inventory describing the functional characteristics of the Archaea relative to the Eukarya and the Bacteria (BOX 2).

As a result, and particularly in view of 'the shaking the universal tree of life has taken', it is more than fair to say that the efforts initiated and continued by Woese and colleagues (see, for example, REFS 31,32) have made a profound impact on the scientific community by introducing the existence of the three domains and, hence, changing our understanding of the evolution of cellular life on Earth. It is therefore with the benefit of hindsight, and of being witness to the seminal advances in discovering and defining the Archaea, that it is possible to reflect on some of the interesting findings about the third domain.

The rapid and continued growth in the publication output over the past 60 years (18,452 entries in PubMed for 'archaea'; accessed 6 August 2010) and the growth in the amount of DNA sequence data available (216,240 non-redundant reads for archaea in the latest release of RefSeq) provide a good perception of the healthy interest in studies of the Archaea (FIG. 1).

Lessons from aminoacyl-tRNA synthetases

The core components of the translation apparatus (rRNAs, ribosomal proteins, translation initiation factors and translation elongation factors) have highly conserved sequences across all domains of life, hence the use of SSU rRNA sequences for inferring the universal tree of life. Aminoacyl-tRNA synthetases (aaRSs) catalyse the formation of cognate aminoacyl-tRNAs and are therefore essential for the translation process. However, unlike the core components that must work cooperatively in very precise ways, individual aaRSs interact with only a small subset of the tRNAs and essentially function in isolation in the cell[33]. Mirroring the selection pressures on these different parts of the translation apparatus, the core components produce a common universal tree of life showing the accepted view of the three domains, whereas aaRSs tend to show a different topology consistent with lateral gene transfer (LGT), including possible transfer from extinct lineages[33]. Rationalizing the evolution of the genetic code, and the phylogenies of the aaRSs and the core components of the translation apparatus, has been a critical step for developing views about the evolution of life, particularly those concerning the roles of vertical versus lateral transmission of genes and the temporal considerations of when particular evolutionary events occurred[31–37].

Small-subunit rRNA
The ribosome is the core biological machine of the translation apparatus and is essential for converting the genetic code described in DNA and mRNA into protein. Ribosomal RNA (rRNA) is the RNA component of the ribosome and forms two subunits, the small subunit (SSU) and the large subunit. SSU rRNA is highly conserved in all cellular forms of life and is commonly used for describing the phylogeny of organisms.

Lateral gene transfer
Horizontal transfer of genes between unrelated species, as opposed to vertical inheritance within a species.

Timeline | **Milestones in 130 years of research into the Archaea**

Year	Milestone
1880	The first member of the Archaea is described — an extreme halophile[76,77]
1936	Methanogens are isolated[49–51]
1962	Ether-linked lipids are discovered in haloarchaea[78]
1970	Carl Woese investigates the evolution of cells, focusing on the translation mechanism[15]; ether-linked lipids are found in a *Thermoplasma* sp.[122,123]
1971	Bacteriorhodopsin is discovered in haloarchaea[81]
1972	Ether-linked lipids are detected in a *Sulfolobus* sp.[124,125]
1974	An archaeal virus is discovered in a haloarchaeon[97]
1975	The high-resolution electron microscopy structure of bacteriorhodopsin is obtained[83]
1977	SSU rRNA sequences reveal the identity of the third domain[20]
1980	SSU rRNA sequences reveal that there are two distinct clades (Crenarchaeota and Euryarchaeota) within the third domain[80]; a square-shaped haloarchaeon is discovered[7]
1982	Growth at >100 °C is seen for a *Pyrodictium* sp.[126]
1984	Archaea are found to fix nitrogen (methanogens)[53,54]
1987	Natural gene transfer is discovered in a methanogen [114]
1988	An antarctic haloarchaeon is isolated[136]
1989	Natural gene transfer is seen in haloarchaea[111]
1990	The third domain, Archaea, is formally proposed[22]
1992	Molecular signatures are described for marine archaea[132,133]; an antarctic methanogen is isolated[137]
1996	The first complete genome sequence is obtained, of a methanogen[58]

SSU rRNA, small-subunit ribosomal RNA.

Bootstrap value
A computationally derived measure of confidence about tree topology: the closer the bootstrap value is to 100, the more confidence we can have in the topology of the tree.

Monophyletic
Pertaining to a natural taxonomic group or clade: consisting of individuals that share a common ancestor.

The archaeal phylogenetic tree

By bypassing the need to obtain axenic cultures of microorganisms and applying SSU rRNA gene sequence analysis to environmental samples, Norman Pace demonstrated that an enormous diversity of previously unknown archaeal (and bacterial) species existed in the environment[28]. This approach expanded data sets describing the presence and type of archaea in the biosphere[38,39]. SSU rRNA gene sequences remain the accepted standard for describing the phylogeny of the Archaea[39]. Since the early studies, it was recognized that subdivisions within the Archaea could be defined, and in the original description of the three domains, the kingdoms Crenarchaeota and Euryarchaeota were described[22]. Since this time, new kingdoms have been proposed based on the discovery of new isolates or new SSU rRNA sequences, including the Nanoarchaeota[40], Korarchaeota[41] and, most recently, Thaumarchaeota[42]. Interestingly, the Nanoarchaeota, as represented by *Nanoarchaeum equitans*, were originally identified by microscopy but could not be classified as part of any domain, because the SSU rRNA probes were not sufficiently similar to be able to hybridize[40]. It was only after shotgun sequencing of the genome and subsequent identification of the 16S rRNA gene that *N. equitans* could be classified as a member of the Archaea, and a new kingdom was proposed[40]. With the SSU rRNA sequence known, *N. equitans* and other species were subsequently identified in many other locations and environments around the globe[43].

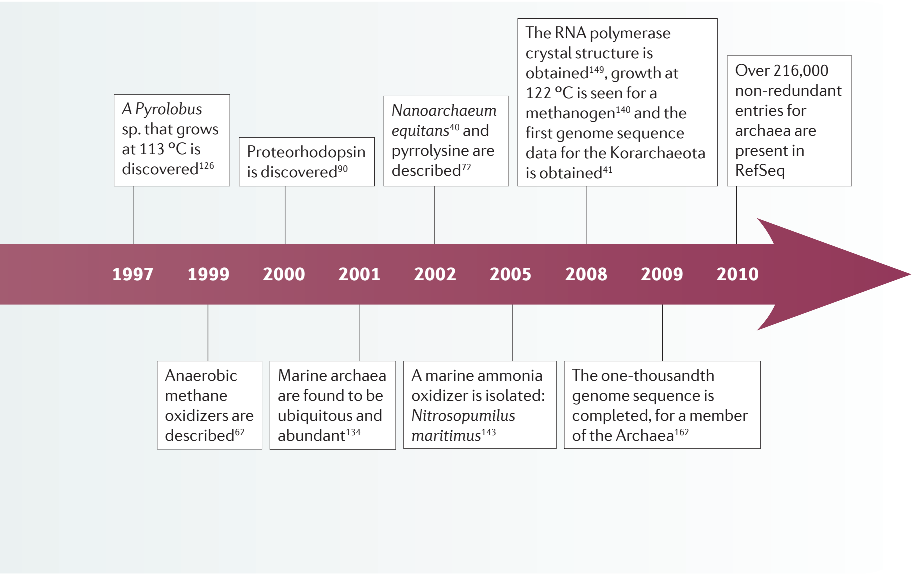
As the SSU rRNA database grows, the archaeal phylogenetic tree will continue to be re-evaluated. Chuck Robertson has generated a new, statistically supported overview of the archaeal tree using maximum likelihood for tree construction and collapsing branches with bootstrap values of less than 60% (FIG. 2). The tree reveals a monophyletic clade for the Crenarchaeota that incorporates both the 'Korarchaeota' and the ammonia oxidation clade representing the proposed 'Thaumarchaeota'. The Euryarchaeota clade consists of nine subclades that include the 'Nanoarchaeota', haloarchaea, the order Thermoplasmatales and several groups of methanogens, as well as sequences for diverse uncultivated archaea. The tree is similar to that reported in 2005 (REF. 44) and to that presented by Pace using sequences available in 2008 (REF. 39). As a broad overview, the tree describes at least two evolutionary paths with the basal radiation still to be resolved (and, no doubt, pondered and debated).

From methanogens and methane to oxygen

Methane has been used as a flammable gas for at least 3,000 years, and in the 1770s Alessandro Volta demonstrated that methane trapped in marsh sediments could be released, collected and combusted[45]. However it was not until 1906 that N. L. Sohngen — working in Martinus Beijerinck's laboratory, where the ability of rhizospere bacteria to fix nitrogen was studied — described a methane-using bacterium (*Methylomonas methanica*) and methanogenic 'bacteria' (REFS 46–48). Subsequent studies on methanogens were carried out throughout the 1930s by Cornelius Van Neil and Horace Barker[49–50], and in

Box 2 | **Role of sequencing in archaeal phylogeny**

It is enlightening to reflect on just how important technological advances in DNA sequencing were for influencing the field of archaeal biology. In the same way that nucleic acid-sequencing technologies provided the first and subsequent wealth of small-subunit ribosomal RNA gene sequence data (both for individual isolates and for surveys of uncultivated communities[38]), shotgun genome sequencing has provided a quantum increase in the available sequence information for individual microorganisms (the one-thousandth genome, for *Methanocaldococcus vulcanius* str. M7, was completed in October 2009 (REF. 162)), enabling phylogenomic evaluation of these organisms[6,163]. These analyses have generated a new level of understanding about genome evolution, in particular highlighting the role of lateral gene transfer in shaping existing genomes[31]. However, although genomic studies (and, more recently, metagenomic studies of whole-environmental samples) have identified issues pertaining to defining a species, and have even raised questions about the concept of a species[164], they have nonetheless consolidated the original view of the three domains of life[39,165]. Moreover, they have demonstrated conclusively that the Archaea are neither types of bacteria (that is, 'archaebacteria') nor sensibly considered 'prokaryotes' (REFS 161,166).



1940 Barker described one of the first axenic cultures of a methanogen[51]. Methanogens were subsequently found to have novel lipids[52] and were the crucial subjects in the initial SSU rRNA sequencing studies, carried out in 1976, that ushered in the discovery of the Archaea as the third domain[19]. The ability of several archaea to fix nitrogen was also first discovered in methanogens[53,54].

The microbiology and biochemistry of methanogenesis have been actively investigated since the 1970s, and important discoveries have been made about the unique enzymes and cofactors involved (such as coenzymes F_{420}, F_{430}, B and M, as well as methanofuran, tetrahydromethanopterin and their associated biosynthesis pathways[45,55]). Studies in this field, which continue today (see, for example, REFS 56,57), were challenging owing to the need for strict anaerobic growth conditions and the requirements of the enzymes for unusual elements (for example, Co, W, Mo, Ni). The 1990s saw the advent of shotgun genomics, including the third complete genome sequence — that of *Methanocaldococcus jannaschii* (formerly *Methanococcus jannaschii*)[58] — soon followed by proteomic and transcriptomic studies of global gene expression. These studies provided a new level of insight into the mechanisms of cell function, adaptation and evolution (such as gene expression, metabolism and mobile genetic elements). Advances in recombinant DNA techniques ultimately enabled integrated ecological studies that incorporated fluorescence *in situ* hybridization, sequencing of marker genes (such as the genes encoding SSU rRNA and methyl-CoM reductase I subunit alpha (McrA)), metagenomics and a range of geochemical data (for example, stable-isotope probing (SIP) to examine the interconversion of $^{14}CO_2$–$^{14}CH_4$) to derive greatly expanded views of the biological cycling of methane and associated one-carbon compounds. It is noteworthy that the ways in which studies of methanogens progressed throughout these decades were echoed in studies of other archaea; in essence, studies of the Archaea were tied to, and often drove the development of, technological advances in the biosciences.

Reverse methanogenesis
The methanogenesis pathway functioning in reverse to consume methane and produce cellular carbon and energy; this process leads to the anaerobic oxidation of methane.

Since the discovery of methanogens and their capacity to produce methane, it has been determined that biological methane constitutes ~85% of global methane levels. Methane is an important greenhouse gas, as it has 21 times the global-warming potential of CO_2, although it persists with a shorter half-life. As the simplest carbon compounds, with only one carbon atom, the atmospheric concentrations and ratios of the reduced (CH_4) and oxidized (CO_2) forms of methane have a major influence on the global carbon cycle and, ultimately, the survival of individual species.

Although methane can be oxidized aerobically by methanotrophic bacteria, the bulk of methane oxidation, which occurs primarily in the ocean, seems to take place anaerobically in association with sulphate reduction (for example, in sediments and methane hydrates on the sea floor)[59,60]. In attempting to understand the anaerobic oxidation of methane (AOM)[60,61], novel groups of methanotrophic euryarchaeota have been identified (for example, ANME-1, ANME-2 and ANME-3)[62–64], pathways for reverse methanogenesis have been investigated[65,66], links have been made between AOM and the global nitrogen cycle through a discovered ability of consortia to fix nitrogen[67], and AOM has been inferred to have possibly played a part in the biological cycling of carbon in an ancient anoxic biosphere[68]. A striking example of serendipitous discovery comes from studies that first described the ANME-3 clade and the AOM linked to denitrification[69]. Subsequent research determined that denitrifying bacteria could carry out AOM in the absence of archaea[70]. This then led to the discovery that nitrite-driven AOM could be carried out by oxygenic bacteria[71]. Using metagenomics and SIP, denitrification was shown to lead to AOM via the conversion of NO to N_2 and O_2 (REF. 71). The study has interesting implications for the evolution of early life, as it indicates that some oxygen may have been able to be produced under otherwise anaerobic conditions.

Pyrrolysine: a new amino acid

Another important discovery that heralded from the methanogens was pyrrolysine[72]. This twenty-second amino acid (after selenocysteine as the twenty-first) is incorporated at a UAG amber stop codon in the methylamine methyltransferases of some methanogens in the order Methanosarcinales[72,73], and it is thought to enable the transfer of a methyl group from the catalytic site of the methyltransferase to the cobalt atom of an associated corrinoid protein[74]. Pyrrolysine has also been shown to be incorporated in a $tRNA^{His}$ guanylyltransferase of a methanogen[75]. At least 11 organisms encode pyrrolysine, including two unrelated species of non-methanogenic bacteria[75]. The evolution of the pyrrolysine system has contributed to the debate about the evolution of aaRSs and the role of LGT[35–37].

Haloarchaea: from viruses to bacteriorhodopsin

Haloarchaea are notable for several important discoveries. The first report of an archaeal isolate in the scientific literature seems to have been in 1880, when a species of *Halococcus*, referred to as *Sarcina morrhuae*, was described[76,77]. The characterizations of haloarchaeal

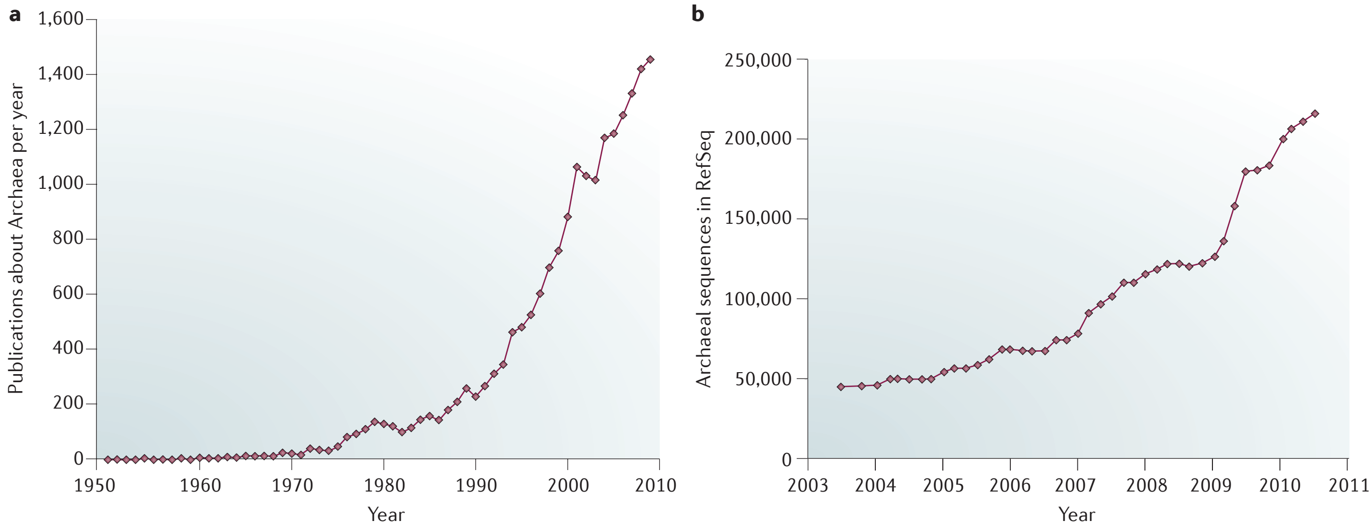


Figure 1 | **Trends in the growth of knowledge about the Archaea. a** | Annual (not cumulative) record of the term 'archaea' in PubMed (accessed 6 August 2010) for the period 1950 to 2009. **b** | The accumulation of DNA sequence data for archaea in RefSeq from June 2003 to July 2010. The taxonomic identification numbers for entries in RefSeq were used to categorize sequences according to the taxonomic divisions in the NCBI Taxonomy Database (Archaea, Bacteria, Eukaryota, Viroids, Viruses, Other and Unclassified), and the data for 'Archaea' were plotted against the dates of RefSeq releases 1 to 42.

lipids[78,79] and SSU rRNA[80] were particularly important for identifying the traits that distinguish the Archaea from the Bacteria (see below).

A great deal of interest in haloarchaea was generated from the discovery of the purple pigment bacteriorhodopsin in 1971 (REF. 81), followed by the demonstration that it functioned as a light-driven proton pump[82,83] to facilitate light-dependent chemotaxis[84]. Subsequent studies described the structural mechanisms for the translocation of protons in bacteriorhodopsin[85] and of chloride in the related halorhodopsin[86], a discovery that contributed greatly to discussions about the evolution of light-sensing mechanisms in rhodopsin systems[87].

The high-resolution electron microscopy structure of bacteriorhodopsin, obtained in 1975, was a landmark in membrane protein science, revealing seven transmembrane helices[83]. Sequence relationships show that bacteriorhodopsin is a member of a large family of membrane proteins that includes the G protein-coupled receptors (GPCRs). The first GPCR crystal structure, that of human β2 adrenergic receptor, was determined 32 years later[88]. Until then, the bacteriorhodopsin structure was the basis for models of the GPCRs, having a notable impact on human health, as the GPCRs represent ~30% of the targets of medicinal drugs[89].

From around 2000, the broad ecological importance of the light-harvesting pigments took on a new emphasis owing to the discovery of proteorhodopsins in marine Proteobacteria[90–92] and subsequently in marine members of the Euryarchaeota[93], the latter contributing further to the debate about the role of LGT in the evolution of these light-harvesting systems. The proteorhodopsin studies have highlighted the functional importance of photoheterotrophy[94,95], particularly in view of claims that proteorhodopsins have been identified in 13% to 80% of metagenomes for marine bacteria and archaea present in oceanic surface waters[96].

The first archaeal viruses were discovered in haloarchaea[97–99], and this heralded the discovery of viruses with a broad range of morphologies and viruses that were genetically unique[99,100]; these viruses were found to infect not only members of the order Halobacteriales, but also members of the hyperthermophilic order Sulfolobales. The discovery of the haloarchaeal viruses, early discoveries about the physical organization of haloarchaeal genomes[101], and subsequent genome and SSU rRNA gene analyses[102–104] have made important contributions to our current views about the evolution of natural microbial populations, including the roles that viruses can play[105].

Haloarchaea have also proved to be models for systems biology studies. Building on the genome sequence of *Halobacterium* sp. NRC-1 from 2000 (REF. 102), one of the first transcriptomic and proteomic studies of archaea was published in 2002 (REF. 106). Subsequent studies used immunoprecipitation to characterize transcription factor-binding sites[107] linked to expression data, and further analyses of operons[108] helped to develop predictive models[109] and visualization tools[110] that describe environmentally responsive gene-regulatory networks. These are pioneering advances that demonstrate approaches for empowering equivalent studies of other biological systems, and they provide valuable tools for future research.

Genetics in the Archaea

Historically, the development of gene transfer systems for archaea has not been rapid, in part owing to the lack of selectable markers (for example, the antibiotics that are used to select for bacteria target cellular properties that are absent in archaea) and to the difficulties in culturing many archaea on solid media (such as the inherent challenge of manipulating slow-growing anaerobes, and the inability of traditional agar to remain solid at

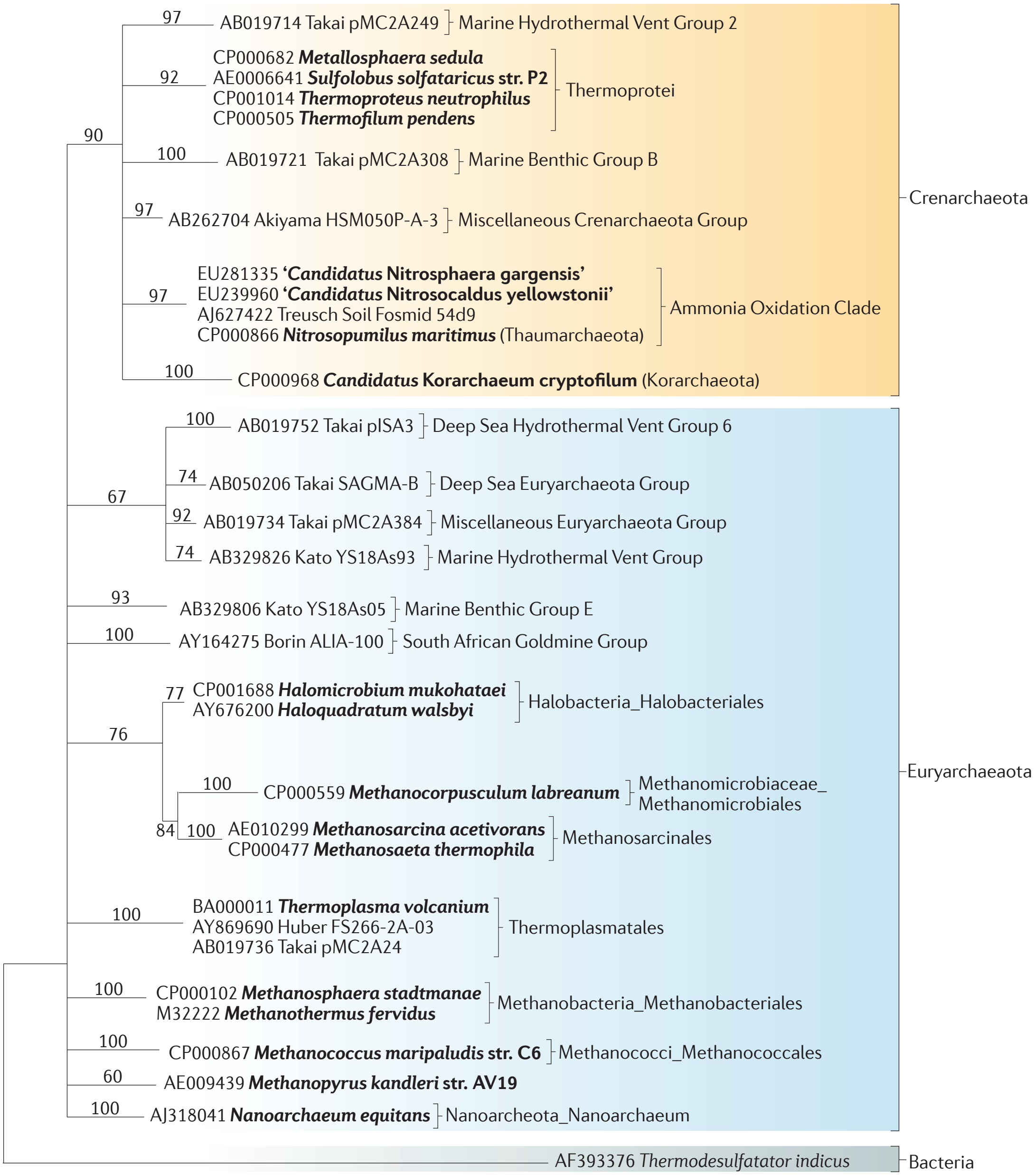


Figure 2 | **Phylogenetic tree of the Archaea.** The phylogenetic structure of the archaeal branch of the universal tree of life, derived from small-subunit (SSU) ribosomal RNA gene sequences. Much of the archaeal phylogenetic tree is formed from organisms that are known only from environmental sequences, which means that many of the clade names are informal, often based on names applied in the initial publication of the environmental sequence or sequences that founded the environmental clades. The clade names used here (except for the Ammonia Oxidation Clade) follow the names used in the Silva 102 guide tree[167]. At least one sequence is provided for each clade, specified by accession numbers at the ends of the branches. Species for which the genomic sequences are available are shown in bold. Paper author and clone names are provided for published environmental sequences. This tree is based on a subsample of all archaeal sequences in the Silva 102 'SSU REF' database. The subsample sequences were selected so that the archaeal sequence space was sampled at 12% to 15% intervals. One thousand maximum-likelihood bootstrap trees were cast with the RAxML software[168] from the subsampled sequences. The bootstrap results were annotated onto a high-scoring maximum-likelihood tree, and branches with <60% bootstrap support were collapsed to better reveal the large-scale, statistically supported phylogenetic structure. This tree suggests that the Euryarchaeota clade is not monophyletic.

temperatures suitable for growing hyperthermophiles). However, the discovery of a natural gene transfer system in haloarchaea[111] led to the development of genetic systems that provided the means for making important discoveries about the Archaea, such as the role of ubiquitin-like small archaeal modifier proteins in *Haloferax volcanii*[112]. Other important milestones in the use of genetics to study the Archaea include exploiting the ability of *Sulfolobus acidocaldarius* to undergo cell–cell mediated isogenic gene exchange[113], and developing methods of transformation using $CaCl_2$ for *Methanococcus voltae*[114] and *Thermococcus kodakaraensis*[115], polyethylene glycol for *H. volcanii*[116], liposomes for some members of the family Methanosarcinacea[117], and electroporation for viral transfection of *Sulfolobus solfataricus*[118]. These gene transfer systems now facilitate various recombinant DNA techniques using a broad range of selection systems to carry out overexpression and reporter fusion assays, and they also enable targeted and random mutagenesis in haloarchaea, methanogens and thermoacidophiles[119,120].

Hyperthermophiles

In the 1960s and early 1970s, in addition to studies of haloarchaea[79,121], analyses of lipid composition were carried out on thermoacidophiles of the genera *Thermoplasma*[122,123] and *Sulfolobus*[124,125] that were isolated by Thomas Brock. These studies identified the presence of tetra-ether lipids in the thermoacidophiles, at the time thought to be an archaeal trait possibly involved in thermal adaptation[123,125]. However, the SSU rRNA sequences revealed that the genus *Sulfolobus* is phylogenetically separate (in the kingdom Crenarchaeota) from a clade that includes the genus *Thermoplasma*, the methanogens and the haloarchaea (collectively, the kingdom known as the Euryarchaeota)[80]. In effect, these early SSU rRNA studies showed that although archaea could be conveniently grouped physiologically (into methanogens, extreme halophiles and thermoacidophiles), the thermoacidophiles represented at least two major lineages of the Archaea[80].

Subsequent studies of many diverse isolates and of the SSU rRNA sequences of environmental samples from aquatic and terrestrial hydrothermal regions have revealed the great diversity of hyperthermophilic members of the Crenarchaeota, as well as of species more broadly within the Euryarchaetoa[4]. Noting the wealth of axenic cultures of archaeal extremophiles available for study today[4,126,127], the achievements of Ralph Wolfe, Otto Kandler, Wolfram Zillig and Karl Stetter cannot be overstated. To achieve this success, they had to surmount the logistical challenges of sampling hydrothermal environments ranging from boiling sulphurous hot springs to deep-sea hydrothermal vents, to account for the fastidious growth requirements of previously uncultured organisms (which often involve culture vessels suitable for growth in complex artificial media that are typically acidic, anaerobic and/or under high hydrostatic pressure), and to overcome the inherent difficulties of isolating environmental microorganisms (which is a general problem, even for nonextreme environments). In essence, the field of archaeal biology owes a great deal of its success to individuals across the globe who have isolated diverse extremophilic species.

The intense interest of Stetter in hydrothermal systems led to the discovery of *N. equitans*, a symbiont, living within the unusually large periplasmic space of its archaeal host, *Ignicoccus hospitalis*[40]. *N. equitans* has a very small genome (0.49 Mb) that lacks many genes necessary for autonomy (for example, lipid synthesis genes)[128], and gene exchange seems to have occurred between *N. equitans* and *I. hospitalis* during the establishment of their apparent symbiosis[129]. The process of producing a minimalistic genome set[129] may have involved the localization of most of the *N. equitans* genomic information to the cytoplasm of *I. hospitalis* (forming the 1.3 Mb genome of *I. hospitalis*). The energy (ATP) generation apparatus of *I. hospitalis* moved to the periplasm, which possibly benefits the growth of *N. equitans*[130,131]. This system is unusual in the archaeal domain and provides an intriguing model for the evolution of eukaryotic cells, which exhibit similar architecture and compartmentalization of their cellular processes.

Chemolithoautotrophically
Pertaining to an organism: able to derive energy from a chemical reaction (chemotroph) using inorganic substrates as electron donors (lithotroph) and CO_2 as a carbon source (autotroph).

Archaea in the cold biosphere

In 1992, efforts led by Jed Fuhrman and Ed DeLong showed that species belonging to new clades of the Archaea were seemingly abundant in the marine environment[132,133]. By demonstrating not only the presence, but also the abundance of archaea in such a large, diverse environment, these findings had an important role in raising awareness in the broader scientific community, as well as with the public, about the contributions that archaea make to natural ecosystems.

With the finding that archaea are present at a depth of 500 m (REF. 132) and abundant at 4,750 m below the surface[134], archaea were shown to be important components of the deep, cold biosphere. Owing to the extent of the Earth that is represented by ocean depths and the ocean subsurface, as well as polar, permafrost, alpine and seasonally cold environments, the majority of Earth's biomass in fact resides at cold temperatures (≥5 °C). Within the cold biosphere, diverse archaea have been identified and linked to a broad range of microbial processes, including ammonia oxidation and methanogenesis[135]. Moreover, through the isolation of haloarchaea[136] and methanogens[137,138] from Antarctica, and in particular through studies on *Methanococcoides burtonii*[135,139], the molecular mechanisms of adaptation in these organisms are gradually being discovered. Interestingly, although some methanogens can grow at subzero temperatures, including in the permafrost, at least one species can grow at temperatures up to 122 °C[135,140], making these archaea the most thermally diverse type of organism known and particularly useful models for the study of adaptation to low and high temperatures.

Up until 2005, hyperthermophiles and thermophiles were the only members of the Crenarchaeota that could be grown axenically. However, in 1998, marine members of the Crenarchaeota were characterized from a sponge[141], shown to be abundant in the world's oceans[134] and predicted to be able to oxidize ammonia chemolithoautotrophically[92,142]. In 2005, David Stahl and colleagues isolated *Nitrosopumilus maritimus* from a temperate marine aquarium; this species can oxidize ammonia chemolithoautotrophically and was the first non-thermophilic isolate of the Crenarchaeota[143]. From genomic analysis, *N. maritimus* was found to be widespread in the Global Ocean Survey metagenome data, representing ~1% of reads, on average[144]. The capacity of *N. maritimus* to thrive autotrophically has been linked to the use of a hydroxypropionate–hydroxybutyrate cycle, which is expected to fix CO_2 more efficiently than the Calvin–Bassham–Benson cycle used by ammonia-oxidizing bacteria[144]. The finding that this pathway is used

Ammonia-oxidizing members of the Crenarchaeota
An archaeon with the ability to grow chemolithoautotrophically with near-stoichiometric conversion of ammonium cations (NH_4^+) to nitrite ions (NO_2^-) using carbonic acid (H_2CO_3) and ammonium (NH_4) as the sole sources of carbon and nitrogen, respectively.

by *N. maritimus* is consistent with the view that the six autotrophic CO_2 fixation pathways that have been identified in the Archaea reflect both phylogenetic lineage and ecological niche[145]. In parallel to the marine studies, investigations of terrestrial environments revealed not only that the Crenarchaeota are present in soil[146], but also that those species with an ability to oxidize ammonia are highly abundant[147]. Collectively, these studies have revealed that ammonia-oxidizing members of the Crenarchaeota play an important part in the global nitrogen cycle, particularly in low-nutrient, acidic and sulphide-containing environments[148].

Concluding remarks

Over the past 70 years, the process of establishing what the Archaea are and what they mean for life on Earth has been a spectacular journey. For those involved, it has often meant struggling to find an acceptable position for the Archaea within the universal tree of life and learning about their unusual abilities to thrive under all manner of extremes, while coming to terms with the finding that these organisms are ubiquitous and abundant and are driving key global microbial processes. This path has attracted 'archaeaologists' with great tenacity and drive, impressive creativity and dynamic thinking, and a capacity to expand a mental comfort zone to broker remarkable observations, often in view of reluctant acceptance. The quality of the minimal number of papers that are cited in this Review from the >18,542 that are available is indicative of the overall standard of science that the field has generated. The Review therefore offers a reflective glimpse into the role that some of the important 'movers and shakers' have made to the field, while providing a contemporary testament to the impact that the field has made — and will undoubtedly continue to make — to the future of biology.

1. Cavicchioli, R. (ed) *Archaea: Molecular and Cellular Biology* (American Society for Microbiology Press, Washington DC, 2007).
A strong collection of chapters from leaders in the field, including a testament about the discovery of the Archaea by Carl Woese.
2. Garrett, R. & Klenk, H.-P. *Archaea: Evolution, Physiology and Molecular Biology* (Blackwell, Oxford, UK, 2007).
3. Rothschild, L. J. & Mancinelli, R. L. Life in extreme environments. *Nature* **409**, 1092–1101 (2001).
4. Kletzin, A. in *Archaea: Molecular and Cellular Biology* (ed. Cavicchioli, R.) 14–92 (American Society for Microbiology Press, Washington DC, 2007).
A comprehensive overview of the microbiology, biochemistry and ecology of many types of archaea.
5. Makarova, K. S., Yutin, N., Bell, S. D. & Koonin, E. V. Evolution of diverse cell division and vesicle formation systems in Archaea. *Nature Rev. Microbiol.* **8**, 731–741 (2010).
6. Gribaldo, S., Poole, A. M., Daubin, V., Forterre, P. & Brochier-Armanet, C. The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse? *Nature Rev. Microbiol.* **8**, 743–752 (2010).
7. Walsby, A. E. A square bacterium. *Nature* **283**, 69–71 (1980).
8. Burns, D. G. *et al. Haloquadratum walsbyi* gen. nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain. *Int. J. Syst. Evol. Microbiol.* **57**, 387–392 (2007).
9. Soppa, J. From genomes to function: haloarchaea as model organisms. *Microbiology* **152**, 585–590 (2006).
10. Cavicchioli, R., Curmi, P. M. G., Saunders, N. & Thomas, T. Pathogenic archaea: do they exist? *Bioessays* **25**, 1119–1128 (2003).
A discussion about the concept of archaeal pathogens.
11. Lepp, P. W. *et al.* Methanogenic Archaea and human periodontal disease. *Proc. Natl Acad. Sci. USA* **101**, 6176–6181 (2004).
12. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
13. Brulc, J. M. *et al*. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc. Natl Acad. Sci. USA* **106**, 1948–1953 (2009).
14. Burkhardt, F. & Smith, S. (eds) *The Correspondence of Charles Darwin:1856–1857* Vol. 6 (Cambridge Univ. Press, Cambridge, UK, 1990).
15. Woese, C. R. in *Organization and Control in Prokaryotic and Eukaryotic Cells. 20th Symposium, Society for General Microbiology* (eds Charles, H. P. & Knight, B. C. J. G.) 39–54 (Cambridge Univ. Press, Cambridge, UK (1970).
16. Woese, C. R., Sogin, M. L. & Sutton, L. A. Procaryote phylogeny. I. Concerning the relatedness of *Aerobacter aerogenes* to *Escherichia coli*. *J. Mol. Evol.* **3**, 293–299 (1974).
17. Crick, F. H. C. The biological replication of macromolecules. *Symp. Soc. Exp. Biol.* **12**, 138–163 (1958).
18. Zuckerkandl, E. & Pauling, L. Molecules as documents of evolutionary history. *J. Theor. Biol.* **8**, 357–366 (1965).
19. Fox, G. E., Magrum, L. J., Balch, W. E., Wolfe, R. S. & Woese, C. R. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proc. Natl Acad. Sci. USA* **74**, 4537–4541 (1977).
A cornerstone of the discovery of the third domain: the first SSU rRNA sequences of methanogens.
20. Woese, C. R. & Fox, G. E. The phylogenetic structure of the procaryotic domain: the primary kingdoms. *Proc. Natl Acad. Sci. USA* **74**, 5088–5090 (1977).
The quintessential description of the third domain that started it all.
21. Woese, C. R. & Fox, G. E. The concept of cellular evolution. *J. Mol. Evol.* **10**, 1–6 (1977). **An important paper describing the importance of the genotype–phenotype relationship in the evolution of life.**
22. Woese, C. R., Kandler, O. & Wheelis, M. L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl Acad. Sci. USA* **87**, 4576–4579 (1990).
The proposal for the formal description of the three domains — a must-read for all educational purposes.
23. Woese, C. R. in *Archaea: Molecular and Cellular Biology* (ed. Cavicchioli, R.) 1–13 (American Society for Microbiology Press, Washington DC, 2007).
A provocative testament about the archaea movement.
24. Cavalier-Smith, T. The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int. J. Syst. Evol. Microbiol.* **52**, 7–76 (2002).
25. Simonson, A. B. *et al.* Decoding the genomic tree of life. *Proc. Natl Acad. Sci. USA* **102**, 6608–6613 (2005).
26. Sapp, J. The prokaryote-eukaryote dichotomy: meanings and mythology. *Microbiol. Mol. Biol. Rev.* **69**, 292–305 (2005).
An extensive and revealing review about the archaea movement.
27. Friend, T. *The Third Domain: the Untold Story of Archaea and the Future of Biotechnology*. (Joseph Henry, Washington DC, 2007).
An enjoyable, easy-to-read and insightful book about the history of the discovery of the Archaea.
28. Woese, C. R. & Goldenfeld, N. How the microbial world saved evolution from the Scylla of molecular biology and the Charybdis of the modern synthesis. *Microbiol. Mol. Biol. Rev.* **73**, 14–21 (2009).
A contemporary essay concerning how the evolutionary process has been studied in the twentieth century.
29. Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S. & Miyata, T. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl Acad. Sci. USA* **86**, 9355–9359 (1989).
30. Brown, J. R. & Doolittle, W. F. Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev.* **61**, 456–502 (1997).
An extensive review of the evolution of the universal tree of life, with particular emphasis on protein phylogeny.
31. Woese, C. R. On the evolution of cells. *Proc. Natl Acad. Sci. USA* **99**, 8742–8747 (2002).
32. Vetsigian, K., Woese, C. R. & Goldenfeld, N. Collective evolution and the genetic code. *Proc. Natl Acad. Sci. USA* **103**, 10696–10701 (2006).
A major advance in considering the role of Darwinian versus Lamarckian influences on the evolution of the genetic code, the translation process and cellular organization.
33. Woese, C. R., Olsen, G. J., Ibba, M. & Söll, D. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* **64**, 202–236 (2000).
A comprehensive review placing the evolution of aaRSs in the context of the rest of the translation apparatus.
34. Woese, C. R. A new biology for a new century. *Microbiol. Mol. Biol. Rev.* **68**, 173–186 (2004).
A reflection on the scientific process involved in revealing the Archaea to the scientific community.
35. Kavran, J. M. *et al.* Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc. Natl Acad. Sci. USA* **104**, 11268–11273 (2007).
36. Nozawa, K. *et al.* Pyrrolysyl-tRNA synthetase-tRNA(Pyl) structure reveals the molecular basis of orthogonality. *Nature* **457**, 1163–1167 (2009).
37. Fournier, G. P., Huang, J. & Gogarten, J. P. Horizontal gene transfer from extinct and extant lineages: biological innovation and the coral of life. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**, 2229–2239 (2009).
A provocative view of the evolution of life (in particular, concerning what might be inferred about extinct lineages).
38. Pace, N. R. A molecular view of microbial diversity and the biosphere. *Science* **276**, 734–740 (1997).
An insightful review capturing the excitement emerging from the use of DNA sequence data to reveal the extent of the previously unknown phylogenetic and functional diversity in the microbial world.
39. Pace, N. R. Mapping the tree of life: progress and prospects. *Microbiol. Mol. Biol. Rev.* **73**, 565–576 (2009).
An up-to-date review about tree construction using SSU rRNA sequences and inferences about the universal tree of life.
40. Huber, H. *et al.* A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**, 63–67 (2002).
Insight into an interesting new lifestyle as well as a new division within the Archaea.

41. Elkins, J. G. *et al.*, A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl Acad. Sci. USA* **105**, 8102–8107 (2008).
    Insight into an interesting new division within the Archaea.
42. Brochier-Armanet, C., Boussau, B., Gribaldo, S. & Forterre, P. Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Rev. Microbiol.* **6**, 245–252 (2008).
43. Casanueva, A. *et al.* Nanoarchaeal 16S rRNA gene sequences are widely dispersed in hyperthermophilic and mesophilic halophilic environments. *Extremophiles* **12**, 651–656 (2008).
44. Robertson, C. E., Harris, J. K., Spear, J. R. & Pace, N. R. Phylogenetic diversity and ecology of environmental Archaea. *Curr. Opin. Microbiol.* **8**, 638–642 (2005).
45. Ferry, J. G. *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics*. *Chapman and Hall Microbiology Series* (Chapman and Hall Inc., New York, 1993).
    A compendium describing essentially everything that was known about methanogens in 1993.
46. Sohngen, N. L. Ueber bakterien, welche methan als kohlenstoffnahrung and energiequelle gebrauchen. *Zentralbl. Bakteriol. Parasitik. Abt. I.* **15**, 513–517 (1906) (in German).
47. Wolfe, R. S. The sixth A. J. Kluyver Memorial Lecture. Methanogens: a surprising microbial group. *Antonie Van Leeuwenhoek* **45**, 353–364 (1979).
48. Dalton, H. The Leeuwenhoek Lecture 2000: the natural and unnatural history of methane-oxidizing bacteria. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **360**, 1207–1222 (2005).
49. Barker, H. A. Studies upon the methane-producing bacteria. *Arch. Mikrobiol.* **7**, 420–438 (1936).
50. Kluyver, A. J. & van Niel, C. B. Prospects for a natural system of classification of bacteria. *Zentralbl. Bakterol. Parasitenkd. Infectinskr. Hyg. Abt. II* **94**, 369–403 (1936).
51. Barker, H. A. Studies upon the methane fermentation. IV. The isolation and culture of *Methanobacterium omelianskii*. *Antonie van Leeuwenhoek J. Microbiol. Serol.* **6**, 201–220 (1940).
52. Kandler, O. & Hippe, H. Lack of peptidoglycan in the cell walls of *Methanosarcina barkeri*. *Arch. Microbiol.* **113**, 57–60 (1977).
53. Murray, P. A. & Zinder, S. Z. Nitrogen fixation by a methanogenic archaebacterium. *Nature* **312**, 284–286 (1984).
54. Belay, N., Sparkling, R. & Daniels, L. Dintrogen fixation by a thermophilic methanogenic bacterium. *Nature* **312**, 286–288 (1984).
55. DiMarco, A. A., Bobik, T. A. & Wolfe, R. S. Unusual coenzymes of methanogenesis. *Annu. Rev. Biochem.* **59**, 355–394 (1990).
56. Ferry, J. G. How to make a living by exhaling methane. *Annu. Rev. Microbiol.* **64**, 453–473 (2010).
57. Scheller, S., Goenrich, M., Boecher, R., Thauer, R. K. & Jaun, B. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* **465**, 606–608 (2010).
58. Bult, C. A. *et al.* Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058 – 1073 (1996).
    A historical, eye-opening article documenting the first genome sequence of an archaeon, and the third complete genome sequence of an organism.
59. Orphan, V. J., House, C. H., Hinrichs, K.-U., McKeegan, K. D. & DeLong, E. F. Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. *Proc. Natl Acad. Sci. USA* **99**, 7663–7668 (2002).
60. Knittel, K. & Boetius, A. Anaerobic oxidation of methane: progress with an unknown process. *Annu. Rev. Microbiol.* **63**, 311–334 (2009).
    A comprehensive review of anaerobic methane oxidation.
61. Alperin, M. & Hoehler, T. The ongoing mystery of sea-floor methane. *Science* **329**, 288–289 (2010).
62. Hinrichs, K. U., Hayes, J. M., Sylva, S. P., Brewer, P. G. & DeLong, E. F. Methane-consuming archaebacteria in marine sediments. *Nature* **398**, 802–805 (1999).
    An important breakthrough in realizing a role for archaea in anaerobic methane oxidation.
63. Orphan, V. J., House, C. H., Hinrichs, K.-U., McKeegan, K. D. & DeLong, E. F. Methane-consuming Archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* **293**, 484–487 (2001).
64. Niemann, H. *et al.* Novel microbial communities of the Haakon Mosby mud volcano and their role as a methane sink. *Nature* **443**, 854–858 (2006).
65. Zehnder, A. J. & Brock, T. D. Methane formation and methane oxidation by methanogenic bacteria. *J. Bacteriol.* **137**, 420–432 (1979).
66. Hallam, S. J. *et al.* Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305**, 1457–1462 (2004).
    An illustration of how metagenomics can be used to tease out a previously unknown pathway when there is evidence that the process (in this case, anaerobic oxidation of methane) occurs.
67. Dekas, A. E., Poretsky, R. S. & Orphan, V. J. Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. *Science* **326**, 422–426 (2009).
68. Michaelis, W. *et al.* Microbial reefs in the Black Sea fueled by anaerobic oxidation of methane. *Science* **297**, 1013–1015 (2002).
69. Raghoebarsing, A. A. *et al.* A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature* **440**, 918–921 (2006).
70. Ettwig, K. F. *et al.* Denitrifying bacteria anaerobically oxidize methane in the absence of Archaea. *Environ. Microbiol.* **10**, 3164–3173 (2008).
71. Ettwig, K. F. *et al.* Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* **464**, 543–548 (2010).
    A lesson in what can be learned beyond the Archaea, from the Archaea.
72. Hao, B. *et al.* A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* **296**, 1462–1466 (2002).
    An important description of pyrrolysine and the mechanism of its incorporation into proteins.
73. Soares, J. A. *et al.* The residue mass of L-pyrrolysine in three distinct methylamine methyltransferases. *J. Biol. Chem.* **280**, 36962–36969 (2005).
74. Krzycki, J. A. Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. *Curr. Opin. Chem. Biol.* **8**, 484–491 (2004).
75. Heinemann, I. U. *et al.* The appearance of pyrrolysine in tRNA^{His} guanylyltransferase by neutral evolution. *Proc. Natl Acad. Sci. USA* **106**, 21103–21108 (2009).
    An interesting description of what a novel amino acid and its cognate aaRSs can explain about evolution.
76. Farlow, W. G. in *United States Commission of Fish and Fisheries Part IV. Report for The Commissioner 1878* 969–973 (Government Printing Office, Washington DC, 1880).
77. Kocur, M. & Hodgkiss, W. Taxonomic status of the genus *Halococcus* Schoop. *Int. J. Syst. Bacteriol.* **23**, 151–156 (1973).
78. Sehgal, S. N., Kates, M. & Gibbons, N. E. Lipids of *Halobacterium cutirubrum*. *Can. J. Biochem. Physiol.* **40**, 69–81 (1962).
79. Marshall, C. L. & Brown, A. D. The membrane lipids of *Halobacterium halobium*. *Biochem. J.* **110**, 441–448 (1968).
80. Fox, G. E. *et al.* The phylogeny of prokaryotes. *Science* **209**, 457–463 (1980).
81. Oesterhelt, D. & Stoeckenius, W. Rhodopsin-like protein from the purple membrane of *Halobacterium hatobium*. *Nature New Biol.* **233**, 149–152 (1971).
82. Oesterhelt, D. & Stoeckenius, W. Functions of a new photoreceptor membrane. *Proc. Natl Acad. Sci. USA* **70**, 2853–2857 (1973).
83. Henderson, R. & Unwin, P. N. T. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257**, 28–32 (1975).
    An important breakthrough, not only for research into light-harvesting proteins (bacteriorhodopsin), but also for structural biology studies of proteins that contain multiple transmembrane domains.
84. Spudich, J. L. & Bogomolni, R. A. Mechanism of colour discrimination by a bacterial sensory rhodopsin. *Nature* **312**, 509–513 (1984).
85. Lanyi, J. K. Bacteriorhodopsin as a model for proton pumps. *Nature* **375**, 461–463 (1995).
86. Kolbe, M., Besir, H., Essen, L. O. & Oesterhelt, D. Structure of the light-driven chloride pump halorhodopsin at 1.8 Å resolution. *Science* **288**, 1390–1396 (2000).
87. Spudich, J. L., Yang, C. S., Jung, K. H. & Spudich, E. N. Retinylidene proteins: structures and functions from Archaea to humans. *Annu. Rev. Cell. Dev. Biol.* **16**, 365–392 (2000).
88. Rasmussen, S. G. *et al.* Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383–387 (2007).
89. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nature Rev. Drug Discov.* **5**, 993–936 (2006).
90. Béjà, O. *et al.* Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**, 1902–1906 (2000).
    An important study describing a previously unknown phototrophic system for marine heterotrophs.
91. Béjà, O., Spudich, E. N., Spudich, J. L., Leclerc, M. & DeLong, E. F. Proteorhodopsin phototrophy in the ocean. *Nature* **411**, 786–789 (2001).
92. Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74 (2004).
93. Frigaard, N. U., Martinez, A., Mincer, T. J. & DeLong, E. F. Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* **439**, 847–850 (2006).
94. Moran, M. A. & Miller, W. L. Resourceful heterotrophs make the most of light in the coastal ocean. *Nature Rev. Microbiol.* **5**, 792–800 (2007).
95. Gómez-Consarnau, L. *et al.* Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation. *PLoS Biol.* **8**, e1000358 (2010).
96. DeLong, E. F. & Béjà, O. The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *PLoS Biol.* **8**, e1000359 (2010).
97. Torsvik, T. & Dundas, I. D. Bacteriophage of *Halobacterium salinarium*. *Nature* **248**, 680–681 (1974).
98. Dyall-Smith, M., Tang, S-L. & Bath, C. Haloarchaeal viruses: how diverse are they? *Res. Microbiol.* **154**, 309–313 (2003).
99. Prangishvili, D., Forterre, P. & Garrett, R. A. Viruses of the Archaea: a unifying view. *Nature Rev. Microbiol.* **4**, 837–848 (2006).
100. Ortmann, A. C., Wiedenheft, B., Douglas, T. & Young, M. Hot crenarchaeal viruses reveal deep evolutionary connections. *Nature Rev. Microbiol.* **4**, 520–528 (2006).
     This article, along with reference 99, reveals the diversity and extent of archaeal viruses.
101. Sapienza, C. & Doolittle, W. F. Unusual physical organization of the *Halobacterium* genome. *Nature* **295**, 384–389 (1982).
     Early insight into genome rearrangement that led to theories about the role of LGT in evolution.
102. Ng, W. V. *et al.* Genome sequence of *Halobacterium species* NRC-1. *Proc. Natl Acad. Sci. USA* **97**, 12176–12181 (2000).
103. Papke, R. T., Koenig, J. E., Rodrīguez-Valera, F. & Doolittle, W. F. Frequent recombination in a saltern population of *Halorubrum*. *Science* **306**, 1928–1929 (2004).
104. Papke, R. T. *et al.* Searching for species in haloarchaea. *Proc. Natl Acad. Sci. USA* **104**, 14092–14097 (2007).
105. Rodriguez-Valera, F. *et al.* Explaining microbial population genomics through phage predation. *Nature Rev. Microbiol.* **7**, 828–836 (2009).
106. Baliga, N. S. *et al.* Coordinate regulation of energy transduction modules in *Halobacterium* sp. analyzed by a global systems approach. *Proc. Natl Acad. Sci. USA* **99**, 14913–14918 (2002).
107. Facciotti, M. T. General transcription factor specified global gene regulation in archaea *Proc. Natl Acad. Sci. USA* **104**, 4630–4635 (2007).
108. Koide, T. *et al.* Prevalence of transcription promoters within archaeal operons and coding sequences. *Mol. Syst. Biol.* **5**, 285 (2009).
109. Bonneau, R. *et al.* A predictive model for transcriptional control of physiology in a free living cell. *Cell* **131**, 1354–1365 (2007).
110. Bare, J. C., Koide, T., Reiss, D. J., Tenenbaum, D. & Baliga, N. S. Integration and visualization of systems biology data in context of the genome. *BMC Bioinformatics* **11**, 382 (2010).
111. Rosenshine, I., Tchelet, R. & Mevarech, M. The mechanism of DNA transfer in the mating system of an archaebacterium. *Science* **245**, 1387–1389 (1989).
112. Humbard, M. A. *et al.* Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature* **463**, 54–60 (2010).
113. Grogan, D. W. Exchange of genetic markers at extremely high temperatures in the archaeon *Sulfolobus acidocaldarius*. *J. Bacteriol.* **178**, 3207–3211 (1996).
114. Bertani, G. & Baresi, L. Genetic transformation in the methanogen *Methanococcus voltae* PS. *J. Bacteriol.* **169**, 2730–2738 (1987).
115. Sato, T., Fukui, T., Atomi, H. & Imanaka, T. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.* **185**, 210–220 (2003).

116. Charlebois, R. L., Lam, W. L., Cline, S. W. & Doolittle, W. F. Characterization of pHV2 from *Halobacterium volcanii* and its use in demonstrating transformation of an archaebacterium. *Proc. Natl Acad. Sci. USA* **84**, 8530–8534 (1987).
117. Metcalf, W. W., Zhang, J. K., Apolinario, E., Sowers, K. R. & Wolfe, R. S. A genetic system for Archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors. *Proc. Natl Acad. Sci. USA* **94**, 2626–2631 (1997).
A striking example of the development of a novel, high-efficiency gene transfer system.
118. Schleper, C., Kubo, K. & Zillig, W. The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc. Natl Acad. Sci. USA* **89**, 7645–7649 (1992).
119. Allers, T. & Mevarech, M. Archaeal genetics — the third way. *Nature Rev. Genet.* **6**, 58–73 (2005).
A good summary of the genetic systems that have been developed for archaea.
120. Sowers, K. & Anderson, K. in *Archaea: Molecular and Cellular Biology* (ed. Cavicchioli, R.) 463–477 (American Society for Microbiology Press, Washington DC, 2007).
121. Seghal, S. N., Kates, M. & Gibbons, N. E. Lipids of *Halobacterium cutirubrum*. *Can. J. Biochem. Physiol.* **40**, 69–81 (1962).
122. Darland, G., Brock, T. D., Samsonoff, W. & Conti, S. F. A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science* **170**, 1416–1418 (1970).
The discovery of an unusual archaeon before the knowledge that a third domain existed.
123. Langworthy, T. A., Smith, M. E. & Mayberry, W. R. A new class of lipopolysaccharide from *Thermoplasma acidophilum*. *J. Bacteriol.* **119**, 106–116 (1974).
124. Brock, T. D., Brock, K. M., Belly, R. T. & Weiss, R. L. *Sulfolobus*: a new genus of sulphur oxidizing bacteria living at low pH and high temperature. *Arch. Microbiol.* **84**, 54–68 (1972).
125. Langworthy, T. A., Smith, M. E. & Mayberry, W. R. Long-chain glycerol diether and polyol dialkyl glycerol triether lipids of *Sulfolobus acidocaldarius*. *J. Bacteriol.* **112**, 1193–1200 (1972).
126. Stetter, K. O. History of discovery of the first hyperthermophiles. *Extremophiles* **10**, 357–262 (2006).
A brief, but informative, review describing how archaeal hyperthermophiles were discovered.
127. Stetter, K. O. Hyperthermophiles in the history of life. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361**, 1837–1842 (2006).
128. Waters, E. *et al.* The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc. Natl Acad. Sci. USA* **100**, 12984–12988 (2003).
129. Podar, M. *et al.* A genomic analysis of the archaeal system *Ignicoccus hospitalis-Nanoarchaeum equitans*. *Genome Biol.* **9**, R158 (2008).
130. Küper, U., Meyer, C., Müller, V., Rachel, R. & Huber, H. Energized outer membrane and spatial separation of metabolic processes in the hyperthermophilic Archaeon *Ignicoccus hospitalis*. *Proc. Natl Acad. Sci. USA* **107**, 3152–3126 (2010).
Insight into the unusual relationship between an archaeal host and archaeal symbiont or parasite, with possible implications for the evolution of a eukaryote-like cell type.
131. van Ooij, C. A road to eukaryotes? *Nature Rev. Microbiol.* **8**, 246–247 (2010).
132. Fuhrman, J. A., McCallum, K. & Davis, A. A. Novel major archaebacterial group from marine plankton. *Nature* **356**, 148–149 (1992).
133. DeLong, E. F. Archaea in coastal marine environments. *Proc. Natl Acad. Sci. USA* **89**, 5685–5689 (1992).
Along with reference 132, this paper describes the new insight that archaea are important members of marine environments.
134. Karner, M. B., DeLong, E. F. & Karl, D. M. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507–510 (2001).
The sheer abundance of archaea in the marine biosphere is described.
135. Cavicchioli, R. Cold adapted Archaea. *Nature Rev. Microbiol.* **4**, 331–343 (2006).
Insight into the diversity, functional roles and adaptive mechanisms of archaea in the cold biosphere.
136. Franzmann, P. D. *et al. Halobacterium lacusprofundi* sp. nov., a halophilic bacterium isolated from Deep Lake, Antarctica. *Syst. Appl. Microbiol.* **11**, 20–27 (1988).
137. Franzmann, P. D., Stringer, N., Ludwig, W., Conway de Macario, E. & Rohde, M. A methanogenic archaeon from Ace Lake, Antarctica: *Methanococcoides burtonii* sp. nov. *System. Appl. Microbiol.* **15**, 573–581 (1992).
138. Franzmann, P. D. *et al. Methanogenium frigidum* sp. nov., a psychrophilic, H_2-using methanogen from Ace Lake, Antarctica. *Int. J. Syst. Bacteriol.* **47**, 1068–1072 (1997).
139. Allen, M. *et al.* The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold-adaptation. *ISME J.* **3**, 1012–1035 (2009).
140. Takai, K. *et al.* Cell proliferation at 122 °C and isotopically heavy CH_4 production by a hyperthermophilic methanogen under high-pressure cultivation. *Proc. Natl Acad. Sci. USA* **105**, 10949–10954 (2008).
141. Schleper, C. *et al.* Genomic analysis reveals chromosomal variation in natural populations of the uncultured psychrophilic archaeon *Cenarchaeum symbiosum*. *J. Bacteriol.* **180**, 5003–5009 (1998).
142. Schleper, C., Jurgens, G. & Jonuscheit, M. Genomic studies of uncultivated archaea. *Nature Rev. Microbiol.* **3**, 479–488 (2005).
143. Könneke, M. *et al.* Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**, 543–546 (2005).
Revealing insight into an elusive phylogenetic and phenotypic type of archaea: chemolithoautrophic ammonia oxidizers.
144. Walker, C. B. *et al. Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine *Crenarchaea*. *Proc. Natl Acad. Sci. USA* **107**, 8818–8823 (2010).
145. Berg, I. A. *et al.* Autotrophic carbon fixation in archaea. *Nature Rev. Microbiol.* **8**, 447–460 (2008).
146. Buckley, D. H., Graber, J. R. & Schmidt, T. M. Phylogenetic analysis of nonthermophilic members of the kingdom *Crenarchaeota* and their diversity and abundance in soils. *Appl. Environ. Microbiol.* **64**, 4333–4339 (1998).
147. Leininger, S. *et al.* Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**, 806–809 (2006).
148. Erguder, T. H., Boon, N., Wittebolle, L., Marzorati, M. & Verstraete, W. Environmental factors shaping the ecological niches of ammonia-oxidizing archaea. *FEMS Microbiol. Rev.* **33**, 855–869 (2009).
149. Hirata, A., Klein, B. J. & Murakami, K. S. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**, 851–854 (2008).
A striking example of how visually and functionally similar key components of the information processing systems are in the Archaea and Eukarya.
150. Cavicchioli, R., Siddiqui, K. S., Andrews, D. & Sowers, K. R. Low-temperature extremophiles and their applications. *Curr. Opin. Biotech.* **13**, 253–261 (2002).
151. Schiraldi, C., Giuliano, M. & De Rosa, M. Perspectives on biotechnological applications of archaea. *Archaea* **1**, 75–86 (2002).
152. Auernik, K. S., Cooper, C. R. & Kelly, R. M. Life in hot acid: pathway analyses in extremely thermoacidophilic archaea. *Curr. Opin. Biotechnol.* **19**, 445–453 (2008).
153. Krishnan, L. & Sprott, G. D. Archaeosome adjuvants: immunological capabilities and mechanism(s) of action. *Vaccine* **26**, 2043–2055 (2008).
154. Cavicchioli, R. Extremophiles and the search for extra-terrestrial life. *Astrobiology* **2**, 281–292 (2002).
155. Tung, H. C., Bramall, N. E. & Price, P. B. Microbial origin of excess methane in glacial ice and implications for life on Mars. *Proc. Natl Acad. Sci. USA* **102**, 18292–18296 (2005).
156. Oremland, R. S. & Voytek, M. A. Acetylene as fast food: implications for development of life on anoxic primordial Earth and in the outer solar system. *Astrobiology* **8**, 45–58 (2008).
157. Oremland, R. S. NO connection with methane. *Nature* **464**, 500–501 (2010).
158. Stetter, K. O. Hyperthermophilic procaryotes. *FEMS Microbiol. Rev.* **18**, 149–158 (1996).
159. Koonin, E. V. & Martin, W. On the origin of genomes and cells within inorganic compartments. *Trends Genet.* **21**, 647–654 (2005).
160. Brazelton, W. J. *et al.* Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *Proc. Natl Acad. Sci. USA* **107**, 1612–1617 (2010).
161. Glansdorff, N., Xu, Y. & Labedan, B. The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol. Direct* **3**, 29 (2008).
162. (No authors listed). A sequence of changes. *Nature Rev. Microbiol.* **8**, 85 (2010).
163. Delsuc, F., Brinkmann, H. & Philippe, H. Phylogenomics and the reconstruction of the tree of life. *Nature Rev. Genet.* **6**, 361–375 (2005).
164. Doolittle, W. F. & Zhaxybayeva, O. On the origin of prokaryotic species. *Genome Res.* **19**, 744–756 (2009).
165. Theobald, D. L. A formal test of the theory of universal common ancestry. *Nature* **465**, 219–222 (2010).
166. Pace, N. R. Time for a change. *Nature* **441**, 289 (2006).
An important and axiomatic essay about learning from what we have learned.
167. Pruesse, E. *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**, 7188–7196 (2007).
168. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).

Acknowledgements

I dedicate this Review to Carl Woese, who pioneered, led and continues to inspire the field with his brilliance. I am indebted to C. Robertson, who generously constructed the phylogenetic tree, and M. DeMaere, who processed the RefSeq data. I also thank my many colleagues who informally and formally commented on the manuscript, in particular N. Pace, T. Kolesnikow, F. Lauro, H. Ertan and M. Dyall-Smith. This Review could not be exhaustive and I regret not being able to cite all of the relevant literature. Research in my laboratory is supported by the Australian Research Council.

Competing interests statement

The author declares no competing financial interests.

FURTHER INFORMATION

Ricardo Cavicchioli's homepage:
http://www.babs.unsw.edu.au/directory.php?personnelID=5
NCBI Taxonomy Datbase:
http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html
PubMed:
http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed
RefSeq: http://www.ncbi.nlm.nih.gov/projects/RefSeq

ALL LINKS ARE ACTIVE IN THE ONLINE PDF