

ARCHAEAL GENETICS — THE THIRD WAY

Thorsten Allers* and Moshe Mevarech[†]

Abstract | For decades, archaea were misclassified as bacteria because of their prokaryotic morphology. Molecular phylogeny eventually revealed that archaea, like bacteria and eukaryotes, are a fundamentally distinct domain of life. Genome analyses have confirmed that archaea share many features with eukaryotes, particularly in information processing, and therefore can serve as streamlined models for understanding eukaryotic biology. Biochemists and structural biologists have embraced the study of archaea but geneticists have been more wary, despite the fact that genetic techniques for archaea are quite sophisticated. It is time for geneticists to start asking fundamental questions about our distant relatives.

DOMAIN

The highest level of taxonomic division, comprising Archaea, Bacteria and Eukarya. In declining order, the other levels include: kingdom, phylum, class, order, family, genus and species.

MONOPHYLETIC

A natural taxonomic group consisting of species that share a common ancestor.

**Institute of Genetics, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, UK.*

[†]*Department of Molecular Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel.*

Correspondence to T.A. e-mail: thorsten.allers@nottingham.ac.uk
doi:10.1038/nrg1504

Ever since microbiology was established by Louis Pasteur and Robert Koch, scientists have wrestled with the problem of defining the phylogenetic relationships among bacteria. Classical taxonomy, which relies on cell morphology, physiology and pathogenicity, is useful for identifying specific microorganisms. However, it fails to establish meaningful evolutionary relationships that can be used to group species into higher taxonomic orders. Carl Woese's solution was to harness the newly emerging techniques of nucleic acid sequencing, and use small-subunit (SSU) ribosomal-RNA nucleotide sequence as a universal molecular chronometer (BOX 1). When he published his findings in 1977, Woese upset the taxonomic appellation by suggesting that prokaryotes are much more diverse than we had previously supposed, and that the phylogenetic structure of the prokaryotic DOMAIN should be reassessed¹.

What he found was that a group of anaerobic 'bacteria', which had been studied for years owing to their unique ability to generate methane, are not bacteria at all. There had been inklings that these microbes have some 'unbacterial' aspects, such as the presence of N-linked glycoproteins and a peculiar spectrum of antibiotic sensitivity. The rRNA phylogeny revealed that they are no more related to typical bacteria than they are to eukaryotes, and Woese therefore renamed this group of microorganisms Archaeobacteria¹. In a subsequent paper, he shortened the name to Archaea and suggested

that this domain be given equal footing with Bacteria and Eukarya². Unsurprisingly, this proposal ran into much resistance.

Despite numerous attempts to square the taxonomic circle, the three-domain organization has stood the test of time (BOX 1). We now recognize that the domain Archaea is home to many microbes that were previously misclassified as bacteria owing to their prokaryotic morphology. Archaea are clearly MONOPHYLETIC and their status is underpinned by unique features such as a distinctive cell membrane containing isoprene side chains that are ether-linked to glycerol³. The SSU rRNA tree also reveals several archaeal phyla, which have biological differences that underpin their taxonomic split. For example, Euryarchaeota contain histones that are strikingly similar to eukaryotic homologues, whereas Crenarchaeota use completely different DNA-binding proteins⁴. Further insights have come from genome-sequencing projects, which have shown that archaea are a chimaera of bacterial and eukaryotic features; their core metabolic functions resemble those of bacteria, whereas their information-processing functions are distinctly eukaryotic. One feature that seems to unite archaea is their ability to thrive in harsh and unusual environments (BOX 2); it is because these organisms are so well suited to conditions that might have existed on the early (Archaean) earth that Woese gave them their name. However, it

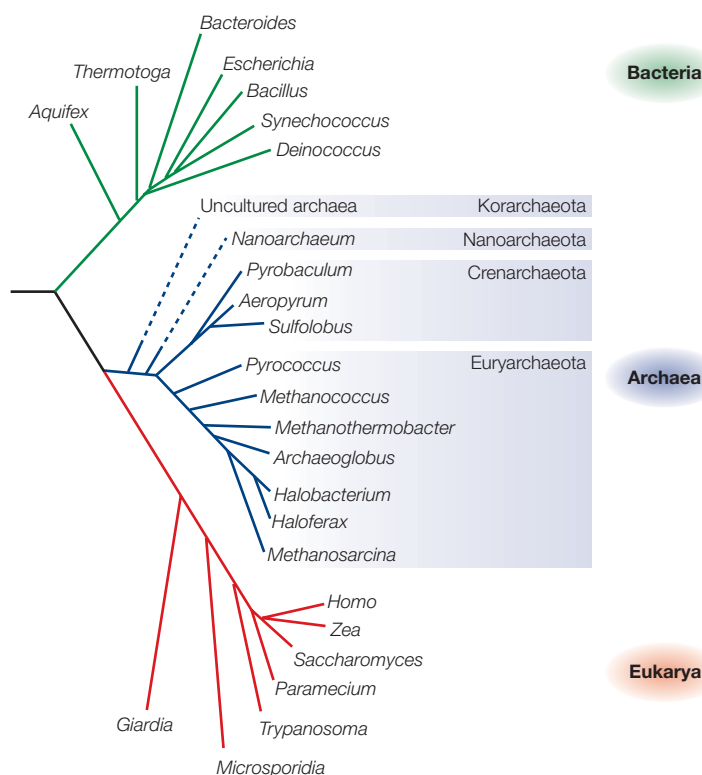
Box 1 | Archaeal taxonomy and the impact of lateral gene transfer

Archaeal taxonomy

The idea of using amino-acid sequences as a tool for molecular phylogeny was first proposed by Francis Crick in 1958 (REF 126), but had to wait until the molecular biology revolution of the 1970s, when Carl Woese revisited the problem of prokaryotic taxonomy¹. His choice of small-subunit (SSU) ribosomal RNA (rRNA) sequence as a molecular chronometer was visionary¹²⁷. As an essential component of all self-replicating organisms, rRNA shows remarkable sequence conservation; different parts of the molecule have varying rates of base substitution, allowing both coarse and fine-scale phylogenetic analyses. Furthermore, rRNA is abundant and easy to isolate, which has proved essential to CULTIVATION-INDEPENDENT STUDIES (BOX 2).

The rRNA tree reveals that the domain Archaea comprises several phyla. Euryarchaeota is the most diverse group, including all known methanogens and halophiles, in addition to thermophilic and

PSYCHROPHILIC species. Members of Crenarchaeota are renowned as hyperthermophiles (all temperature record-breaking species belong to this phylum), but include the psychrophile *Cenarchaeum symbiosum*. Of the remaining phyla, Nanoarchaeota has one known member (*Nanoarchaeum equitans*¹²⁸), and so far Korarchaeota are indicated only by environmental DNA sequences¹²⁹. Owing to the paucity of identified species, the positions of these phyla on the rRNA tree are uncertain (indicated by dashed branches).



Lateral gene transfer

Unfortunately, life is not as simple as the rRNA tree suggests. It is commonly assumed that only eukaryotes indulge in sex, whereas prokaryotes rely on vertical inheritance (coupled with prodigious reproductive powers) to meet new environmental challenges. In reality, prokaryotes are highly promiscuous, and the role of lateral gene transfer (LGT) as a driving force in prokaryotic evolution has been grossly underestimated¹². The fraction of 'foreign' (mostly bacterial) genes in archaea might be as high as 20–30% (REF 130). Most are the result of orthologous replacement or acquisition of a paralogous gene. Consequently, archaeal genomes resemble fine-scale genetic mosaics¹³¹. Occasionally, LGT can lead to a novel gain-of-function event. It has been suggested that the switch from an anaerobic to aerobic lifestyle by the (methanogenic) ancestor of haloarchaea was facilitated by LGT of respiratory-chain genes from bacteria¹³². Archaea, like bacteria, often group co-regulated genes in operons, and this arrangement would promote co-inheritance by LGT.

The degree to which archaea are 'polluted' by bacterial genes depends on the species examined, and LGT might account for some variation in archaeal genome sizes (see TABLE 1). For example, *Methanosarcina mazei* has a bloated genome of 4.10 Mb, 30% of which is bacterial in origin¹³³. Metabolically diverse methanogenic and halophilic archaea are most promiscuous, having motive, means and opportunity for LGT: their motive is to acquire genes for novel metabolic functions, the opportunity arises because methanogens and halophiles often cohabit with bacteria, and the means is their ability to engage in cell mating, as demonstrated for *Haloferax volcanii*⁵⁷.

LGT has serious consequences for taxonomy. Phylogenetic trees that are based on individual genes differ significantly from the rRNA tree¹³⁴, and the idea of a 'core' of never-exchanged genes has not passed muster¹³¹; such is the extent of LGT that no one gene is unique to either archaea or bacteria. Even the gold standard of rRNA can be corrupted when many copies of the small-subunit rRNA gene are present, as in haloarchaea¹³⁵. Nevertheless, as more archaeal genomes are sequenced and analysed, so a genomic signature is emerging that defines this domain in a holistic way¹⁴.

would be misleading to think that all archaea are EXTREMOPHILES. Recent environmental studies have shown that archaea are much more widespread than previously thought, and might constitute as much as 20% of the total biomass⁵.

As the domain Archaea has become widely accepted, researchers have turned to these fascinating microorganisms for answers to some of the most pressing questions in biology. Owing to the molecular features that they share with their more complex cousins,

CULTIVATION-INDEPENDENT STUDY

Method for determining environmental biodiversity without the need to obtain microbiologically pure cultures. This can be done by using sequences that are retrieved from environmental samples to construct a molecular phylogenetic survey (for example, through environmental genome-shotgun sequencing).

PSYCHROPHILE

An organism that can grow at permanently low temperatures; typically less than 10°C.

EXTREMOPHILE

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

archaea have served well as a streamlined model for eukaryotes, particularly in the field of DNA replication⁶ (discussed in detail later). On the other hand, the ability of archaea to thrive at high temperatures and salinity has endeared them to structural biologists, who have found thermostable and HALOPHILIC proteins to be almost indispensable. For example, the first crystal structure for a ribosome was obtained using the large ribosomal subunit from *Haloarcula marismortui*⁷. Archaeal enzymes are now routinely exploited as a source of high quality structural data that can be used to predict functional interactions in eukaryotic systems, particularly in fields related to information processing such as DNA repair⁸. Exploitation of the extremophilic features of archaea for biotechnology has yet to bear fruit⁹. Of the few examples in current use, those that are familiar to most scientists are the thermostable enzymes used for DNA amplification by PCR (for example, *Pfu* DNA polymerase from *Pyrococcus furiosus*). However, the potential for archaeal models in biotechnological applications is significant. For example, the ability of METHANOGENIC archaea to thrive under anaerobic conditions means that they are ideally suited for use in the bioremediation of anoxic sludge, such as marine coastal sediment. Furthermore, the methane they generate through anaerobic digestion of manure can be used as a fuel source. Finally, we should not underestimate the ecological impact of archaea; as methane is a powerful greenhouse gas, these organisms might be partly responsible for global warming.

With all the interesting aspects of archaea, why do so few scientists work on this domain of life? An important factor is the perceived lack of genetic systems. Archaea, or at least the ones that are cultivable, are renowned as extremophiles, and organisms that thrive in boiling acid are not conducive to routine genetic techniques. Furthermore, when faced with sophisticated model organisms such as *Escherichia coli* or *Saccharomyces cerevisiae*, it is understandable that scientists are reluctant to switch domain. Unfortunately, this means that numerous biochemical and structural studies on archaea are not being underpinned by *in vivo* data; *E. coli* and *S. cerevisiae* have been such successful models precisely because of the synergy of biochemistry and genetics. In this review, we aim to show that archaeal genetics is more advanced than is commonly believed. We will survey the current state of genetic systems — the difficulties in establishing genetic tools for archaea will be set out, followed by an update of methodologies in current use. We will also review the field of archaeal genomics, showing how data from genome projects have led to a reappraisal of the phylogenetic status of archaea, and how the striking similarity between archaeal and eukaryotic information-processing systems has stimulated a new generation of researchers to seek answers in the third domain.

Lessons from comparative archaeal genomics

Whereas the DNA sequencing revolution of the late 1970s gave birth to the Archaea domain, it was the genome sequencing revolution of the past decade that

enabled the archaeal concept to pass from adolescence to maturity. In retrospect, it was fortuitous that one of the first genome sequences to be published was of the methanogenic archaeon *Methanococcus jannaschii*¹⁰ (now renamed *Methanocaldococcus jannaschii*; see TABLE 1). The new discipline of genomics stimulated interest in these exotic microorganisms, as biologists found their genes of interest in a new context. Archaea proved to be a mosaic of molecular features, which are encoded by two different groups of genes: a lineage that codes for information processing which is eukaryotic in nature, and a lineage that codes for operational (housekeeping) functions with a bacterial aspect¹¹. This tidy division is not inviolable, for LATERAL GENE TRANSFER can lead to conflicting phylogenetic signals when any one archaeal species is examined in isolation¹² (BOX 1). According to the complexity hypothesis of Jain and Rivera¹³, informational genes are less prone to lateral transfer than operational genes, as the former are typically members of large complex systems¹³. However, the comparison of complete genome sequences has revealed that archaea are more than a sum of their (eukaryotic and bacterial) parts¹⁴. More than anything else, it is the high fraction (as much as 50%) of archaeal genes with no clear function that should prompt experimental biologists to reclaim the initiative in a post-genomic era.

Transcription and translation. The revelation that information-processing systems are similar in archaea and eukaryotes predates genome sequencing, and was noted in the 1980s by Wolfram Zillig and colleagues in studies of DNA-dependent RNA polymerases¹⁵. The core components of the archaeal and eukaryal enzyme (RNA polymerase II) are more closely related to each other than to the bacterial version, and the archaeal holoenzyme contains extra subunits that have counterparts in eukaryotes but not bacteria. Similar to eukaryote RNA polymerase II, the archaeal enzyme requires further basal factors for efficient promoter recognition, including TATA-box binding protein (TBP) and transcription factor B (TFB)¹⁶. Many archaea contain several homologues of TFB and/or TBP that might have distinct roles in transcription; for example, expression of one TFB-encoding gene is upregulated in response to heat shock of *Haloferax volcanii*¹⁷. Nevertheless, the basal transcription machinery in archaea is much simpler than in the eukaryotic system, and therefore is more amenable to analysis. This should lead to a better understanding of the many small subunits that are conserved between archaeal and eukaryal RNA polymerases.

Surprisingly, genome analysis has revealed that archaea also possess numerous homologues of bacterial transcription regulators¹⁸. This indicates that archaea might use a bacterial mode of transcriptional regulation, in which repressors bind at operator sites near the promoter and interfere directly with initiation. Such repressors have been studied *in vivo* in *Archaeoglobus fulgidus* and *Methanococcus maripaludis*^{19,20}. However, other systems are more reminiscent of eukaryotic regulation. In an elegant genetic analysis of gas-vesicle synthesis genes from *Halobacterium salinarum*, Felicitas Pfeifer

HALOPHILE

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

METHANOGEN

An anaerobic organism that generates methane by reduction of carbon dioxide, various one-carbon compounds or acetic acid.

LATERAL GENE TRANSFER

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

Box 2 | The ecological distribution of archaea – not just extremophiles



The archaea are notorious for inhabiting some of the most forbidding places on earth, and thrive under conditions that few bacteria and no eukaryotes would tolerate. Some of these habitats are illustrated in the figure. Methanogens are strict anaerobes that inhabit various anoxic habitats such as swamps (a) and sewage plants, and have a unique ability to generate energy by reducing carbon dioxide to form methane¹³⁶. This was first noted by the Italian physicist Alessandro Volta, who found that ‘combustible air’ is produced in lakes and bogs. Halophilic archaea grow readily in hypersaline ponds (b) and salt lakes such as the Dead Sea. Unlike bacteria, which maintain an osmotic balance using organic compatible solutes such as betaine, archaea accumulate inorganic salts (mainly K^+) in the cytoplasm. Although this allows them to grow in saturating NaCl solutions, it means that proteins of haloarchaea need to be adapted to function in high salt. ACIDOPHILES and ALKALIPHILES grow, as their names indicate, at extremes of pH; acidophiles are often thermophiles and a significant fraction of alkaliphiles are also halophiles. Thermophiles such as *Pyrolobus fumarii* are found near deep-sea hydrothermal vents (c), where temperatures exceed 100°C, and *Sulfolobus solfataricus* populates hot springs such as those in Yellowstone National Park (d). There are many molecular adaptations to thermophily, but the most striking is reverse gyrase, an enzyme that introduces positive supercoils in DNA and thereby protects it from unwinding¹³⁷. At the opposite end of the spectrum, psychrophilic archaea thrive in permanently cold conditions such as seawater or dry lakes in the Antarctic. However, it is misleading to believe that all archaea have been damned to such a hellish existence. Cultivation-independent methods such as ENVIRONMENTAL GENOME-SHOTGUN SEQUENCING have indicated that mesophilic archaea are remarkably commonplace¹³⁸ and might represent more than 20% of microbial cells in the oceans⁵. So, classical microbiology, with its emphasis on pure culture, is inadequate at determining microbial diversity.

It has been suggested that our inability to culture mesophilic archaea might account for our inability to detect pathogenic archaea¹³⁹. Although indicators of archaeal population density (such as methane levels) have been found to correlate with diseases such as chronic periodontitis¹⁴⁰, there are no examples of an archaeon being directly responsible for a human malady. Lateral gene transfer between archaea and bacteria might have contributed to bacterial pathogenesis, where it has led to the emergence of novel virulence genes in the latter¹⁴¹. Although there might be pathogenic archaea present, it is also possible that an intrinsic feature prevents them from posing a threat to vertebrates. For instance, archaea have unique cell membranes with isoprene side-chains³, and liposome adjuvants that are prepared from archaeal membranes elicit a significantly greater immune response than liposomes that are prepared from bacterial membranes¹⁴². Therefore, an archaeal pathogen would be an easy target for the immune system.

and colleagues have shown that transcriptional activation by GvpE involves binding that occurs upstream from the TFB-recognition element, and probably leads to direct contact with the basal machinery²¹. It is noteworthy that GvpE resembles a basic leucine-zipper

(bZIP) protein, a motif that is commonly found in eukaryotic regulators.

Translation in archaea has been studied much less intensively than transcription, but the message is similar²². The core components (such as rRNA) are

ACIDOPHILE

An organism that requires a low pH for growth; typically less than pH 3.

ALKALIPHILE

An organism that requires a high pH for growth; typically greater than pH 10.

ENVIRONMENTAL GENOME-SHOTGUN SEQUENCING

High-throughput sequencing and computational reconstruction of genomic DNA fragments that are extracted from environmental samples to assess microbial diversity in a cultivation-independent manner.

Table 1 | **Archaea with sequenced genomes or ongoing genome projects**

Species name	Genome size (Mb) (Date)	Phylum	Growth characteristics and optimal temperature	Genetic potential	Sequence resource
<i>Methanocaldococcus jannaschii</i>	1.66 (1996)	Euryarchaeota	Hyperthermophilic methanogen, anaerobic, 85°C		NCBI
<i>Archaeoglobus fulgidus</i>	2.18 (1997)	Euryarchaeota	Hyperthermophilic, sulphate-reducing, anaerobic, 83°C		NCBI
<i>Methanothermobacter thermoautotrophicus</i>	1.75 (1997)	Euryarchaeota	Methanogen, anaerobic, 65°C	+	NCBI
<i>Pyrococcus horikoshii</i>	1.74 (1998)	Euryarchaeota	Hyperthermophilic, anaerobic, 96°C		NCBI
<i>Aeropyrum pernix</i>	1.67 (1999)	Crenarchaeota	Hyperthermophilic, aerobic, 95°C		NCBI
<i>Halobacterium</i> sp. NRC-1*	2.57 (2000)	Euryarchaeota	Halophilic, aerobic, 42°C	+++	NCBI
<i>Halobacterium salinarum</i> *	~2.5 (2000†)	Euryarchaeota	Halophilic, aerobic, 42°C	+++	MPG
<i>Thermoplasma acidophilum</i>	1.56 (2000)	Euryarchaeota	Thermoacidophilic, aerobic, 59°C		NCBI
<i>Thermoplasma volcanium</i>	1.58 (2000)	Euryarchaeota	Thermoacidophilic, aerobic, 60°C		NCBI
<i>Pyrococcus abyssi</i>	1.77 (2001)	Euryarchaeota	Hyperthermophilic, anaerobic, 96°C	++	NCBI
<i>Pyrococcus furiosus</i>	1.91 (2001)	Euryarchaeota	Hyperthermophilic, anaerobic, 96°C	++	NCBI
<i>Pyrolobus fumarii</i>	1.85 (2001‡)	Crenarchaeota	Hyperthermophilic, aerobic, 106°C		
<i>Sulfolobus solfataricus</i>	2.99 (2001)	Crenarchaeota	Thermoacidophilic, aerobic, 80°C	+++	NCBI
<i>Sulfolobus tokodaii</i>	2.69 (2001)	Crenarchaeota	Thermoacidophilic, aerobic, 80°C	+	NCBI
<i>Ferroplasma acidarmanus</i>	1.87 (2002§)	Euryarchaeota	Acidophilic, anaerobic, 42°C		ORNL
<i>Methanopyrus kandleri</i>	1.69 (2002)	Euryarchaeota	Hyperthermophilic methanogen, anaerobic, 98°C		NCBI
<i>Methanosarcina acetivorans</i>	5.75 (2002)	Euryarchaeota	Methanogen, anaerobic, 35°C	+++	NCBI
<i>Methanosarcina barkeri</i>	4.86 (2002§)	Euryarchaeota	Methanogen, anaerobic, 35°C	+++	ORNL
<i>Methanosarcina mazei</i>	4.10 (2002)	Euryarchaeota	Methanogen, anaerobic, 37°C	+++	NCBI
<i>Pyrobaculum aerophilum</i>	2.22 (2002)	Crenarchaeota	Hyperthermophilic, nitrate-reducing, aerobic, 100°C		NCBI
<i>Hyperthermus butylicus</i>	1.67 (2003‡)	Crenarchaeota	Hyperthermophilic, sulphate-reducing, anaerobic, 100°C		
<i>Methanogenium frigidum</i>	~2.5 (2003§)	Euryarchaeota	Psychrophilic methanogen, anaerobic, 15°C		UNSW
<i>Nanoarchaeum equitans</i>	0.49 (2003)	Nanoarchaeota	Symbiotic hyperthermophile, anaerobic, 90°C		NCBI
<i>Sulfolobus acidocaldarius</i>	2.23 (2003‡)	Crenarchaeota	Thermoacidophilic, aerobic, 80°C	++	
<i>Haloarcula marismortui</i>	4.27 (2004)	Euryarchaeota	Halophile, aerobic, 37°C	+	NCBI
<i>Haloferax volcanii</i>	4.03 (2004‡)	Euryarchaeota	Halophile, aerobic, 45°C	+++	UMBI; TIGR
<i>Methanococcoides burtonii</i>	2.56 (2004§)	Euryarchaeota	Psychrotolerant methanogen, anaerobic, 23°C		ORNL
<i>Methanococcus marisaludis</i>	1.66 (2004)	Euryarchaeota	Methanogen, anaerobic, 37°C	+++	NCBI
<i>Methanococcus voltae</i>	~1.9 (2004‡)	Euryarchaeota	Methanogen, anaerobic, 37°C	+++	
<i>Natronomonas pharaonis</i>	2.75 (2004‡)	Euryarchaeota	Haloalkaliphilic, aerobic, 40°C		
<i>Picrophilus torridus</i>	1.55 (2004)	Euryarchaeota	Acidophilic, aerobic, 60°C		NCBI
<i>Thermococcus kodakaraensis</i>	2.09 (2004‡)	Euryarchaeota	Hyperthermophilic, anaerobic, 85°C	++	
<i>Thermoproteus tenax</i>	~1.84 (2004§)	Crenarchaeota	Hyperthermophilic, anaerobic, 86°C		REF. 44
<i>Acidianus brierleyi</i>	~1.9	Crenarchaeota	Thermoacidophilic, aerobic, 70°C		
<i>Halobaculum gomorrense</i>	~2.7	Euryarchaeota	Halophilic, aerobic, 37°C	+	
<i>Haloquadratum walsbyi</i>	~3.18	Euryarchaeota	Halophilic, aerobic, 40°C		
<i>Halorubrum lacusprofundi</i>	~2.6	Euryarchaeota	Psychrotolerant halophile, aerobic, 30°C	+	
<i>Natrialba asiatica</i>	~3.1	Euryarchaeota	Halophile, aerobic, 37°C	+	
<i>Sulfolobus metallicus</i>	~1.9	Crenarchaeota	Thermoacidophilic, aerobic, 80°C		
<i>Cenarchaeum symbiosum</i>		Crenarchaeota	Symbiotic psychrophile, aerobic, 10°C		
<i>Methanococcus thermolithotrophicus</i>		Euryarchaeota	Thermophilic, halotolerant methanogen, anaerobic, 62°C		
<i>Methanosaeta concilii</i>		Euryarchaeota	Methanogen, anaerobic, 37°C		
<i>Methanosarcina thermophila</i>		Euryarchaeota	Thermophilic methanogen, anaerobic, 50°C		
<i>Methanospaera stadtmanae</i>		Euryarchaeota	Methanogen, anaerobic, 37°C		
<i>Methanospirillum hungatei</i>		Euryarchaeota	Methanogen, anaerobic, 37°C		

Date indicates either completion or publication of genome sequence (no entry indicates a continuing genome project). **Halobacterium* spp NRC-1 and *Halobacterium salinarum* genome sequences are essentially identical. †The genome sequence is complete but not published. ‡The genome sequence is published but remains incomplete. +, growth on solid media; ++, potential for rudimentary genetics, such as transformation and selectable markers; +++, potential for advanced genetics, including shuttle vectors, gene replacement and reporter genes. MPG, Max Plank Gesellschaft; NCBI, National Center for Biotechnology Information; ORNL, Oak Ridge National Laboratory; TIGR, The Institute for Genomic Research; UMBI, University of Maryland Biotechnology Institute; UNSW, University of New South Wales (for URLs, see Online links box).

eukaryotic in nature, as are the levels of complexity — more than ten initiation factors are found in archaea and eukaryotes, whereas bacteria require only three. Similarly, translation initiation in archaea and eukaryotes uses methionine, whereas bacteria use N-formyl-methionine. On the other hand, both bacteria and archaea use polycistronic mRNAs, and recognition of mRNA by the ribosome often occurs by means of a purine-rich Shine–Dalgarno sequence in the 5' untranslated region. It is notable that a second mechanism for translation initiation is used in archaea, which operates on leaderless mRNAs and is therefore more reminiscent of the eukaryotic pathway²³. Studies of these two mechanisms, and the circumstances under which they are used, should shed light the origins of translation initiation.

Chromatin. The octameric nucleosome, consisting of two copies each of the histones H2A, H2B, H3 and H4, has long been considered a hallmark of the eukaryotic cell. Because prokaryotes were thought not to require such an ornate machinery for DNA compaction, it came as a considerable surprise when the laboratory of John Reeve reported that the methanogenic archaeon *Methanothermobacter ferredoxigenes* contains a homologue of eukaryotic histones²⁴. Genome sequencing has revealed that histones are widespread among Euryarchaeota but absent from Crenarchaeota⁴.

Archaeal histones dimerize to form a structure that resembles the eukaryotic H3–H4 dimer, and assemble into a tetramer to bind ~60 bp of DNA. However, archaeal histones lack the N-terminal and C-terminal tails that are sites of regulatory post-translational modification in eukaryotes, indicating that chromatin remodelling is not used as a mode of gene regulation in archaea²⁵. By contrast, nucleoid proteins that are found in Crenarchaeota, such as Alba, undergo post-translational modification; a significant proportion of Alba is acetylated at lysine residues, and deacetylation (which is mediated by Sir2) leads to transcriptional repression²⁶. The lysine acetylase that functions on Alba was recently identified (S. Bell, personal communication), and should prove a fruitful area for further study.

DNA recombination and repair. There is considerable interest in studying DNA recombination and repair systems in archaea, as they commonly have to contend with harsh conditions that threaten genomic stability²⁷. Furthermore, these are highly complex repair processes, especially so in eukaryotic cells, which have a specialized programme of meiotic recombination. The potential of archaeal genomics was plainly demonstrated when homology to an archaeal topoisomerase led to the identification of Spo11 as the eukaryotic enzyme responsible for double-strand breaks that are formed during meiosis²⁸.

Strand exchange is the cornerstone of DNA recombination, providing the means to identify and synapse with a homologous template, and is carried out by proteins of the RecA family. The archaeal homologue (RadA) shows greater similarity to the eukaryotic

sequence (Rad51) than the bacterial one (RecA). The resemblance is even more striking at the structural level, to the point at which functional interactions between eukaryotic proteins can be extrapolated from the archaeal crystal structure⁸. RadA has been shown to promote strand exchange *in vitro*²⁹, and a *radA* mutant of *H. volcanii* has been generated that is defective in recombination and is highly sensitive to DNA damage³⁰. A RadA paralogue, RadB, has been identified in the genome sequences of Euryarchaeota. RadB has no strand exchange activity³¹, and *radB* mutants of *H. volcanii* are not defective in recombination (T.A., unpublished observations). Genetic studies of RadB, which are underway in T.A.'s laboratory, should provide some insight into the function of eukaryotic Rad51 paralogues; these are largely of unknown function and have no counterparts in bacteria.

Other forms of DNA repair involve either excision or direct reversal of the lesion. The archaeal homologue of eukaryotic XPF (Rad1), a nuclease that recognizes junctions between single and double-stranded regions of DNA, might function in excision repair. Intriguingly, crenarchaeal XPF lacks the N-terminal 'helicase' domain that is present in the euryarchaeal and eukaryotic proteins³². An example of direct reversal of DNA damage is the photoreactivation system, which uses photolyase to act on pyrimidine dimers. Although this enzyme is not widespread among archaea, it is found in halophiles that are commonly exposed to solar radiation³³. Most archaea also lack homologues of *mutS* and *mutL* genes, which encode the mismatch repair machinery that is conserved from *E. coli* to humans. Despite this, genetic studies of *Sulfolobus acidocaldarius* have shown that archaea are just as efficient at repairing DNA damage as *E. coli*³⁴, indicating that novel pathways of DNA repair have still to be discovered. Such a repair system for THERMOPHILIC archaea has been predicted by genome sequencing and analysis³⁵ and awaits genetic study.

DNA replication. Archaea and bacteria share a genomic structure, usually consisting of a single circular chromosome, but differ in the machinery that is used to carry out DNA replication⁶. As with other aspects of information processing, the archaeal proteins are more similar to eukaryotic homologues than bacterial ones. Because only a subset of the eukaryotic proteins are found in archaea, the archaeal system is simpler and is therefore more amenable to analysis. The laboratory of Hannu Myllykallio used genome analysis to predict (and biochemistry to confirm) the location of the chromosomal replication origin in *Pyrococcus abyssi*³⁶. This prediction was based on the observation that leading strands of replication often contain an excess of G over C nucleotides. The origin of replication is highly conserved among the three *Pyrococcus* species examined (*P. furiosus*, *P. abyssi* and *P. horikoshii*), and the identity of the gene cluster that is located in this region is of particular interest. In addition to sequences that encode RadB (discussed above), a single gene similar to both eukaryotic *cdc6* and *orc1* is found directly adjacent to the origin; *orc1* codes for a subunit of the eukaryotic origin

THERMOPHILE

An organism that requires high temperatures for growth; typically greater than 60°C.

recognition complex, suggesting that Cdc6/Orc1 functions as the initiator protein in archaea. This is indeed the case³⁷, and almost every archaeal chromosomal replication origin identified so far is adjacent to a *cdc6/orc1* gene⁶. The exception is intriguing: both *S. acidocaldarius* and *S. solfataricus* have three *cdc6/orc1* genes and three replication origins, but only two of these co-localize^{38,39}. Even more mystifying is why some species of haloarchaea (such as *Halobacterium* species NRC-1) should require ten distinct *cdc6/orc1* genes. This is a question that genetics is best placed to answer.

Central metabolism and energy conversion. It is commonly stated that operational genes in archaea (that code for central metabolism, energy conversion and biosynthesis) are bacterial in origin¹¹. As with the comparison between archaeal and eukaryotic informational genes, this statement is more of a soundbite than a true representation of archaea. For example, methanogenesis is not found in any bacteria. Genomic analysis of *Methanosarcina acetivorans* has revealed a surprising diversity of methanogenic pathways that use acetate and various one-carbon compounds (acetoclastic and methylotrophic pathways, respectively)⁴⁰. However, *M. acetivorans* is unable to reduce CO₂ using H₂ (the hydrogenotrophic pathway), as this species lacks the ferredoxin-dependent hydrogenase that is encoded by the *ech* operon. The pivotal role of this enzyme has been confirmed by genetic studies in *Methanosarcina barkeri*, which have demonstrated that mutants lacking Ech are unable to grow using H₂ and CO₂ alone⁴¹.

Among the HETEROTROPHIC archaea, a significant fraction can metabolize sugars. Although glycolytic pathways are well conserved in bacteria and eukaryotes, archaea use several variant enzymes, the presence of which can best be explained by independent, convergent evolution⁴². Support for variant metabolic pathways in archaea has come from several studies, such as the prediction of a novel aconitase family by comparative genome analysis⁴³. Aconitase is an essential part of the tricarboxylic acid cycle, and the canonical gene is found only in a minority of euryarchaea. Similarly, genome analysis has indicated that many archaea, such as the HYPERTHERMOPHILIC crenarchaeon *Thermoproteus tenax*, lack the enzymes for an oxidative pentose phosphate pathway (PPP)⁴⁴. As pentoses are essential for anabolic purposes, it is likely that archaea use a variant PPP that is encoded by genes with no obvious bacterial or eukaryotic homologues. Therefore, while it is true that well-conserved operational genes in archaea are most similar to their counterparts in bacteria, there are many novel or variant enzymes that await discovery. Genomics can point the way, but genetics and biochemistry must work hand in hand to unravel these mysteries.

Archaeal genetics — back to basics

In the early years of archaeal genetics, the development of selectable markers and transformation protocols were intimately linked — without a selectable phenotype it is impossible to quantify transformation efficiency and vice versa. This impasse was due to the

scarcity of archaeal antibiotics. Bacterial antibiotics such as ampicillin are safe for medical use because their targets (in this case, peptidoglycan cell walls) are not encountered in eukaryotic cells. Given that these drug targets are also absent from archaea, it is not surprising that such antibiotics are ineffective in the third domain. The issue of selectable markers is discussed in a later section. Here, we summarize the methods that are available for introducing foreign DNA into archaeal cells.

Transformation. To circumvent the bottleneck created by developing transformation protocols without genetic markers, researchers turned to unconventional methods. Cline and Doolittle⁴⁵ assayed TRANSFECTION OF *Halobacterium halobium* with naked DNA from phage ΦH, allowing them to develop an efficient polyethylene glycol-mediated transformation method for *H. volcanii*⁴⁶, which was subsequently adapted for use in various archaea including *M. maripaludis* and *P. abyssi*^{47,48}. It is only effective in species for which SPHEROPLASTS can be generated readily, usually by removing the paracrystalline glycoprotein surface layer (S-layer). By contrast, archaea with a rigid cell wall made of pseudopeptidoglycan, such as *Methanothermobacter thermoautotrophicus*, have been recalcitrant to transformation. Although it is possible to remove the cell wall enzymatically using pseudopeptidoglycan endopeptidase, the protoplasts fail to regenerate (J. Chong, personal communication).

Other transformation protocols have been used with varying success (see table in BOX 3). Electroporation is a versatile technique and can be used for *Methanococcus voltae* and *S. solfataricus*^{49–51}, but is inefficient in species such as *M. acetivorans*⁵². Furthermore, it is not universally applicable; *P. abyssi* cannot be transformed by this method⁴⁸, and electroporation is impossible for halophilic archaea, which cannot tolerate salt concentrations of <1 M NaCl. Heat shock after treatment with CaCl₂, a method that is commonly used for *E. coli*, can be used with some archaea but is not efficient. It is noteworthy that in *Thermococcus kodakaraensis*, CaCl₂ treatment is not essential for DNA uptake⁵³; this is reminiscent of natural transformation, which has been observed in *M. voltae*⁵⁴. Efficient transformation of the *Methanosarcina* species is only possible using a liposome-mediated protocol, which yields >10⁷ transformants per microgram of DNA (REF 52). The drawback of this method is that the requisite cationic liposomes are expensive.

Other gene-transfer mechanisms. Once the DNA is safely inside an archaeal cell, it can be transferred to its neighbours by various means. Phage-mediated TRANSDUCTION is a mainstay of *E. coli* genetics, and similar phenomena have been reported in *Methanobacterium thermoautotrophicum* Marburg⁵⁵ and *M. voltae*⁵⁶. In the case of *M. voltae*, there is still no evidence that gene transfer is mediated by viral particles⁵⁶, and the observation that the transfer agent is resistant to DNase does not rule out alternative routes. For example, bidirectional genetic exchange has been observed during cell mating in *H. volcanii*⁵⁷, which involves cell–cell contact or fusion rather than transduction, and is actually

HETEROTROPH

An organism that requires complex organic molecules such as amino acids and sugars to build macromolecules and derive energy.

HYPERTHERMOPHILE

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

TRANSFECTION

Infection of a host cell by naked DNA or RNA that is isolated from a virus.

SPHEROPLAST

A cell that is denuded of most of its cell wall or surface layer, usually by chemical or enzymatic treatment. Also known as a protoplast.

TRANSDUCTION

Transfer of host genes between archaeal or bacterial species, using a virus as a vector.

Box 3 | Genetics needs a solid (media) foundation

Transformation or gene transfer method	Species	References	Notes
Polyethylene glycol (PEG)	<i>Haloarcula</i> spp., <i>Halobacterium</i> spp., <i>Haloferax volcanii</i> , <i>Methanococcus maripaludis</i> , <i>Pyrococcus abyssi</i>	45,47,48, 145,146	Requires spheroplast formation by removal of the S-layer (usually by treatment with EDTA, but not for <i>Methanococcus maripaludis</i>)
Electroporation	<i>Methanosarcina acetivorans</i> , <i>Methanococcus voltae</i> , <i>Sulfolobus solfataricus</i>	49–52	Not universally applicable in archaea
Liposomes	<i>M. acetivorans</i> , <i>Methanosarcina barkeri</i> , <i>M. voltae</i>	52,114	Efficient, but expensive.
CaCl ₂ and heat shock	<i>M. voltae</i> , <i>Pyrococcus furiosus</i> , <i>Sulfolobus acidocaldarius</i> , <i>Thermococcus kodakaraensis</i>	53,54,105	Not efficient
Cell mating or conjugation	<i>H. volcanii</i> , <i>S. acidocaldarius</i>	57,59	Chromosomal marker exchange requires stable cell contact for <i>Haloferax volcanii</i> but not <i>Sulfolobus acidocaldarius</i>
Virus or plasmid-mediated conjugation	<i>S. solfataricus</i>	50,51,60, 99,100	Mediated by self-spreading vectors based on the SSV1 virus or the pNOB8 plasmid
Transduction	<i>Methanobacterium thermoautotrophicum</i>	55,56	Only seen in <i>Methanobacterium thermoautotrophicum</i> Marburg

The primary requirement for any genetic system is the ability to obtain clonal cultures. Robert Koch first realized the potential of solid media for pure culture methods, by noting that different colonial forms breed true and therefore represent the clonal expansion of a single cell. So, he laid the cornerstone for microbiology and microbial genetics. Two of Koch's associates developed the necessary technology at the end of the nineteenth century: Richard Petri invented the eponymous dish and Walter Hesse (or more accurately, his wife Fannie) adopted agar as a gelling agent. To this day, little has changed.

Archaeal genetics is no exception — growth on solid media is essential before techniques such as transformation can be developed (see accompanying table). The culture of haloarchaea is trivial, as they are aerobic and most grow at 35–45°C and neutral pH. Provided sufficient salt is present in the media, there is little to distinguish halophilic methods from those used with *E. coli*¹²⁵. Handling methanogens is trickier, primarily because they are obligate anaerobes and require an environment with a reducing potential of less than –330 mV. Efficient cultivation has only been possible since 1950, when the 'Hungate' technique for preparing and dispensing chemically-prereduced media into stoppered tubes was introduced¹²³. Another innovation due to Hungate was the use of roll tubes, in which agar is spun horizontally to coat the inner surface of the vessel¹⁴³. A refined version of the Hungate technique is still in use today¹⁴⁴, although the introduction of the anaerobic glove box has allowed the use of conventional Petri dishes.

Gellan gum (also known as Gelrite) has been instrumental in the establishment of genetics for hyperthermophiles, which grow above the gelling temperature of agar¹²⁴. Gelrite is a deacetylated polysaccharide that is produced by *Pseudomonas elodea* and solidifies in the presence of divalent cations to form a matrix that is stable at temperatures as high as 120°C. The main disadvantage of Gelrite is that it contains trace nucleic acids, which can interfere with selection for uracil PROTOTROPHY^{53,84}. Transformants are therefore grown in selective liquid media (deficient in uracil) before plating on Gelrite. However, given the importance of selection on solid media to the development of genetics, a more elegant solution will no doubt be found.

stimulated by DNase treatment⁵⁸. Similar cell mating has also been seen in *S. acidocaldarius* and *S. solfataricus*. In the former, chromosomal marker exchange between two AUXOTROPHIC mutants can be measured by the appearance of stable genetic recombinants⁵⁹. In the latter, conjugative plasmids such as pNOB8 have been shown to propagate throughout the culture using a cell–cell contact-dependent mechanism⁶⁰. The kinetics of cell mating in *Sulfolobus* spp differs from that in *H. volcanii*, as it does not require cell–cell contacts to be stabilized by growth on solid media^{58,61}.

Restriction–modification systems. From the perspective of foreign DNA, the inside of a cell can be a hostile environment. Restriction–modification systems are widespread among prokaryotes, and archaea are no

exception. Enzymes that recognize 5'-CTAG-3' are common, having been identified initially in *M. thermoformicum*, where they are plasmid-encoded⁶². Genome analysis has revealed the presence of putative CTAG methylases in many species (restriction endonucleases are virtually impossible to identify by sequence homology), and DNA that is isolated from *H. volcanii* is resistant to cleavage at 5'-CTAG-3' sites, indicating that this site is modified, possibly by methylation⁴⁶. *H. volcanii* also has a restriction system that recognizes adenine-methylated GATC sites (which occur frequently in vectors that are based on *E. coli* plasmids), resulting in DNA fragmentation followed by plasmid loss or chromosomal integration by recombination⁶³. This can be circumvented by passaging the DNA through an *E. coli dam*⁻ strain that is deficient in GATC methylation⁶⁴. Other

PROTOTROPH

An organism that can grow on minimal media that contain a carbon source and inorganic compounds.

AUXOTROPH

A mutant that requires nutrients that are not needed by wild-type strains for growth on minimal media.

Table 2 | **Selectable markers used in archaeal genetics**

Selection marker	Marker	Type	Species	References	Notes
Alcohol dehydrogenase	<i>adh</i> , <i>adh-hT</i>	A	<i>Pyrococcus furiosus</i> , <i>Sulfolobus solfataricus</i>	106,147	The <i>adh-hT</i> gene is from <i>Bacillus stearothermophilus</i> .
Anisomycin	<i>Ani^R</i> (23S rRNA)	A	<i>Halobacterium</i> spp	148	A mutant 23S rRNA gene, used for mutation of the chromosomal rRNA gene but not developed as a selectable marker.
Bleomycin	<i>ShBle</i>	A	<i>Haloferax volcanii</i>	149	Bleomycin-resistance <i>ShBle</i> gene from <i>Streptoalloteichus hindustanus</i> .
Hygromycin B	<i>hph</i>	A	<i>S. solfataricus</i>	51,150	Mutated (thermostable) version of the hygromycin B phosphotransferase (<i>hph(mut)</i>) gene from <i>Escherichia coli</i> .
Mevinolin	<i>Mev^R</i> (<i>hmg</i>)	A	<i>Halobaccula</i> spp, <i>H. volcanii</i> <i>Halobacterium</i> spp	63,72,73,145	Use of heterologous <i>Mev^R</i> gene prevents recombination between chromosomal <i>hmg</i> gene and <i>Mev^R</i> .
Neomycin	<i>Neo^R</i> (APH3')	A	<i>Methanococcus maripaludis</i>	70	Geneticin is not inhibitory to <i>Methanococcus maripaludis</i> .
Novobiocin	<i>Nov^R</i> (<i>gyrB</i>)	A	<i>H. volcanii</i> , <i>Sulfolobus acidocaldarius</i>	66,82	Inhibits DNA gyrase; <i>Nov^R</i> over-expresses DNA gyrase B subunit.
Pseudomonic acid	<i>PA^R</i> (<i>lleS</i>)	A	<i>Methanosarcina acetivorans</i> , <i>Methanosarcina barkeri</i> , <i>Methanobacterium thermoautotrophicum</i> Marburg	71,151	Mutant isoleucyl-tRNA synthetase <i>lleS</i> gene from <i>Methanosarcina barkeri</i> Fusaro. Pseudomonic acid is not commercially available.
Puromycin	<i>Pur^R</i> (<i>pac</i>)	A	<i>M. acetivorans</i> , <i>M. barkeri</i> , <i>M. maripaludis</i> , <i>Methanosarcina mazei</i> , <i>Methanococcus voltae</i>	52,69,97,152, 153	Puromycin-resistance <i>pac</i> gene from <i>Streptomyces alboniger</i> , which is widely used as a marker for methanogenic archaea.
Thiostrepton	<i>Ths^R</i> (23S rRNA)	A	<i>Halobacterium</i> spp	148	Mutant 23S rRNA gene, used for mutation of the chromosomal rRNA gene but not developed as a selectable marker.
Trimethoprim	<i>hdrA</i>	A	<i>H. volcanii</i>	154	Trimethoprim is not widely used as a selectable marker.
8-aza-2,6-diaminopurine (8ADP) or 8-aza-hypoxanthine	<i>hpt</i>	C	<i>M. acetivorans</i> , <i>M. maripaludis</i>	88,89	Used in conjunction with puromycin or neomycin as a positive selectable marker.
Uracil/5-fluoroorotic acid (5-FOA)	<i>ura3</i> , <i>pyrE</i> , <i>pyrF</i>	X/C	<i>Halobacterium</i> spp, <i>H. volcanii</i> , <i>Pyrococcus abyssi</i> , <i>S. solfataricus</i> , <i>Thermococcus kodakaraensis</i>	48,53,77, 79–83, 84	Useful marker. Isolation of spontaneous 5-FOA-resistant mutants is easy, allowing the system to be implemented widely.
Histidine	<i>hisA</i> , <i>his1</i>	X	<i>M. voltae</i> , <i>H. volcanii</i>	155,156	<i>his1</i> gene of <i>Haloferax volcanii</i> has not been developed as a marker.
Lactose	<i>lacS</i> (+ <i>lacTr</i>)	X	<i>S. solfataricus</i>	86	The <i>lacS</i> gene in the host strain is disrupted by transposons or is deleted.
Leucine	<i>leuB</i> , <i>leuA</i>	X	<i>H. volcanii</i> , <i>M. maripaludis</i>	76,80	<i>leuB</i> and <i>leuA</i> genes that are in host strains are deleted. <i>M. maripaludis leuA</i> has not yet been developed as a marker.
Proline	<i>proC</i>	X	<i>M. acetivorans</i>	88	<i>E. coli proC</i> gene also functions as a marker.
Thymidine	<i>hdrA</i> , <i>hdrB</i>	X	<i>H. volcanii</i>	80,154	<i>hdrB</i> gene that is in the host strain is deleted. Complex media that are based on yeast extract are deficient in thymidine.
Tryptophan	<i>trpA</i> , <i>trpE</i>	X	<i>H. volcanii</i> , <i>T. kodakaraensis</i>	53,80	<i>trpE</i> mutants (tryptophan auxotrophs) of <i>Thermococcus kodakaraensis</i> have been isolated, but no selectable marker has been developed.

A, antibiotic; C, counter-selectable; X, auxotrophic or similar.

restriction–modification systems have been documented, such as the *SuaI* enzyme of *S. acidocaldarius* that recognizes 5'-GGCC-3' (REF. 65).

Antibiotics. Although most bacterial antibiotics are ineffective in archaea, several exceptions have been exploited to develop selectable markers for archaeal genetics (TABLE 2). Novobiocin is a potent inhibitor of DNA gyrase (*gyrB*), an enzyme that is present in both bacteria and archaea. To develop a vector for halophilic archaea, the laboratory of Mike Dyall-Smith isolated a novobiocin-resistant mutant of *Haloferax* strain Aa2.2 (REF. 66)

that mapped to the *gyrB* gene⁶⁷. Puromycin is another drug that is effective in both bacteria and archaea, and has been shown to inhibit growth in *M. voltae* and other methanogens⁶⁸; it is the most widely used antibiotic for this group of archaea. The resistance marker (puromycin transacetylase) is a bacterial gene from *Streptomyces alboniger*. Owing to differences between bacterial and archaeal gene regulation, it is transcribed using an *M. voltae* promoter⁶⁹. A similar approach was used to generate a construct for neomycin resistance in *M. maripaludis* (using APH3'I and APH3'II genes from Tn903 and Tn5, respectively⁷⁰).

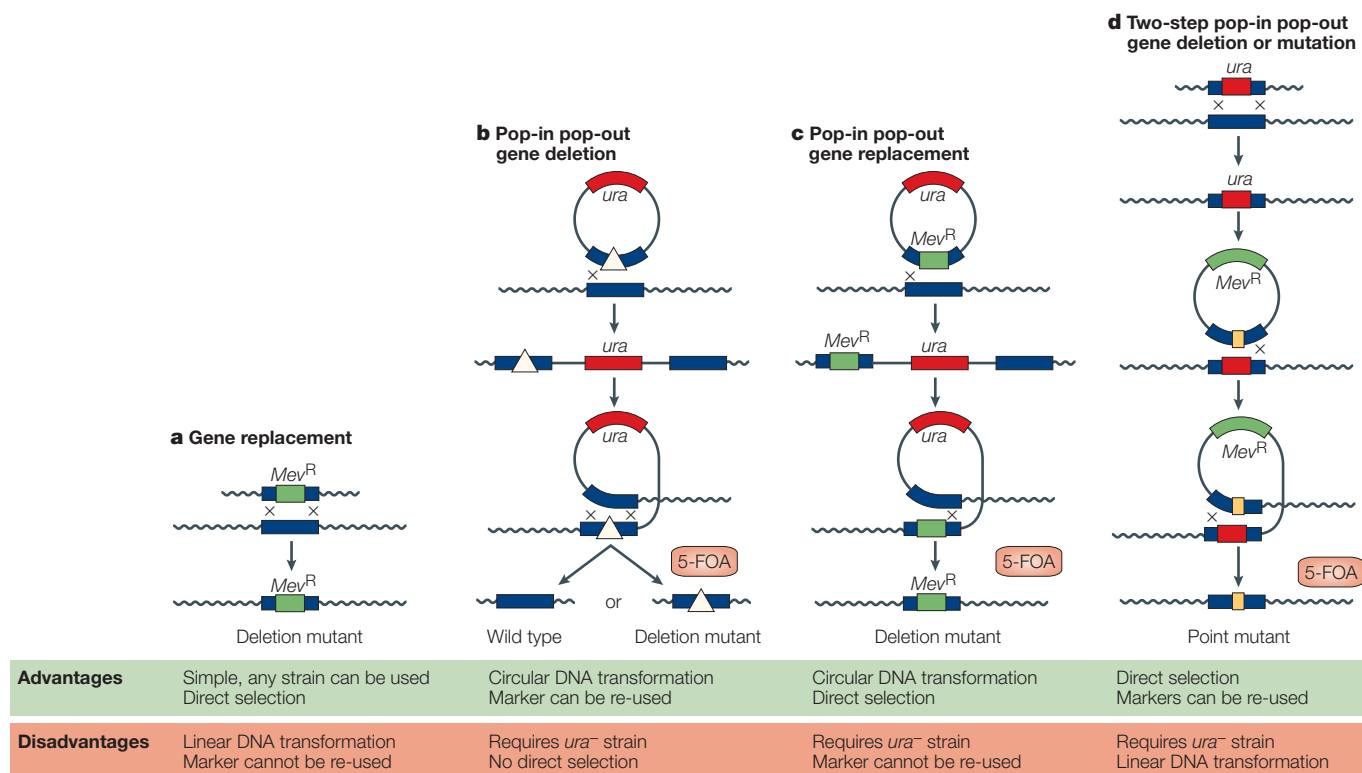


Figure 1 | **Gene knockout methods that are used in archaeal genetics.** **a** | Direct replacement of a gene with a selectable marker, by recombination between linear DNA which comprises flanking regions of the gene and a chromosomal target. In some archaeal species, recombination using linear DNA is less efficient than circular DNA. **b** | The pop-in pop-out method uses circular DNA and selection for transformation to uracil prototrophy. Therefore, a *ura*⁻ strain must be used^{53,77,79}. Intramolecular recombinants that have lost the plasmid are counter-selected using 5-fluoroorotic acid (5-FOA), which is converted to toxic 5-fluorouracil in *ura*⁺ (but not *ura*⁻) cells. Unless the mutant has a readily screened phenotype, the deletion must be verified by southern blotting. **c** | Variant of the pop-in pop-out method for gene deletion, in which the gene is replaced with a marker that allows direct selection⁸⁰. **d** | Combination of gene replacement (with *ura* marker) and the pop-in pop-out method, suitable for generating point mutations⁷⁷. Counter-selection with 5-FOA ensures that the *ura*-marked gene deletion is replaced with the desired mutation.

By contrast, the gene for pseudomonic acid resistance originates in archaea, and was generated by mutagenesis of the isoleucyl-tRNA synthetase gene from *M. barkeri*⁷¹. Similarly, resistance to mevinolin, a 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitor, was isolated from a spontaneous *hmgA* mutant of *H. volcanii*⁷². However, pseudomonic acid is not available commercially, and mevinolin is difficult to obtain as it is licensed as a cholesterol-lowering drug. By inhibiting the conversion of acetyl-CoA to mevalonic acid, mevinolin prevents the synthesis of cholesterol in humans and isoprenoid lipid side chains in archaea. There are additional drawbacks to these antibiotics: spontaneous resistance can arise at high frequency owing to gene amplification, and plasmids that bear mevinolin or novobiocin markers suffer from instability owing to homologous recombination with the chromosome (both these markers were derived from *Haloferax* spp and are virtually identical to the chromosomal sequence of *H. volcanii*). This instability can be alleviated by using markers from distantly related species, such as the mevinolin-resistant *hmgA* mutant allele from *Haloarcula hispanica*⁷³.

Auxotrophic selectable markers. Plasmid instability owing to recombination could also be prevented by deleting the homologous chromosomal gene. Although this is not possible for mevinolin and novobiocin-resistance markers (as both *hmgA* and *gyrB* are essential), it is feasible for genes that are involved in amino-acid biosynthesis and other metabolic pathways where auxotrophic strains can easily be complemented. For example, a *leuB* deletion mutant that is defective in leucine synthesis can be grown on complete media, unless selection for a plasmid-encoded *leuB* marker is required, in which case leucine-deficient media is used. Such an approach has been adopted in yeast genetics, as few bacterial antibiotics are effective against eukaryotic cells. In addition to leucine, strains that are auxotrophic for histidine, proline, tryptophan and thymidine have been isolated in several species (TABLE 2). The principal drawback with auxotrophic markers is that they cannot be developed easily in obligatory AUTOTROPHS, which includes most methanogens. However, many *Methanosarcina* and *Methanococcus* species are facultative autotrophs that readily take up amino acids, and are therefore compatible with

AUTOTROPH

An organism that can synthesize its own macromolecules from simple, inorganic molecules such as carbon dioxide, hydrogen and ammonia.

COUNTER-SELECTABLE MARKER
A marker that if present leads to cell death under selective conditions, usually by conferring sensitivity to an antibiotic or by promoting the synthesis of a toxic product from a non-toxic precursor.

auxotrophic markers⁷⁴. To enable full exploitation of auxotrophic markers, it is best if the organism can be grown on chemically defined (minimal) media, as is the case for *P. abyssi*⁷⁵, *H. volcanii*⁵⁸ and *M. maripaludis*⁷⁶.

Gene-knockout systems. Auxotrophic markers for uracil biosynthesis (*ura3*, *pyrE* or *pyrF* genes) are the most useful, as they can be COUNTER-SELECTED using 5-fluoroorotic acid (5-FOA); *ura⁻* cells are resistant to this compound owing to their inability to convert 5-FOA to the toxic analogue 5-fluorouracil. Such markers can be implemented in organisms that grow poorly on minimal media, as complex media that are deficient in uracil (such as casamino acids) can be used. This has enabled the establishment of gene-knockout systems in *Halobacterium* spp^{77,78}, *H. volcanii*^{79,80} and *T. kodakaraensis*⁵³ (TABLE 2). The salient features of these systems are that the uracil marker can be reused (FIG. 1b), and that transformation is carried out using circular

DNA (transformation using linear DNA is inefficient in some species). Uracil-auxotrophic mutants have also been isolated in *P. abyssi*^{48,81}, *S. acidocaldarius*⁸² and *S. solfataricus*⁸³, but gene-knockout systems have still to be developed. A trivial problem is that Gelrite, the gelling agent used in solid media for hyperthermophiles, contains trace uracil⁸⁴ (BOX 3). A more serious problem is that, in the widely used *S. solfataricus* P1 and P2 strains, gene-targeting constructs fail to recombine with the chromosome⁸³. Although it is possible that recombination is suppressed in *S. solfataricus* owing to active transposable elements⁸⁴, *Halobacterium* spp also suffers from active transposition but is proficient for recombination⁸⁵. Moreover, a different isolate of *S. solfataricus* is capable of homologous recombination, and has been used for gene-knockout experiments (using selection for lactose utilization)⁸⁶.

Counter-selectable markers have recently been developed for methanogens, using the purine-salvage enzyme

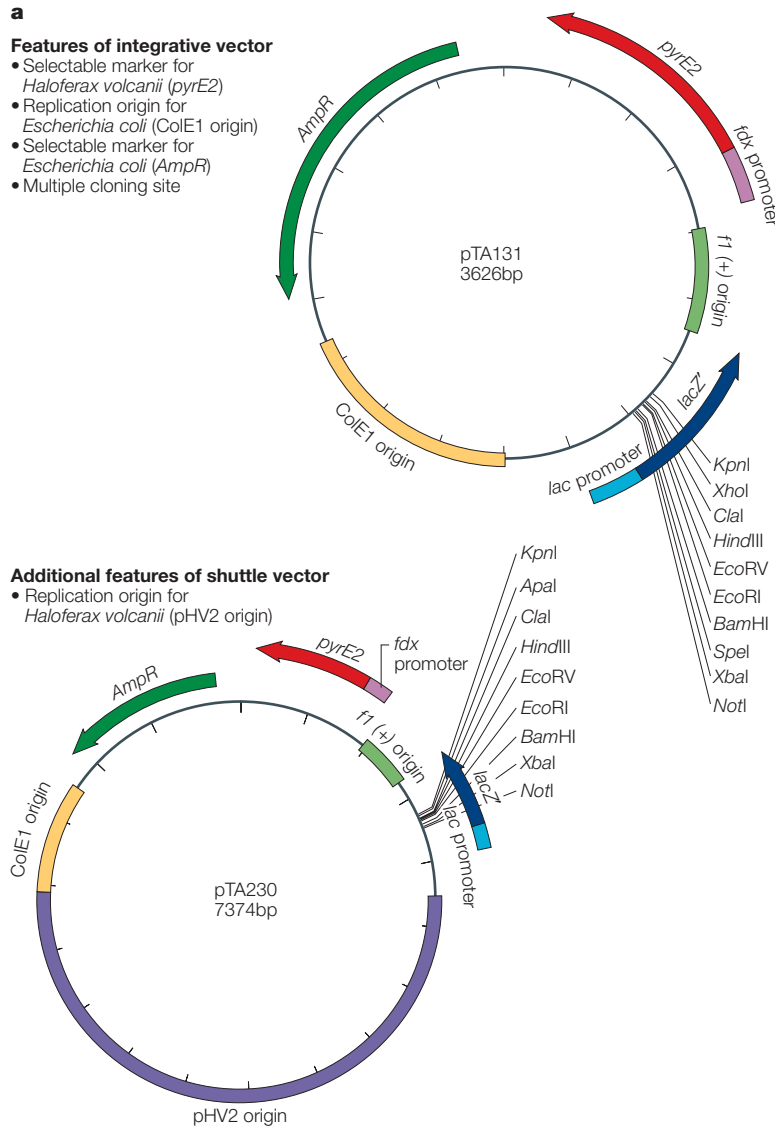
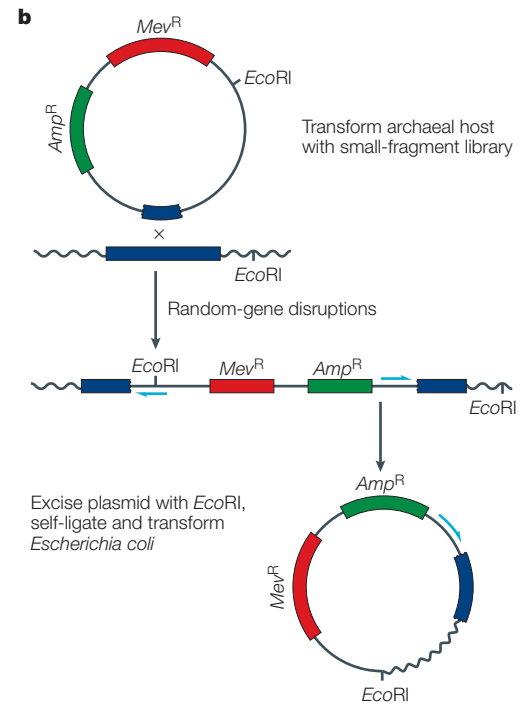


Figure 2 | Plasmid vectors. a | Typical integrative and shuttle-plasmid vectors for archaeal genetics (in this case, *Haloferax volcanii*), with relevant features⁸⁰. See online **supplementary information S1** (table) of commonly used plasmids. **b** | Random insertional mutagenesis using a small-fragment library. Recombination between an integrative plasmid that carries a small (internal) fragment of a gene and the chromosome leads to disruption of the gene. Rapid identification of the mutant gene is possible by using DNA-sequencing primers (light blue) directed to plasmid sequences. The insertion can be cloned directly from genomic DNA by cutting with an enzyme (for example, *EcoRI*) to liberate the plasmid and some surrounding sequence, which is self-ligated and used to transform *Escherichia coli*.



hypoxanthine phosphoribosyltransferase that is encoded by the *hpt* gene; mutants are resistant to the toxic base analogues 8-aza-2,6-diaminopurine and 8-azahypoxanthine⁸⁷. Unlike systems based on uracil and 5-FOA, gene knockout with *hpt* requires an additional marker for positive selection of plasmid integration. Puromycin and neomycin-resistance markers have been used with *hpt*, to construct an *AproC* mutant of *M. acetivorans*⁸⁸ and alanine-utilization mutants of *M. maripaludis* respectively⁸⁹.

Random mutagenesis. As mentioned in previous sections, similarity to bacterial or eukaryotic enzymes can be used to predict the function of only half the proteins that are encoded by archaeal genomes. If we are to explain the function of the remaining half, we must move beyond targeted gene knockouts. Random mutagenesis provides the means to uncover genes and reaction pathways that are unique to archaea. UV radiation and chemical mutagenesis (using ethyl methanesulphonate) have been used to isolate auxotrophic mutants of *H. volcanii*⁵⁸, *M. voltae*⁵⁴, *M. maripaludis*⁹⁰ and *P. abyssi*⁹¹. Because these mutations are difficult to map, transposon-insertion mutagenesis has been attempted. *In vitro* transposition was used to study the *nifH* gene of *M. maripaludis*⁹², although mutagenesis was not random because a defined target (rather than the whole genome) was used. The laboratory of Bill Metcalf has developed an elegant *in vivo* transposition system for *M. acetivorans* using a modified version of *mariner*-family transposon *Himar1*, which carries a puromycin-resistance marker as well as features that allow easy cloning of transposon insertions⁹³. Unfortunately, this system is restricted to methanogens, as eukaryotic or bacterial transposons cannot function in the hypersaline interior of halophiles or at the high temperatures that are required by hyperthermophiles. Synthetic transposons that are based on insertion sequences from *H. salinarum* have been constructed for use in *H. volcanii*, but have had little success⁹⁴.

Random insertional mutagenesis is possible without transposition, so long as the species is proficient for homologous recombination. In this approach, recombination between a truncated version of the gene and its chromosomal copy leads to an insertion–disruption mutation (FIG. 2b). A targeted version of insertional mutagenesis has been used in *M. voltae*, to characterize genes that encode flagellins and hydrogenases^{95–97}. For random mutagenesis, a genomic library of small fragments (less than a full-length gene) is used to target recombination. As with the *in vivo* transposition system for *M. acetivorans*, insertion–disruptions are easily cloned by cutting and self-ligating genomic DNA fragments from the mutant, followed by introduction into *E. coli* (FIG. 2b). Such random insertional mutagenesis has been used to isolate acetate auxotrophs of *M. maripaludis*⁹⁸.

Plasmid vectors. There are two basic types of vector for archaeal genetics: INTEGRATING-VECTOR PLASMIDS and SHUTTLE-VECTOR PLASMIDS (see FIG. 2a; online [supplementary information S1](#) (table)). As their name implies, integrating (or suicide) plasmids must integrate into the host

chromosome, as they do not have an origin of replication for archaea. They are usually used in gene knockout or insertion–disruption mutagenesis, where efficient homologous recombination is paramount. In strains where this is not possible, such as *S. solfataricus* P1 and P2 isolates, vectors that are based on the SSV1 virus have been used that integrate into the chromosome by site-specific recombination⁸³. SSV1 is also capable of stable replication as a circular plasmid, and this faculty has permitted the construction of recombinant shuttle vectors for *S. solfataricus*^{51,99}. An intriguing feature of SSV1 (and derived vectors) is that it spreads efficiently in cultures without lysis of the host cells⁵⁰. This conjugative behaviour eliminates the need for efficient transformation. Self-spreading is also seen with pNOB8, another plasmid of *Sulfolobus* spp⁶⁰, but vectors that are derived from pNOB8 impose a significant burden on the host cell and have not been widely used¹⁰⁰.

Shuttle vectors for other species are more conventional and use replication origins taken from plasmids that are indigenous to the host. For example, the shuttle vector in FIG. 2a (pTA230) uses the origin from pHV2, a naturally occurring 6.4 kb *H. volcanii* plasmid. As the plasmid is non-essential, *H. volcanii* could be cured of pHV2 by using ethidium bromide, therefore generating the widely used strain WFD11 (REF. 46). A strain that was cured of pHV2 without using ethidium bromide (DS70) has since been isolated⁷³. The WFD11 strain enabled the laboratory of Ford Doolittle to develop shuttle vectors for halophilic archaea, using the pHV2 origin and a mevinolin-resistance marker⁷². So far, it remains the most commonly used replicon in *H. volcanii*. Although pHV2-based plasmids can replicate in *Halobacterium* spp, several additional shuttle vectors have been derived from plasmids pGRB1 (REF. 101) and pHH1 (REF. 63). Interestingly, plasmids that are based on pHV2 and pHH1 fail to replicate in recombination-deficient *radA* mutants of *H. volcanii*³⁰, but pHK2 replicons⁶⁶ do not have this problem.

Shuttle vectors for methanogens are less common. The most useful replicon is based on the naturally occurring plasmid pC2A from *M. acetivorans*. The laboratory of Bill Metcalf has developed a series of pC2A derivatives, using puromycin resistance as a selectable marker, and demonstrated that they can transform various *Methanosarcina* species⁵². Shuttle vectors for use in *M. maripaludis* have been derived from the cryptic plasmid pURB500. Early incarnations suffered from instability in *E. coli* hosts, most probably owing to the high A+T content of the replicon (~70%)¹⁰²; stable vectors featuring a gene expression cassette for *M. maripaludis* and a *lacZ* gene for blue–white screening in *E. coli* have since been constructed¹⁰³. Plasmids that are for use in hyperthermophilic euryarchaea are similarly rare. So far, only shuttle vectors that are based on the small pGT5 plasmid of *P. abyssi* strain GE5 have been developed¹⁰⁴. They can be stably propagated in *P. abyssi* strain GE9 (which is devoid of pGT5), as well as the crenarchaeote *S. acidocaldarius*^{48,105,106}, indicating that mechanisms of plasmid replication are conserved between the main archaeal phyla.

INTEGRATIVE-VECTOR PLASMID

A plasmid vector that is unable to replicate in an archaeal host, which therefore must integrate into the host chromosome by homologous or site-specific recombination.

SHUTTLE-VECTOR PLASMID

A plasmid vector that can replicate in both *Escherichia coli* and an archaeal host.

Table 3 | Phenotypic markers and reporter genes used in archaeal genetics

Reporter gene	Host species	Marker type	References	Notes
<i>bgaH</i>	<i>Halobacterium</i> spp, <i>Haloferax volcanii</i>	Colour (X-gal)	108–110	β -galactosidase from <i>Haloferax alicantei</i> . Wild-type <i>H. volcanii</i> lacks detectable β -galactosidase activity
<i>lacS</i>	<i>Sulfolobus solfataricus</i>	Colour (X-gal)	83,111–113	Thermostable β -galactosidase from <i>S. solfataricus</i>
<i>lacZ</i>	<i>Methanococcus maripaludis</i>	Colour (X-gal)	20	The <i>lacZ</i> gene is from <i>Escherichia coli</i> . Colour development requires exposure to oxygen
<i>uidA</i>	<i>Methanosarcina acetivorans</i> , <i>Methanococcus voltae</i>	Colour (X-gluc)	88,107	The <i>uidA</i> gene is from <i>E. coli</i> . Colour development requires exposure to oxygen
<i>treA</i>	<i>M. voltae</i>	Enz	114	The trehalase gene <i>treA</i> is from <i>Bacillus subtilis</i> , and is enzyme insensitive to moderate salinity
GFP	<i>H. volcanii</i>	Fluor	115	A modified variant of GFP that is soluble in <i>H. volcanii</i> due to reduced hydrophobicity
<i>hdrA</i>	<i>H. volcanii</i>	Res	157	The ferredoxin (<i>fdx</i>) promoter of <i>Halobacterium</i> spp was analysed in <i>H. volcanii</i> using trimethoprim. Dihydrofolate reductase (<i>hdrA</i>) is competitively inhibited by trimethoprim

Colour, blue colouration on exposure to chromogenic indicator; Enz, enzymatic assay from cell lysate; Fluor, fluorescent reporter protein; GFP, green fluorescent protein; Res, resistance to antibiotic; X-gal, 5-bromo-4-chloro-3-indoyl- β -D-galactopyranoside; X-gluc, 5-bromo-4-chloro-3-indoyl- β -D-glucuronide.

Analysis of gene expression. Few compounds have had a greater impact on microbial genetics than X-gal (5-bromo-4-chloro-3-indoyl- β -D-galactopyranoside), a chromogenic substrate that is converted by β -galactosidase into an insoluble blue dye. Originally developed by Julian Davies and Jacques Monod for studies of the *lac* operon of *E. coli*, it has since been put to a multitude of uses. In archaea, as in many other organisms, it has been used as a phenotypic reporter for gene expression (TABLE 3). The *E. coli* genes can be used directly in methanogens: *lacZ* has been used to monitor gene expression in *M. maripaludis*²⁰ and the β -glucuronidase gene *uidA* has been used similarly in *M. voltae* and *M. acetivorans*^{88,107}. However, as methanogens are strict anaerobes and oxygen is necessary for blue-colour development from X-gal, replica-planting is often essential if viable cells are to be recovered; this is not necessary for *M. maripaludis*, which can tolerate short exposure to oxygen (J. Leigh, personal communication). Although such problems are not encountered with halophiles (they are aerobic), *E. coli lacZ* is not active in the high salt concentrations found in the haloarchaeal cytosol (up to 5 M KCl). The laboratory of Mike Dyall-Smith therefore isolated a β -galactosidase gene *bgaH* from *Haloferax alicantei* (now called *Haloferax lucentensis*) that develops a blue colour from X-gal (REF. 108). Moreover, it is functional in *H. salinarum* and *H. volcanii* (which lacks detectable β -galactosidase activity), and has been used as a reporter gene for transcription analyses in both species^{109,110}. A similar approach has been taken for thermophilic archaea, using a thermostable β -galactosidase from *S. solfataricus* that is encoded by *lacS*^{111,112}. A mutant strain of *S. solfataricus* is available, in which *lacS* has been inactivated by transposition of an insertion element¹¹³, and was recently used by the laboratory of Christa Schleper to develop a sophisticated gene-reporter system⁸³.

Phenotypic markers other than β -galactosidase have been used (TABLE 3). For example, the salt-resistant trehalase gene *treA* from *Bacillus subtilis* is functional in

M. voltae, and its activity can be assayed in cell lysates¹¹⁴. Recently, a modified derivative of GFP was developed that is soluble and active in the high salt cytosol of *H. volcanii*¹¹⁵. However, researchers are increasingly turning to *ersatz* genetics. Genome sequence data have led to the development of archaeal microarrays, enabling studies of the response to UV radiation in *Halobacterium* spp NRC-1 (REF. 116), a characterization of the central metabolism of *H. volcanii*¹¹⁷, and the identification of chromosomal replication origins in *Sulfolobus* spp³⁹. It is noteworthy that the latter study was only possible because the cell cycle of *Sulfolobus* spp has been studied in some detail, resulting in various means for synchronizing cell cultures¹¹⁸. No doubt further microarray studies will be published in the near future, but if researchers are to make full use of such modern technology, they must first confront basic aspects of archaeal cell biology such as the cell cycle. Without bread-and-butter genetics, we will continue to operate without a solid foundation of knowledge about these fascinating organisms.

Future directions

Since the field was last reviewed^{119,120}, there has been considerable progress in development of tools for archaeal genetics. Gene-knockout systems in particular have made possible the systematic analysis of pathways that operate in this domain of life (FIG. 1). However, there is much work to be done. For example, *S. solfataricus* P1 and P2 strains stubbornly refuse to integrate foreign DNA into the chromosome by homologous recombination, thereby limiting the scope of genetics in this important organism. The way forward might be to use a different isolate of *S. solfataricus* that is proficient for recombination⁸⁶. With the exception of SSV1, archaeal viruses have still to be harnessed for genetic purposes. Gene transfer by phage-mediated transduction would speed up the construction of archaeal mutant strains. Gene-expression systems with tightly regulated promoters are badly lacking. Heat-inducible chaperonin promoters are available for *H. volcanii* and *S. solfataricus*^{83,121},

but the use of heat-shock to induce transcription is far from desirable. In this respect, progress is being made in *M. acetivorans*, in which an acetate-inducible overexpression system has been developed (K. Sowers, personal communication). An improved method of gene regulation would also allow the development of archaeal two-hybrid systems, because yeast or bacterial two-hybrid systems are seldom of any value for analysing interactions between halophilic or thermophilic proteins. Finally, the genome sequences of several key archaea, including *H. volcanii*, *M. voltae* and *S. acidocaldarius*, have still to be published (TABLE 1). No doubt, this will be rectified in the near future.

Above all, more researchers should be working on archaea. Neophyte 'archae-ologists' can find an entertaining introduction to the subject in *The Surprising*

Archaea by John Howland¹²². There are laboratory manuals containing detailed protocols for methanogens¹²³, thermophiles¹²⁴ and halophiles¹²⁵; also for halophiles, the excellent 'HaloHandbook' is available online (see Online links). Above all, researchers thinking of switching to archaea (and those who are merely curious) should remember that there is no single model organism for this entire domain. The wide range of habitats that are colonized by archaea is testament to their diversity, which is reflected at the molecular level by the bewildering array of metabolic and energy conversion mechanisms they use. Nevertheless, there are core functions that are related to information processing that unite and define archaea, and it is here that they share a common heritage with eukaryotes. Exciting discoveries await those who take the third way.

- Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl Acad. Sci. USA* **74**, 5088–5090 (1977).
- 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).
Together with reference 1, this is the seminal paper about archaea. Woese and colleagues use phylogenetic analysis that is based on rRNA sequences to propose a three-domain system for living organisms.
- van de Vossenbergh, J. L., Driessen, A. J. & Konings, W. N. The essence of being extremophilic: the role of the unique archaeal membrane lipids. *Extremophiles* **2**, 163–170 (1998).
- White, M. F. & Bell, S. D. Holding it together: chromatin in the Archaea. *Trends Genet.* **18**, 621–626 (2002).
- DeLong, E. F. & Pace, N. R. Environmental diversity of bacteria and archaea. *Syst. Biol.* **50**, 470–478 (2001).
- Grabowski, B. & Kelman, Z. Archaeal DNA replication: eukaryal proteins in a bacterial context. *Annu. Rev. Microbiol.* **57**, 487–516 (2003).
- Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905–920 (2000).
Archaeal proteins are often used for structural studies owing to their high rigidity. This paper is an impressive example, and describes the structure of the Haloarcula marismortui 50S ribosomal subunit.
- Shin, D. S. *et al.* Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *EMBO J.* **22**, 4566–4576 (2003).
- van den Burg, B. Extremophiles as a source for novel enzymes. *Curr. Opin. Microbiol.* **6**, 213–218 (2003).
- Bult, C. J. *et al.* Complete genome sequence of the methanogenic archaeon, *Methanococcus jarnaschii*. *Science* **273**, 1058–1073 (1996).
- Rivera, M. C., Jain, R., Moore, J. E. & Lake, J. A. Genomic evidence for two functionally distinct gene classes. *Proc. Natl Acad. Sci. USA* **95**, 6239–6244 (1998).
- Boucher, Y. *et al.* Lateral gene transfer and the origins of prokaryotic groups. *Annu. Rev. Genet.* **37**, 283–328 (2003).
A comprehensive and up-to-date review of the breadth and depth of lateral gene transfer.
- Jain, R., Rivera, M. C. & Lake, J. A. Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl Acad. Sci. USA* **96**, 3801–3806 (1999).
This paper, together with reference 11, helps to explain why archaea are a mosaic of bacterial and eukaryotic features, and why genes that encode the latter are not subject to lateral gene transfer.
- Graham, D. E., Overbeek, R., Olsen, G. J. & Woese, C. R. An archaeal genomic signature. *Proc. Natl Acad. Sci. USA* **97**, 3304–3308 (2000).
By comparing several complete genome sequences, this study seeks to eliminate the background noise that is generated by lateral gene transfer, and derives a set of signature genes that defines the archaea in a holistic manner.
- Huet, J., Schnabel, R., Sentenac, A. & Zillig, W. Archaeobacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. *EMBO J.* **2**, 1291–1294 (1983).
In this study, Zillig and colleagues show for the first time that information processing in archaea uses enzymes that are similar to those found in eukaryotes.
- Bell, S. D. & Jackson, S. P. Mechanism and regulation of transcription in archaea. *Curr. Opin. Microbiol.* **4**, 208–213 (2001).
- Thompson, D. K., Palmer, J. R. & Daniels, C. J. Expression and heat-responsive regulation of a TFIIB homologue from the archaeon *Haloferax volcanii*. *Mol. Microbiol.* **33**, 1081–1092 (1999).
- Aravind, L. & Koonin, E. V. DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Res.* **27**, 4658–4670 (1999).
- Bell, S. D., Cairns, S. S., Robson, R. L. & Jackson, S. P. Transcriptional regulation of an archaeal operon *in vivo* and *in vitro*. *Mol. Cell* **4**, 971–982 (1999).
- Cohen-Kupiec, R., Blank, C. & Leigh, J. A. Transcriptional regulation in Archaea: *in vivo* demonstration of a repressor binding site in a methanogen. *Proc. Natl Acad. Sci. USA* **94**, 1316–1320 (1997).
- Hofacker, A., Schmitz, K. M., Cichonczyk, A., Sartorius-Neef, S. & Pfeifer, F. GvpE- and GvpD-mediated transcription regulation of the *p-gvp* genes encoding gas vesicles in *Halobacterium salinarum*. *Microbiology* **150**, 1829–1838 (2004).
- Dennis, P. P. Ancient ciphers: translation in Archaea. *Cell* **89**, 1007–1010 (1997).
- Benelli, D., Maone, E. & Londei, P. Two different mechanisms for ribosome/mRNA interaction in archaeal translation initiation. *Mol. Microbiol.* **50**, 635–643 (2003).
- Sandman, K., Krzycki, J. A., Dobrinski, B., Lurz, R. & Reeve, J. N. HMF, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermobacter fervidus*, is most closely related to histones. *Proc. Natl Acad. Sci. USA* **87**, 5788–5791 (1990).
This represents the first demonstration that (eury)archaea package their DNA using histones in a similar way to eukaryotic homologues.
- Reeve, J. N. Archaeal chromatin and transcription. *Mol. Microbiol.* **48**, 587–598 (2003).
- Bell, S. D., Botting, C. H., Wardleworth, B. N., Jackson, S. P. & White, M. F. The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science* **296**, 148–151 (2002).
Histones are not found in Crenarchaeota. This report describes a chromatin protein that is prevalent in Crenarchaeota, and also found in a range of eukaryotic species.
- DiRuggiero, J., Brown, J. R., Bogert, A. P. & Robb, F. T. DNA repair systems in Archaea: mementos from the last universal common ancestor? *J. Mol. Evol.* **49**, 474–484 (1999).
- Bergerat, A. *et al.* An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature* **386**, 414–417 (1997).
A perfect illustration of how archaea can provide vital insights into eukaryotic biology. Homology to an archaeal topoisomerase helped to identify Spo11 as the nuclease responsible for initiating meiotic recombination.
- Seitz, E. M., Brockman, J. P., Sandler, S. J., Clark, A. J. & Kowalczykowski, S. C. RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev.* **12**, 1248–1253 (1998).
- Woods, W. G. & Dylla-Smith, M. L. Construction and analysis of a recombination-deficient (*radA*) mutant of *Haloferax volcanii*. *Mol. Microbiol.* **23**, 791–797 (1997).
- Komori, K. *et al.* Both RadA and RadB are involved in homologous recombination in *Pyrococcus furiosus*. *J. Biol. Chem.* **275**, 33782–33790 (2000).
- Roberts, J. A., Bell, S. D. & White, M. F. An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. *Mol. Microbiol.* **48**, 361–371 (2003).
- McCready, S. & Marcello, L. Repair of UV damage in *Halobacterium salinarum*. *Biochem. Soc. Trans.* **31**, 694–698 (2003).
- Grogan, D. W. The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol.* **8**, 180–185 (2000).
- Makarova, K. S., Aravind, L., Grishin, N. V., Rogozin, I. B. & Koonin, E. V. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* **30**, 482–496 (2002).
- Mylykallio, H. *et al.* Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* **288**, 2212–2215 (2000).
This study shows how archaea use 'eukaryotic' enzymes to replicate a circular chromosome in a high-speed 'bacterial' fashion. As demonstrated in references 38 and 39, some archaea use more than one replication origin per chromosome.
- Matsunaga, F., Fortner, P., Ishino, Y. & Mylykallio, H. *In vivo* interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. *Proc. Natl Acad. Sci. USA* **98**, 11152–11157 (2001).
- Robinson, N. P. *et al.* Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell* **116**, 25–38 (2004).
- Lundgren, M., Andersson, A., Chen, L., Nilsson, P. & Bergander, R. Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *Proc. Natl Acad. Sci. USA* **101**, 7046–7051 (2004).
- Galagan, J. E. *et al.* The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res.* **12**, 532–542 (2002).
- Meuer, J., Kuettnner, H. C., Zhang, J. K., Hedderich, R. & Metcalf, W. W. Genetic analysis of the archaeon *Methanosarcina barkeri* Fusaro reveals a central role for Ech hydrogenase and ferredoxin in methanogenesis and carbon fixation. *Proc. Natl Acad. Sci. USA* **99**, 5632–5637 (2002).
- Verhees, C. H. *et al.* The unique features of glycolytic pathways in Archaea. *Biochem. J.* **375**, 231–246 (2003).
- Makarova, K. S. & Koonin, E. V. Filling a gap in the central metabolism of archaea: prediction of a novel aconitase by comparative-genomic analysis. *FEMS Microbiol. Lett.* **227**, 17–23 (2003).
- Siebers, B. *et al.* Reconstruction of the central carbohydrate metabolism of *Thermoproteus tenax* by use of genomic and biochemical data. *J. Bacteriol.* **186**, 2179–2194 (2004).
- Cline, S. W. & Doolittle, W. F. Efficient transfection of the archaeobacterium *Halobacterium halobium*. *J. Bacteriol.* **169**, 1341–1314 (1987).
- 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 archaeobacterium. *Proc. Natl Acad. Sci. USA* **84**, 8530–8534 (1987).
Along with reference 45, this is the first example of high-efficiency transformation of archaea. The authors went on to develop a series of shuttle vectors that was based on the pHV2 plasmid.

47. Tumbula, D. L., Makula, R. A. & Whitman, W. B. Transformation of *Methanococcus maripaludis* and identification of a Pst I-like restriction system. *FEMS Microbiol. Lett.* **121**, 309–314 (1994).
48. Lucas, S. *et al.* Construction of a shuttle vector for, and spheroplast transformation of, the hyperthermophilic archaeon *Pyrococcus abyssi*. *Appl. Environ. Microbiol.* **68**, 5528–5536 (2002).
49. Patel, G. B., Nash, J. H. E., Agnew, B. J. & Sprott, G. D. Natural and electroporation-mediated transformation of *Methanococcus voltae* protoplasts. *Appl. Environ. Microbiol.* **60**, 903–907 (1994).
50. 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).
51. Carnio, R., Contursi, P., Rossi, M. & Bartolucci, S. An autonomously replicating transforming vector for *Sulfolobus solfataricus*. *J. Bacteriol.* **180**, 3237–3240 (1998).
52. 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).
- This paper describes the development of vital tools that make it possible to do genetic analysis in *Methanosarcina* species. They have been adopted by researchers working on other methanogens.**
53. 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).
54. Bertani, G. & Baresi, L. Genetic transformation in the methanogen *Methanococcus voltae* PS. *J. Bacteriol.* **169**, 2730–2738 (1987).
55. Meile, L., Abendschein, P. & Leisinger, T. Transduction in the archaeobacterium *Methanobacterium thermoautotrophicum* Marburg. *J. Bacteriol.* **172**, 3507–3508 (1990).
56. Bertani, G. Transduction-like gene transfer in the methanogen *Methanococcus voltae*. *J. Bacteriol.* **181**, 2992–3002 (1999).
57. Rosenshine, I., Tchelet, R. & Mevarech, M. The mechanism of DNA transfer in the mating system of an archaeobacterium. *Science* **245**, 1387–1389 (1989).
58. Mevarech, M. & Werczberger, R. Genetic transfer in *Halobacterium volcanii*. *J. Bacteriol.* **162**, 461–462 (1985).
59. Grogan, D. W. Exchange of genetic markers at extremely high temperatures in the archaeon *Sulfolobus acidocaldarius*. *J. Bacteriol.* **178**, 3207–3211 (1996).
60. Schleper, C., Holz, I., Janekovic, D., Murphy, J. & Zillig, W. A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating. *J. Bacteriol.* **177**, 4417–4426 (1995).
61. Ghané, F. & Grogan, D. W. Chromosomal marker exchange in the thermophilic archaeon *Sulfolobus acidocaldarius*: physiological and cellular aspects. *Microbiology* **144**, 1649–1657 (1998).
62. Nolling, J. & de Vos, W. M. Identification of the CTAG-recognizing restriction-modification systems MthZI and MthFI from *Methanobacterium thermoformicum* and characterization of the plasmid-encoded *mthZIM* gene. *Nucleic Acids Res.* **20**, 5047–5052 (1992).
63. Blaseio, U. & Pfeifer, F. Transformation of *Halobacterium halobium*: development of vectors and investigation of gas vesicle synthesis. *Proc. Natl Acad. Sci. USA* **87**, 6772–6776 (1990).
64. Holmes, M. L., Nuttall, S. D. & Dyllal-Smith, M. L. Construction and use of halobacterial shuttle vectors and further studies on *Haloferax* DNA gyrase. *J. Bacteriol.* **173**, 3807–3813 (1991).
65. Grogan, D. W. Cytosine methylation by the *SuaI* restriction-modification system: implications for genetic fidelity in a hyperthermophilic archaeon. *J. Bacteriol.* **185**, 4657–4661 (2003).
66. Holmes, M. L. & Dyllal-Smith, M. L. A plasmid vector with a selectable marker for halophilic archaeobacteria. *J. Bacteriol.* **172**, 756–761 (1990).
67. Holmes, M. L. & Dyllal-Smith, M. L. Mutations in DNA gyrase result in novobiocin resistance in halophilic archaeobacteria. *J. Bacteriol.* **173**, 642–648 (1991).
68. Possot, O., Gernhardt, P., Klein, A. & Sibold, L. Analysis of drug resistance in the archaeobacterium *Methanococcus voltae* with respect to potential use in genetic engineering. *Appl. Environ. Microbiol.* **54**, 734–740 (1988).
69. Gernhardt, P., Possot, O., Foglino, M., Sibold, L. & Klein, A. Construction of an integration vector for use in the archaeobacterium *Methanococcus voltae* and expression of a eubacterial resistance gene. *Mol. Gen. Genet.* **221**, 273–279 (1990).
70. Argyle, J. L., Tumbula, D. L. & Leigh, J. A. Neomycin resistance as a selectable marker in *Methanococcus maripaludis*. *Appl. Environ. Microbiol.* **62**, 4233–4237 (1996).
71. Boccazzi, P., Zhang, J. K. & Metcalf, W. W. Generation of dominant selectable markers for resistance to pseudomonadic acid by cloning and mutagenesis of the *ileS* gene from the archaeon *Methanosarcina barkeri* Fusaro. *J. Bacteriol.* **182**, 2611–2618 (2000).
72. Lam, W. L. & Doolittle, W. F. Shuttle vectors for the archaeobacterium *Halobacterium volcanii*. *Proc. Natl Acad. Sci. USA* **86**, 5478–5482 (1989).
73. Wendoloski, D., Ferrer, C. & Dyllal-Smith, M. L. A new simvastatin (mevinolin)-resistance marker from *Halococcus hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology* **147**, 959–964 (2001).
74. Lange, M. & Ahring, B. K. A comprehensive study into the molecular methodology and molecular biology of methanogenic Archaea. *FEMS Microbiol. Rev.* **25**, 553–571 (2001).
75. Watrin, L., Martin-Jezequel, V. & Prieur, D. Minimal amino acid requirements of the hyperthermophilic archaeon *Pyrococcus abyssi*, isolated from deep-sea hydrothermal vents. *Appl. Environ. Microbiol.* **61**, 1138–1140 (1995).
76. Haydock, A. K., Porat, I., Whitman, W. B. & Leigh, J. A. Continuous culture of *Methanococcus maripaludis* under defined nutrient conditions. *FEMS Microbiol. Lett.* **238**, 85–91 (2004).
77. Peck, R. F., Dassarma, S. & Krebs, M. P. Homologous gene knockout in the archaeon *Halobacterium salinarum* with *ura3* as a counterselectable marker. *Mol. Microbiol.* **35**, 667–676 (2000).
- The first use of a counter-selectable gene knockout system in archaea. This technique has now been widely applied.**
78. Wang, G., Kennedy, S. P., Fasiludeen, S., Rensing, C. & DasSarma, S. Arsenic resistance in *Halobacterium* sp. strain NRC-1 examined by using an improved gene knockout system. *J. Bacteriol.* **186**, 3187–3194 (2004).
79. Bitan-Banin, G., Ortenberg, R. & Mevarech, M. Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the *pyrE* gene. *J. Bacteriol.* **185**, 772–778 (2003).
80. Allers, T., Ngo, H., Mevarech, M. & Lloyd, R. G. Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl. Environ. Microbiol.* **70**, 943–953 (2004).
81. Watrin, L., Lucas, S., Purcarea, C., Legrain, C. & Prieur, D. Isolation and characterization of pyrimidine auxotrophs, and molecular cloning of the *pyrE* gene from the hyperthermophilic archaeon *Pyrococcus abyssi*. *Mol. Gen. Genet.* **262**, 378–381 (1999).
82. Grogan, D. W. Selectable mutant phenotypes of the extremely thermophilic archaeobacterium *Sulfolobus acidocaldarius*. *J. Bacteriol.* **173**, 7725–7727 (1991).
83. Jonuscheit, M., Martusewitsch, E., Stedman, K. M. & Schleper, C. A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Mol. Microbiol.* **48**, 1241–1252 (2003).
- This report describes a sophisticated genetic system for *S. solfataricus*. Genetic studies of hyperthermophiles are relatively rare, but the tools described here should facilitate future investigations.**
84. Martusewitsch, E., Sensen, C. W. & Schleper, C. High spontaneous mutation rate in the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by transposable elements. *J. Bacteriol.* **182**, 2574–2581 (2000).
85. Sapienza, C., Rose, M. R. & Doolittle, W. F. High-frequency genomic rearrangements involving archaeobacterial repeat sequence elements. *Nature* **299**, 182–185 (1982).
86. Worthington, P., Hoang, V., Perez-Pomares, F. & Blum, P. Targeted disruption of the α -amylase gene in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* **185**, 482–488 (2003).
87. Bowen, T. L. & Whitman, W. B. Incorporation of exogenous purines and pyrimidines by *Methanococcus voltae* and isolation of analog-resistant mutants. *Appl. Environ. Microbiol.* **53**, 1822–1826 (1987).
88. Pritchett, M. A., Zhang, J. K. & Metcalf, W. W. Development of a markerless genetic exchange method for *Methanosarcina acetivorans* C2A and its use in construction of new genetic tools for methanogenic archaea. *Appl. Environ. Microbiol.* **70**, 1425–1433 (2004).
89. Moore, B. C. & Leigh, J. A. Markerless mutagenesis in *Methanococcus maripaludis* demonstrates roles for alanine dehydrogenase, alanine racemase, and alanine permease. *J. Bacteriol.* (in the press).
90. Ladapo, J. & Whitman, W. B. Method for isolation of auxotrophs in the methanogenic archaeobacteria: role of the acetyl-CoA pathway of autotrophic CO₂ fixation in *Methanococcus maripaludis*. *Proc. Natl Acad. Sci. USA* **87**, 5598–5602 (1990).
91. Watrin, L. & Prieur, D. UV and ethyl methanesulfonate effects in hyperthermophilic archaea and isolation of auxotrophic mutants of *Pyrococcus* strains. *Curr. Microbiol.* **33**, 377–382 (1996).
92. Blank, C. E., Kessler, P. S. & Leigh, J. A. Genetics in methanogens: transposon insertion mutagenesis of a *Methanococcus maripaludis* *nifH* gene. *J. Bacteriol.* **177**, 5773–5777 (1995).
93. Zhang, J. K., Pritchett, M. A., Lampe, D. J., Robertson, H. M. & Metcalf, W. W. *In vivo* transposon mutagenesis of the methanogenic archaeon *Methanosarcina acetivorans* C2A using a modified version of the insect *mariner*-family transposable element *Himar1*. *Proc. Natl Acad. Sci. USA* **97**, 9665–9670 (2000).
94. Dyllal-Smith, M. L. & Doolittle, W. F. Construction of composite transposons for halophilic Archaea. *Can. J. Microbiol.* **40**, 922–929 (1994).
95. Jarrell, K. F., Bayley, D. P., Florian, V. & Klein, A. Isolation and characterization of insertional mutations in flagellin genes in the archaeon *Methanococcus voltae*. *Mol. Microbiol.* **20**, 657–666 (1996).
96. Thomas, N. A., Mueller, S., Klein, A. & Jarrell, K. F. Mutants in *flaI* and *flaJ* of the archaeon *Methanococcus voltae* are deficient in flagellum assembly. *Mol. Microbiol.* **46**, 879–887 (2002).
97. Berghöfer, Y. & Klein, A. Insertional mutations in the hydrogenase *vhc* and *fcc* operons encoding selenium-free hydrogenases in *Methanococcus voltae*. *Appl. Environ. Microbiol.* **61**, 1770–1775 (1995).
98. Kim, W. & Whitman, W. B. Isolation of acetate auxotrophs of the methane-producing archaeon *Methanococcus maripaludis* by random insertional mutagenesis. *Genetics* **152**, 1429–1437 (1999).
99. Stedman, K. M., Schleper, C., Rumpf, E. & Zillig, W. Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*. Construction and testing of viral shuttle vectors. *Genetics* **152**, 1397–1405 (1999).
100. Eflerink, M. G., Schleper, C. & Zillig, W. Transformation of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus* via a self-spreading vector. *FEMS Microbiol. Lett.* **137**, 31–35 (1996).
101. Krebs, M. P., Hauss, T., Heyn, M. P., RajBhandary, U. L. & Khorana, H. G. Expression of the bacteriopycin gene in *Halobacterium halobium* using a multicopy plasmid. *Proc. Natl Acad. Sci. USA* **88**, 859–863 (1991).
102. Tumbula, D. L., Bowen, T. L. & Whitman, W. B. Characterization of pURB500 from the archaeon *Methanococcus maripaludis* and construction of a shuttle vector. *J. Bacteriol.* **179**, 2976–2986 (1997).
103. Gardner, W. L. & Whitman, W. B. Expression vectors for *Methanococcus maripaludis*: overexpression of acetylhydroxyacid synthase and β -galactosidase. *Genetics* **152**, 1439–1447 (1999).
104. Erauso, G. *et al.* Sequence of plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidence for rolling-circle replication in a hyperthermophile. *J. Bacteriol.* **178**, 3232–3237 (1996).
105. Aagaard, C. *et al.* General vectors for archaeal hyperthermophiles: strategies based on a mobile intron and a plasmid. *FEMS Microbiol. Rev.* **18**, 93–104 (1996).
106. Aravalli, R. N. & Garrett, R. A. Shuttle vectors for hyperthermophilic archaea. *Extremophiles* **1**, 183–191 (1997).
107. Beneke, S., Bestgen, H. & Klein, A. Use of the *Escherichia coli uidA* gene as a reporter in *Methanococcus voltae* for the analysis of the regulatory function of the intergenic region between the operons encoding selenium-free hydrogenases. *Mol. Gen. Genet.* **248**, 225–228 (1995).
108. Holmes, M. L. & Dyllal-Smith, M. L. Sequence and expression of a halobacterial β -galactosidase gene. *Mol. Microbiol.* **36**, 114–122 (2000).
109. Patenge, N., Haase, A., Bolhuis, H. & Oesterhelt, D. The gene for a halophilic β -galactosidase (*bgaH*) of *Haloferax alicantei* as a reporter gene for promoter analyses in *Halobacterium salinarum*. *Mol. Microbiol.* **36**, 105–113 (2000).
110. Gregor, D. & Pfeifer, F. Use of a halobacterial *bgaH* reporter gene to analyse the regulation of gene expression in halophilic archaea. *Microbiology* **147**, 1745–1754 (2001).
111. Pisani, F. M. *et al.* Thermostable β -galactosidase from the archaeobacterium *Sulfolobus solfataricus*. Purification and properties. *Eur. J. Biochem.* **187**, 321–328 (1990).
112. Grogan, D. W. Evidence that β -galactosidase of *Sulfolobus solfataricus* is only one of several activities of a thermostable β -D-glycosidase. *Appl. Environ. Microbiol.* **57**, 1644–1649 (1991).
113. Schleper, C., Roder, R., Singer, T. & Zillig, W. An insertion element of the extremely thermophilic archaeon *Sulfolobus solfataricus* transposes into the endogenous β -galactosidase gene. *Mol. Gen. Genet.* **243**, 91–96 (1994).

114. Sniezko, I., Dobson-Stone, C. & Klein, A. The *treA* gene of *Bacillus subtilis* is a suitable reporter gene for the archaeon *Methanococcus voltae*. *FEMS Microbiol. Lett.* **164**, 237–242 (1998).
115. Reuter, C. J. & Maupin-Furlow, J. Analysis of proteasome-dependent proteolysis in *Haloflex volcanii* cells using short-lived green fluorescent proteins. *Appl. Environ. Microbiol.* (in the press).
116. Baliga, N. S. *et al.* Systems level insights into the stress response to UV radiation in the halophilic archaeon *Halobacterium* NRC-1. *Genome Res.* **14**, 1025–1035 (2004).
117. Zaigler, A., Schuster, S. C. & Soppa, J. Construction and usage of a onefold coverage shotgun DNA microarray to characterize the metabolism of the archaeon *Haloflex volcanii*. *Mol. Microbiol.* **48**, 1089–1105 (2003).
118. Hjort, K. & Bernander, R. Cell cycle regulation in the hyperthermophilic crenarchaeon *Sulfolobus acidocaldarius*. *Mol. Microbiol.* **40**, 225–234 (2001).
119. Sowers, K. R. & Schreier, H. J. Gene transfer systems for the Archaea. *Trends Microbiol.* **7**, 212–219 (1999).
120. Metcalf, W. W. in *Genetic Methods for Diverse Prokaryotes* (eds Smith, M. C. M. & Sockett, R. E.) 277–326 (Academic Press, London, 1999).
121. Kuo, Y.-P., Thompson, D. K., St Jean, A., Charlebois, R. L. & Daniels, C. J. Characterization of two heat shock genes from *Haloflex volcanii*: a model system for transcription regulation in the Archaea. *J. Bacteriol.* **179**, 6318–6324. (1997).
122. Howland, J. L. *The surprising archaea: discovering another domain of life* (Oxford Univ. Press Inc., USA, 2000).
123. Sowers, K. R. & Schreier, H. J. (eds) in *Archaea: a Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1995).
124. Robb, F. T. & Place, A. R. (eds) in *Archaea: a Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1995).
125. DasSarma, S. & Fleischmann, E. M. (eds) in *Archaea: a Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1995).
126. Crick, F. H. C. On protein synthesis. *Symp. Soc. Exp. Biol.* **12**, 138–163 (1958).
127. 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).
128. Huber, H. *et al.* A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**, 63–67 (2002).
- This paper describes *N. equitans*, an obligate symbiont that grows in co-culture with the crenarchaeon *Ignicoccus* spp. The *N. equitans* genome, which is only 491 kb in size, encodes the machinery for information processing but lacks the genes for amino acid, lipid and nucleotide biosynthesis.**
129. Barns, S. M., Delwiche, C. F., Palmer, J. D. & Pace, N. R. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl Acad. Sci. USA* **93**, 9188–9193 (1996).
130. Koonin, E. V., Makarova, K. S. & Aravind, L. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**, 709–742 (2001).
131. Doolittle, W. F. *et al.* How big is the iceberg of which organellar genes in nuclear genomes are but the tip? *Philos. Trans. R. Soc. Lond. B* **358**, 39–57; discussion 57–58 (2003).
132. Kennedy, S. P., Ng, W. V., Salzberg, S. L., Hood, L. & DasSarma, S. Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence. *Genome Res.* **11**, 1641–1650 (2001).
133. Deppenmeier, U. *et al.* The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. *J. Mol. Microbiol. Biotechnol.* **4**, 453–461 (2002).
134. Brown, J. R. & Doolittle, W. F. Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev.* **61**, 456–502 (1997).
135. Boucher, Y., Douady, C. J., Sharma, A. K., Kamekura, M. & Doolittle, W. F. Intragenomic heterogeneity and intergenomic recombination among haloarchaeal rRNA genes. *J. Bacteriol.* **186**, 3980–3990 (2004).
136. Deppenmeier, U. The unique biochemistry of methanogenesis. *Prog. Nucleic Acid Res. Mol. Biol.* **71**, 223–283 (2002).
137. Forterre, P. A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. *Trends Genet.* **18**, 236–237 (2002).
138. Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74 (2004).
- The high-throughput techniques used in this study provide a snapshot of microbial biodiversity, without the bias that is inherent in traditional culture methods. In conjunction with other environmental studies, it indicates that archaea are extant in almost any habitat examined.**
139. Eckburg, P. B., Lepp, P. W. & Relman, D. A. Archaea and their potential role in human disease. *Infect. Immun.* **71**, 591–596 (2003).
140. Lepp, P. W. *et al.* Methanogenic Archaea and human periodontal disease. *Proc. Natl Acad. Sci. USA* **101**, 6176–6181 (2004).
141. Gophna, U., Charlebois, R. L. & Doolittle, W. F. Have archaeal genes contributed to bacterial virulence? *Trends Microbiol.* **12**, 213–219 (2004).
142. Sprott, G. D., Dicaire, C. J., Gurnani, K., Deschatelets, L. A. & Krishnan, L. Liposome adjuvants prepared from the total polar lipids of *Haloflex volcanii*, *Planococcus* spp and *Bacillus firmus* differ in ability to elicit and sustain immune responses. *Vaccine* **22**, 2154–2162 (2004).
143. Hungate, R. E. in *Methods in Microbiology* (eds Norris, J. R. & Ribbons, D. W.) 117–132 (Academic Press, New York, 1969).
144. Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R. & Wolfe, R. S. Methanogens: reevaluation of a unique biological group. *Microbiol. Rev.* **43**, 260–296 (1979).
145. Cline, S. W. & Doolittle, W. F. Transformation of members of the genus *Haloarcula* with shuttle vectors based on *Halobacterium halobium* and *Haloflex volcanii* plasmid replicons. *J. Bacteriol.* **174**, 1076–1080 (1992).
146. Cline, S. W., Lam, W. L., Charlebois, R. L., Schalkwyk, L. C. & Doolittle, W. F. Transformation methods for halophilic archaeobacteria. *Can. J. Microbiol.* **35**, 148–152 (1989).
147. Contursi, P. *et al.* Development of a genetic system for hyperthermophilic Archaea: expression of a moderate thermophilic bacterial alcohol dehydrogenase gene in *Sulfolobus solfataricus*. *FEMS Microbiol. Lett.* **218**, 115–120 (2003).
148. Mankin, A. S., Zyryanova, I. M., Kagrananova, V. K. & Garrett, R. A. Introducing mutations into the single-copy chromosomal 23S rRNA gene of the archaeon *Halobacterium halobium* by using an rRNA operon-based transformation system. *Proc. Natl Acad. Sci. USA* **89**, 6535–6539 (1992).
149. Nuttall, S. D., Deutscher, S. E., Irving, R. A., Serrano-Gomicia, J. A. & Dyal-Smith, M. L. The ShBle resistance determinant from *Streptoalloteichus hindustanus* is expressed in *Haloflex volcanii* and confers resistance to bleomycin. *Biochem. J.* **346**, 251–254 (2000).
150. Cannio, R., Contursi, P., Rossi, M. & Bartolucci, S. Thermoadaptation of a mesophilic hygromycin B phosphotransferase by directed evolution in hyperthermophilic Archaea: selection of a stable genetic marker for DNA transfer into *Sulfolobus solfataricus*. *Extremophiles* **5**, 153–159 (2001).
151. Jenal, U. *et al.* Isoleucyl-tRNA synthetase of *Methanobacterium thermoautotrophicum* Marburg. Cloning of the gene, nucleotide sequence, and localization of a base change conferring resistance to pseudomonadic acid. *J. Biol. Chem.* **266**, 10570–10577 (1991).
152. Sandbeck, K. A. & Leigh, J. A. Recovery of an integration shuttle vector from tandem repeats in *Methanococcus maripaludis*. *Appl. Environ. Microbiol.* **57**, 2762–2763 (1991).
153. de Macario, E. C., Guernini, M., Dugan, C. B. & Macario, A. J. Integration of foreign DNA in an intergenic region of the archaeon *Methanosarcina mazei* without effect on transcription of adjacent genes. *J. Mol. Biol.* **262**, 12–20 (1996).
154. Ortenberg, R., Rozenblatt-Rosen, O. & Mevarech, M. The extremely halophilic archaeon *Haloflex volcanii* has two very different dihydrofolate reductases. *Mol. Microbiol.* **35**, 1493–1505 (2000).
155. Pfeiffer, M., Bestgen, H., Burger, A. & Klein, A. The *vhuU* gene encoding a small subunit of a selenium-containing [NiFe]-hydrogenase in *Methanococcus voltae* appears to be essential for the cell. *Arch. Microbiol.* **170**, 418–426 (1998).
156. Conover, R. K. & Doolittle, W. F. Characterization of a gene involved in histidine biosynthesis in *Halobacterium* (*Haloflex*) *volcanii*: isolation and rapid mapping by transformation of an autotroph with cosmid DNA. *J. Bacteriol.* **172**, 3244–3249 (1990).
157. Danner, S. & Soppa, J. Characterization of the distal promoter element of halobacteria *in vivo* using saturation mutagenesis and selection. *Mol. Microbiol.* **19**, 1265–1276 (1996).

Acknowledgements

We thank Steve Bell, John Leigh, Kevin Sowers, Ed Bolt and numerous other colleagues whose comments have improved the manuscript. We are grateful to John Leigh, Bill Metcalf, Kevin Sowers, Tadayuki Imanaka, David Walsh, Steve Bell, James Chong, Alan Majerik, Friedhelm Pfeiffer, Dieter Oesterhelt and Julie Maupin-Furlow for communicating results before publication, and to Malcolm White, Patrick Forterre, Richard Shand, Joel Querellou, Michel Gouillou, Frank Robb and Cold Spring Harbor Laboratory Press for pictures of archaeal habitats. T.A. is supported by a Royal Society University Research Fellowship.

Competing interests statement

The authors declare that they have no competing financial interests.

Online links

FURTHER INFORMATION

American Type Culture Collection: <http://www.atcc.org>
Archaeal genomes at National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov/genomes/static/a_g.html
Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH: <http://www.dsmz.de/>
Draft archaeal genomes at Oak Ridge National Laboratory: <http://genome.ornl.gov/microbial>
Draft genomes of psychrophilic archaea at University of New South Wales: <http://psycho.bioinformatics.unsw.edu.au/genomes>
Draft *Haloflex volcanii* genome at The Institute for Genomic Research: http://tigrblast.tigr.org/ufmg/index.cgi?database=h_volcanii&seq
Genomes OnLine Database (GOLD): <http://www.genomesonline.org>
Gordon Research Conference on Archaea: Ecology, HaloHandbook: <http://www.microbiol.unimelb.edu.au/micro/staff/mds/HaloHandbook>
***Halobacterium salinarum* genome at Max Planck Gesellschaft:** <http://www.halobex.mpg.de>
Halophile genomes at University of Maryland Biotechnology Institute: <http://zdna2.umbi.umd.edu/~haloweb>
Japan collection of microorganisms: <http://www.jcm.riken.go.jp>
Metabolism & Molecular Biology: <http://www.grc.org/programs/2005/archaea.htm>
National Collections of Industrial, Food and Marine Bacteria: <http://www.ncimb.co.uk>

SUPPLEMENTARY INFORMATION

See online article: S1 (table)
 Access to this links box is available online.