A re-evaluation of the archaeal membrane lipid biosynthetic pathway

Laura Villanueva, Jaap S. Sinninghe Damsté and Stefan Schouten

Abstract | Archaea produce unique membrane lipids in which isoprenoid alkyl chains are bound to glycerol moieties via ether linkages. As cultured representatives of the Archaea have become increasingly available throughout the past decade, archaeal genomic and membrane lipid-composition data have also become available. In this Analysis article, we compare the amino acid sequences of the key enzymes of the archaeal ether-lipid biosynthesis pathway and critically evaluate past studies on the biochemical functions of these enzymes. We propose an alternative archaeal lipid biosynthetic pathway that is based on a 'multiple-key, multiple-lock' mechanism.

In 1977, Woese & Fox[1] proposed a new domain of life, the Archaea (at that time called the Archaebacteria), in addition to the Eubacteria (Bacteria) and the Eukarya, on the basis of their genomic make-up. In addition, the domain Archaea has other traits that distinguishes it from the Bacteria and the Eukarya. One of the most intriguing traits of the Archaea is the unique structure of their membrane lipids[2]. Bacterial and eukaryotic membranes are composed of fatty acid chains that are linked to the glycerol moiety via ester bonds and that are organized in a bilayer structure. By contrast, archaeal membrane lipids are characterized by ether linkages (instead of ester linkages) between the glycerol moiety and the alkyl chains, isoprene-based alkyl chains (instead of acetate-based straight alkyl chains) as building blocks of the apolar side chains and opposite stereochemistry of the glycerol phosphate backbone (that is, *sn*-glycerol-1-phosphate (G1P) rather than *sn*-glycerol-3-phosphate (G3P))[3]. Soon after the discovery of archaeal membrane ether lipids, it was suggested that they could provide an advantage in extreme environments (for example, high temperature, high salinity or extreme pH)[3], on the basis that they are more chemically stable than the ester-linked membrane lipids that are present in the Bacteria and Eukarya[4]. This is probably due to restrictions in the hydrocarbon chain mobility in ether-linked membranes, which might also result in reduced membrane permeability. However, the discovery of ether lipids in ubiquitous mesophilic or neutrophilic archaea that are found in non-extreme settings, such as the ocean[5], where they have been shown to have important roles in global biogeochemical cycles such as the methane cycle[6] and nitrogen cycle[7], suggested that this hypothesis needed to be re-evaluated.

In addition to lacking an exact answer as to why archaea produce ether membrane lipids, we also do not know how they are produced biochemically. Many steps in the archaeal membrane lipid biosynthetic pathway are still unknown and studies have mainly focused on the evolutionary processes that are involved in the differentiation of bacterial and archaeal membranes[2]. Phylogenetic analyses of the enzymes that are involved in the archaeal membrane lipid biosynthetic pathway have been carried out[8,9], but they were limited to the few archaeal genomes that were available at that time. In light of the recent availability of many more archaeal genome sequences, particularly for mesophilic and environmentally important archaea, and the much more detailed information that is available on archaeal membrane-lipid composition[10], it is timely to analyse the relationship between archaeal membrane ether-lipid composition and the enzymes that are involved in their biosynthesis. In this Analysis article, we present the results of our analysis of the amino acid sequences of key biosynthetic enzymes, as well as a critical evaluation of the current concept of the archaeal membrane ether-lipid biosynthetic pathway, which was based on enzymatic studies in specific archaeal isolates. Taken together, the data indicate that our current understanding of the archaeal membrane lipid biosynthesis pathway must be reconsidered.

Distribution of membrane lipids in Archaea

Early studies based on 16S rRNA gene sequences initially supported a deep split within the Archaea, forming two major phyla: the Crenarchaeota and the Euryarchaeota[11]. Culture studies and the analysis of environmental gene

Department of Marine Organic Biogeochemistry, Royal Netherlands Institute for Sea Research (NIOZ), PO Box 59, NL-1790 AB Den Burg, The Netherlands. Correspondence to L.V.
e-mail:
laura.villanueva@nioz.nl
doi:10.1038/nrmicro3260
Published online 7 May 2014

Table 1 | **Distribution of archaeal membrane lipids in different orders of the Euryarchaeota, Crenarchaeota and Thaumarchaeota**

Archaea	Temperature*	pH‡	Metabolism	Archaeol	Extended archaeol	GDGT-0	GDGT-1–4	GDGT-5–8	Cren-archaeol
Euryarchaeota									
Halobacteriales	Mesophile	Neutrophile or alkalophile	Heterotrophy	✓	✓				
Methanosarcinales	Mesophile	Neutrophile	Methanogenesis	✓					
Methanopyrales	Hyperthermophile	Neutrophile	Methanogenesis	✓					
Methanococcales	Mesophile or thermophile	Neutrophile or alkalophile	Methanogenesis	✓					
Thermococcales	Thermophile or hyperthermophile	Neutrophile	Sulphur dependent	✓		✓			
Methanobacteriales	Mesophile or thermophile	Neutrophile	Methanogenesis	✓		✓			
Archaeoglobales	Mesophile or thermophile	Alkalophile	Sulphur dependent			✓			
Methanomicrobiales	Mesophile	Neutrophile	Methanogenesis			✓			
Thermoplasmatales§	Mesophile or thermophile	Acidophile	Sulphur dependent			✓	✓	✓	
Crenarchaeota									
Thermoproteales	Thermophile or hyperthermophile	Neutrophile or acidophile	Sulphur dependent			✓	✓	✓	
Sulfolobales	Thermophile or hyperthermophile	Acidophile	Sulphur dependent			✓	✓	✓	
Acidilobales	Hyperthermophile	Acidophile	Organotroph			✓	✓	✓	
Desulfurococcales	Hyperthermophile	Neutrophile	Sulphur dependent			✓	✓		
Thaumarchaeota									
Cenarchaeales	Mesophile	Neutrophile	Autotroph; ammonia oxidizer			✓	✓		✓
Nitrosopumilales	Mesophile	Neutrophile	Autotroph‖; ammonia oxidizer			✓	✓		✓
Nitrososphaerales	Mesophile or thermophile	Neutrophile	Autotroph; ammonia oxidizer			✓	✓		✓

Information on archaeal membrane lipid distribution taken from REF. 10. GDGT, glycerol dibiphytanyl glycerol tetraether. *Hyperthermophiles have an optimal growth temperature of at least 80 °C; mesophiles grow at a moderate temperature, typically between 20 °C and 45 °C. ‡Neutrophiles grow at almost neutral pH (pH 5–8); alkalophiles grow at alkaline pH (pH >8); acidophiles grow at acidic pH (pH <5). §Deep hydrothermal vent Euryarchaeota 2 (DHVE-2) cluster (*Candidatus* Aciduliprofundum boonei), closely related to the order Thermoplasmatales, synthesize GDGT-0 and GDGT-1 to GDGT-4 (REF. 28). ‖Autotroph, but some reports have suggested possible mixotrophy[22].

sequences indicate that Crenarchaeota mostly consist of hyperthermophiles and thermoacidophiles[12]. Most hyperthermophilic Crenarchaeota have been isolated from geothermally heated soils or waters, sulphur-rich springs or hydrothermal vents, where they mostly obtain their energy from sulphur-containing compounds[13]. Euryarchaeota are abundant in a range of environments and have widely diverse physiological strategies, including halophilic, thermophilic and methanogenic lifestyles[14]. Horizontal gene transfer (HGT) is thought to have been especially important in the evolution of certain members of the Euryarchaeota. For example, Halobacteriales have acquired several genes from bacteria and it has been proposed that HGT transformed a methanogen into the common ancestor of the Halobacteria[15]. The evolution of another order of the Euryarchaeota — the Thermoplasmatales — is believed to have involved extensive HGT from Sulfolobales (which are hyperthermophilic Crenarchaeota) and Bacteria[16,17].

During the past decade, several other archaeal phyla have been discovered — Korarchaeota and Nanoarchaeota[18,19], Thaumarchaeota[20] and the recently proposed 'Aigarchaeota' phylum[21]. Species of the Korarchaeota, Nanoarchaeota and 'Aigarchaeota' have a limited environmental distribution — they are mainly found in hot springs, and their physiology is still unclear (for

Hyperthermophiles
Organisms that have an optimal growth temperature of at least 80 °C.

Thermoacidophiles
Microorganisms that thrive in acidic, sulphur-rich and high-temperature environments. The name is a combination of thermophile and acidophile.

an example, see REF 18). By contrast, Thaumarchaeota are widespread in marine, lacustrine and terrestrial environments, as revealed by environmental genomic analyses[22].

Although there are a broad range of archaeal membrane lipids[23], they typically feature a variation of two main core structures: *sn*-2,3-diphytanyl glycerol diether (also known as archaeol) with phytanyl chains (C_{20}) in a bilayer structure, and *sn*-2,3-dibiphytanyl diglycerol tetraether (also known as glycerol dibiphytanyl glycerol tetraether (GDGT)), in which the two glycerol moieties are connected by two C_{40} isoprenoid chains, enabling the formation of monolayer membranes[24,25]. GDGTs can contain 0–8 cyclopentane moieties (that is, GDGT-x, where x is the number of cyclopentane moieties[10] (TABLE 1)). The presence of these cyclopentane moieties is thought to be essential to maintaining functional membranes and cellular homeostasis in situations of extreme pH or thermal stress; the number of cyclopentane moieties increases as growth temperature increases[26] and pH decreases[27,28].

Comparing archaeal phylogeny with membrane lipid composition reveals that most lipids are not specific for a phylogenetic group[10] (TABLE 1). Only the GDGT crenarchaeol[5], which contains four cyclopentane moieties and a cyclohexane moiety, is considered to be characteristic of the Thaumarchaeota[29], which suggests that the biosynthesis of the cyclohexane moiety is unique for this phylum. GDGTs are the dominant lipid species in Crenarchaeota and Thaumarchaeota, whereas members of the Euryarchaeota synthesize archaeol (for example, Methanococcales, Halobacteriales and Methanosarcinales), GDGTs (for example, Methanopyrales, Thermoplasmatales, Archaeoglobales and Methanomicrobiales) or both (Thermococcales and Methanobacteriales) (TABLE 1). GDGT-0 is found in all (hyper)thermophilic Crenarchaeota and in several thermophilic Euryarchaeota, in some mesophilic methanogenic Euryarchaeota and in Thaumarchaeota. GDGTs that have 1–4 cyclopentane moieties are synthesized by hyperthermophilic Crenarchaeota, Thaumarchaeota, by the thermophilic euryarchaeotal order Thermoplasmatales and by the euryarchaeote '*Candidatus* Aciduliprofundum boonei' (which is a member of the deep hydrothermal vent Euryarchaeota 2 (DHVE-2) cluster and is closely related to the Thermoplasmatales order)[30]. However, GDGTs are apparently not synthesized by methanogenic Euryarchaeota (TABLE 1). GDGTs that have more than four cyclopentane moieties (GDGTs 5–8; TABLE 1) are rare and are only found in hyperthermophilic Crenarchaeota and some hyperthermophilic Euryarchaeota of the order Thermoplasmatales. GDGTs are absent in Halobacteriales (which are Euryarchaeota) and they mainly contain archaeol or extended archaeol with one C_{25} isoprenoid chain[31].

Archaeal lipid synthesis

Previous studies have characterized some of the enzymes that are involved in the biosynthesis of archaeal membrane ether lipids (FIG. 1). Isopentenyl diphosphate and dimethylallyl diphosphate (DMAPP) function as basic building blocks of the isoprenoid chains and are synthesized by the mevalonate pathway[2,32]. DMAPP is thought to be consecutively condensed with several isopentenyl diphosphate units to form geranylgeranyl diphosphate (GGPP; C_{20}) by a short-chain (C_{20}) isoprenyl diphosphate (IPP) synthase, GGPP synthase (FIG. 1). The subsequent ether-bond formation is catalysed by two prenyltransferases: GGPP is attached to the G1P to form geranylgeranylglyceryl phosphate (GGGP), which is catalysed by the GGGP synthase. The attachment of the second side chain to GGGP generates digeranylgeranylglyceryl phosphate (DGGGP); this is catalysed by the DGGGP synthase (FIG. 1). This is thought to be followed — after addition of a polar headgroup to the glycerol moiety — by a reduction of the unsaturated isoprenoid chains, which is mediated by geranylgeranyl reductases[2], to form archaeol. The formation of GDGTs is thought to initially involve the coupling of two archaeol molecules by head-to-head condensation of the phytanyl chains (FIG. 1). Cyclopentane moieties are thought to subsequently be formed by internal cyclization. The last two steps are highly unusual as they involve non-activated methyl groups and the enzymes involved are unknown[2,33].

Evidence for the head-to-head coupling of archaeol comes from pulse-chase experiments using cell extracts of the euryarchaeon *Thermoplasma acidophilum* (of the order Thermoplasmatales) incubated with ^{14}C-mevalonate, which showed the incorporation of radioactivity, first into archaeol and then into GDGT-0 (REF. 34). Furthermore, pulse-chase experiments that were carried out with cell extracts of *T. acidophilum* labelled with ^{14}C-mevalonate showed that using a squalene epoxidase inhibitor (known as terbinafine) led to the accumulation of archaeol with a modified

Figure 1 | **Current understanding of the archaeal lipid biosynthetic pathway.** The two basic building blocks are the five-carbon compound isopentenyl diphosphate and its isomer dimethylallyl diphosphate (DMAPP), which are synthesized by the mevalonate pathway in archaea[2]. DMAPP consecutively condenses with several isopentenyl diphosphate units to form geranylgeranyl diphosphate (GGPP; C_{20}), which is catalysed by an isoprenyl diphosphate synthase, GGPP synthase. Dihydroxyacetone phosphate (DHAP) is converted to glycerol-1-phosphate (G1P). The formation of the two ether bonds between G1P and the GGPP units is catalysed by the geranylgeranylglyceryl phosphate (GGGP) synthase and the digeranylgeranylglyceryl phosphate (DGGGP) synthase. Then, cytidine-diphosphate (CDP)-diglyceride synthase replaces the phosphate group of the unsaturated DGGGP with CDP, generating (unsaturated) CDP-DGGGP; CDP is then replaced with a polar headgroup by a CDP-alcohol phosphatidyl transferase[32]. Saturation of the side chains is thought to be mediated by geranylgeranyl reductases. The formation of GDGTs is thought to involve a head-to-head coupling between the two archaeol lipids, followed by internal cyclization to form cyclopentane moieties. The internal cyclization reactions are highly unusual and the enzymes that are involved are unknown. Figure based on data from REF. 2.

Halophilic
A term used to describe extremophilic organisms that thrive at high concentrations of salt.

Methanogenic
A term used to describe Archaea that produce methane under anoxic conditions.

Horizontal gene transfer
(HGT). The transfer of genetic material between different species of microorganisms; the acquired genes are transmitted to the next generation as the cell divides.

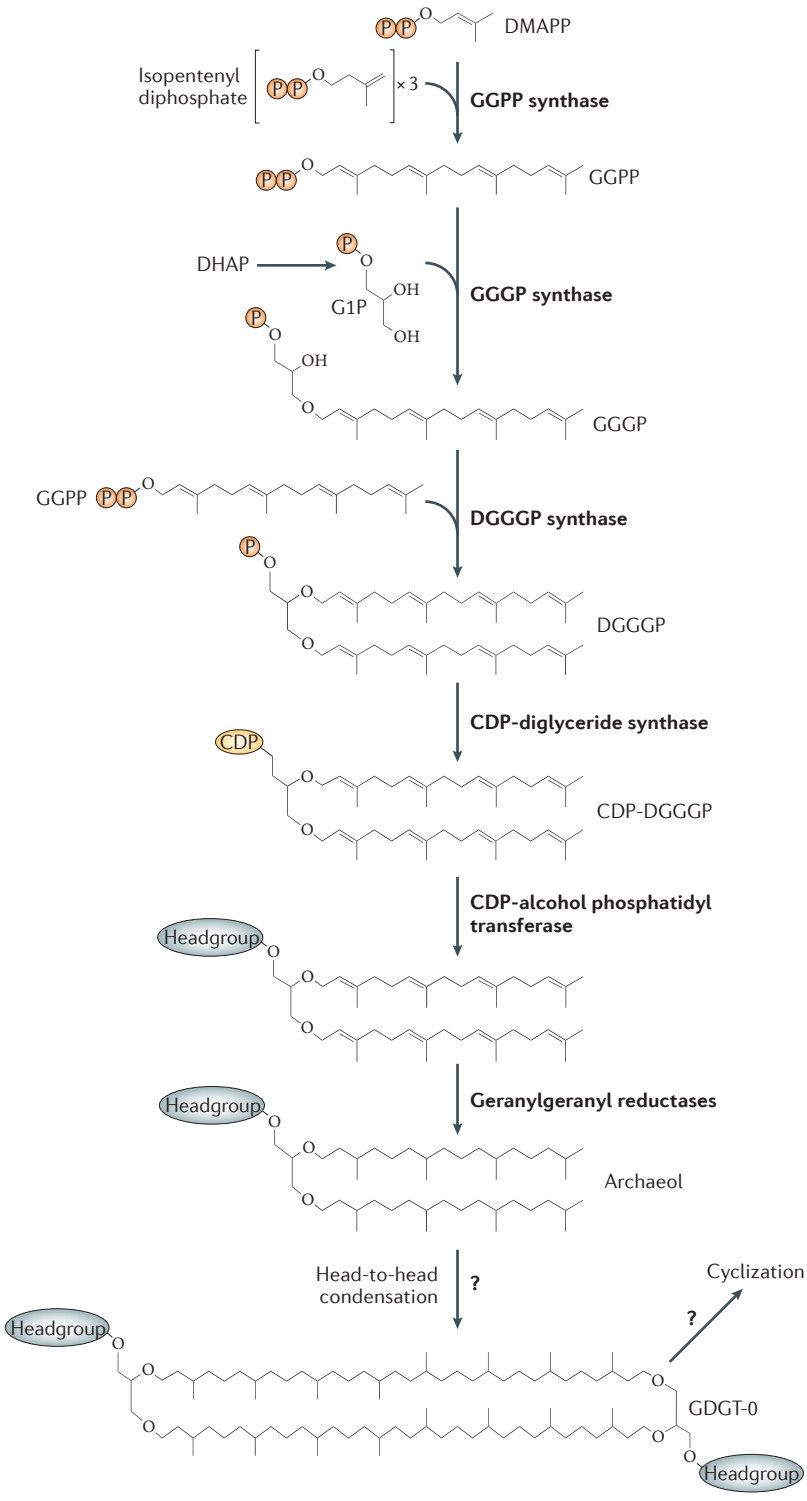
Phytanyl chains
Saturated chains that are composed of four head-to-tail-linked isoprene units (that is, C_{20} isoprenoids).

Isoprenoid
(Also known as isoprene). A term used to describe a group of natural products that have diverse structures composed of various numbers of isopentenyl (C_5) pyrophosphate units.

Prenyltransferases
Enzymes that transfer (iso) prenyl moieties to acceptor molecules.

Head-to-head condensation
The coupling of two isoprenyl units at the C1 position of both units.

Squalene
A biochemical precursor of the steroid and triterpenoid families; it is synthesized by tail-to-tail condensation of farnesyl pyrophosphate (C_{15}) by squalene synthase.



headgroup, rather than GDGTs[35]. These experiments suggest that archaeol is the precursor of GDGTs. However, Poulter *et al.*[36] studied the *in vivo* incorporation of radiolabelled archaeol into cells of the euryarchaeon *Methanospirillum hungatei* (of the order Methanomicrobiales) and found no incorporation of radioactivity in GDGT-0. Furthermore, radiolabelled phytol, which has one double bond, was not incorporated into archaeol or GDGT-0, whereas geranylgeraniol was efficiently incorporated into both archaeol and GDGT-0. Similar results were obtained by incorporating deuterium-labelled DGGGP analogues in *Methanothermobacter thermoautotrophicus* (of the order Methanobacteriales)[37,38]. The deuterium-labelled DGGGP analogues that had a terminal double bond or a saturated terminal isoprene unit were not incorporated into GDGT-0, and only the DGGGP analogue that had a terminal isopropylidene group was incorporated into the GDGT. These studies suggest that the presence of double bonds in the DGGGP molecule is a prerequisite for the formation of GDGTs, which contradicts the idea that fully saturated phytanyl chains are coupled.

In the following section, we focus on three key enzymes known to be involved in the formation of glycerol ether lipids: GGPP synthase, GGGP synthase and DGGGP synthase. We searched for homologues of these enzymes in all archaeal genomes that are currently available, compared the amino acid moieties that are involved in the selection of the substrate and used maximum-likelihood analyses to reveal their phylogeny then compared this with the distribution of ether membrane lipids (TABLE 1).

IPP synthases

IPP synthases catalyse consecutive condensations of isopentenyl diphosphates with allylic primer substrates to form isoprenoid compounds, including steroids, triterpenoids, carotenoids, prenylated proteins and quinones[39]. IPP synthases have two conserved aspartate-rich motifs that are typical of prenyltransferases; these motifs form a deep hydrophobic cleft or substrate-binding pocket[40]. The prenyl groups that are synthesized can be short (that is, C_{10}–C_{20}) or longer (>C_{20}). Short-chain (up to C_{20}) IPP synthases (FIG. 2) are characterized by the presence of bulky amino acids — that is, phenylalanine or tyrosine — as the fifth amino acid residue before the first aspartate-rich motif, which limits the degree of isoprenoid chain elongation to the 20 carbon atoms of the GGPP[40]. Long-chain (>C_{20}) IPP synthases are characterized by a small amino acid (such as alanine, valine or serine in this position). Some IPP synthases are flexible in the chain length they synthesize; for example, the single, bifunctional short-chain IPP synthase of *M. thermoautotrophicus* synthesizes the C_{15} precursor for the synthesis of squalene as well as GGPP (C_{20}) for the synthesis of archaeal membrane lipids[41].

We searched for homologues of IPP synthases in 43 archaeal genomes (Supplementary information S1 (table)). Long-chain IPP synthases were only detected in species of the Thaumarchaeota phylum, in most orders of the Crenarchaeota and in the orders Halobacteriales, Methanosarcinales, Archaeoglobales and Thermoplasmatales of the Euryarchaeota (Supplementary information S1 (table)). The role of the long-chain IPP synthase in these groups is unknown, but it has been hypothesized that it is involved in the synthesis of isoprenoid chains for molecules other than ether lipids[32,42], such as respiratory quinones[43].

Phytol
An acyclic diterpene (terpene consists of two or more isoprene C_5H_8 units) alcohol.

Geranylgeraniol
A diterpenoid alcohol (also known as 3,7,11,15-tetramethyl-2,6,10,14-hexadecatraen-1-ol)

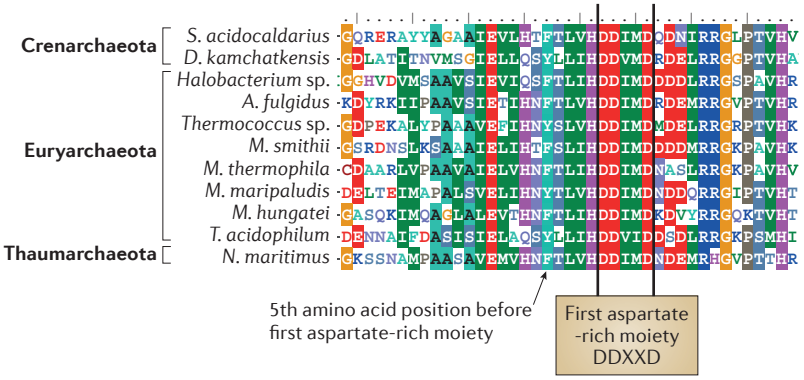


Figure 2 | **Partial IPP synthase protein alignment.** Alignment of amino acid sequences of putative isoprenyl diphosphate (IPP) synthases identified in genomes of different archaeal orders, showing a bulky amino acid residue (such as tyrosine or phenylalanine) at the fifth position before the first aspartate-rich motif, which indicates that they are short-chain IPP synthases elongating the isoprenoid chain up to 20 carbon atoms. Sequences were aligned by MUSCLE (multiple sequence comparison by log-expectation)[53]. Accession numbers are listed in Supplementary information S1 (table). *A. fulgidus*, *Archaeoglobus fulgidus*; *D. kamchatkensis*, *Desulfurococcus kamchatkensis*; *M. hungatei*, *Methanospirillum hungatei*; *M. maripaludis*, *Methanococcus maripaludis*; *M. smithii*, *Methanobrevibacter smithii*; *M. thermophila*, *Methanosaeta thermophila*; *N. maritimus*, *Nitrosopumilus maritimus*; *S. acidocaldarius*, *Sulfolobus acidocaldarius*; *T. acidophilum*, *Thermoplasma acidophilum*.

Putative short-chain IPP synthases were detected in all the archaeal orders (Supplementary information S1 (table)), which suggests that the archaeal lipid biosynthetic pathway starts with the formation of isoprenoid chains that have 20 carbon atoms (that is, GGPPs). According to the current understanding of the archaeal lipid biosynthetic pathway (FIG. 1), short-chain IPP synthases should always encounter the same substrate (that is, isopentenyl diphosphate units) and yield the same product (that is, GGPP). However, the substantial differences between IPP synthases at the amino acid level (FIG. 2) seem to contradict this idea. Rather, the large observed variability in the amino acid sequences of the IPP synthases, and thus the expected plasticity in the structures of these enzymes, suggest that there is structural diversity in the intermediates that are synthesized from isopentenyl diphosphate units.

GGGP synthase

The next step in the proposed biosynthetic pathway consists of the formation of an ether linkage between C3 of the G1P and GGPP to form GGGP (FIG. 1). This step is mediated by the GGGP synthase, which is selective not only for the G1P acceptor but also for the isoprenoid chain that is added, strongly favouring GGPP over shorter or longer chains[44]. GGGP synthase is the first triose phosphate isomerase (TIM)-barrel structure that has been identified to have a prenyltransferase function, which is thought to be unique to the Archaea[45]. GGGP synthase is a homologue of PcrB (heptaprenylglyceryl phosphate (HepGP) synthase), which catalyzes the condensation of G1P with C_{35} heptaprenyl pyrophosphate (HepPP) to HepGP in Gram-positive bacteria (for example, *Bacillus subtilis*)[46].

The only GGGP synthase that has been characterized in detail so far is that of the euryarchaeon *Archaeoglobus fulgidus*[47], which produces archaeol and GDGT-0 as membrane lipids (TABLE 1). The crystal structure of this enzyme shows a unique fold, which functions as a 'greasy slide' and a 'swinging door' owing to the replacement of helix α3 by a strand that creates a large gap to accommodate the product of IPP synthase[45]. It is thought that a bulky hydrophobic amino acid residue — that is, tryptophan — at position 99 (helix α4a of *A. fulgidus*, which is referred to here as the 'chain-length determination area'; FIG. 3), usually marks the end of the gap in the barrel and would presumably select for the chain length of the substrate (in this case presumably GGPP; C_{20}). The GGGP synthase of *A. fulgidus* and PcrB from *B. subtilis* share 35% sequence identity, and the binding sites for G1P are conserved[46] (FIG. 3). Interestingly, the residue that corresponds to the alanine at position 100 (A100) in PcrB from *B. subtilis* and tyrosine 104 (Y104) enable the binding of substrates that are longer than GGPP — that is, $>C_{20}$ (REF. 47). This A100 residue corresponds to tryptophan 99 (W99) in the *A. fulgidus* GGGP synthase (FIG. 3). The conversion of A100 to W100 in PcrB from *B. subtilis* has been shown to prevent the formation of C_{35} products, and the conversion of Y104 to A104 has been shown to enable the formation of longer products of up to C_{40} (REF. 47). Guldan *et al.*[46] also showed that the conversion of W99 to A99 in the *A. fulgidus* GGGP synthase enabled the protein to use substrates that were longer than GGPP.

We searched for GGGP synthase homologues in 72 archaeal genomes and aligned them with the GGGP synthase sequences from *A. fulgidus* (a simplified alignment is shown in FIG. 3). Interestingly, the bulky W99 amino acid residue that is found in *A. fulgidus* GGGP synthase — which is thought to restrict the length of substrates to C_{20} — was only detected in sequences of the euryarchaeotal orders Archaeoglobales, Halobacteriales and Methanomicrobiales, whereas, in the remaining sequences, a small amino acid residue — either glycine or alanine — was found in the corresponding position. This amino acid position also coincides with the A100 residue in PcrB of *B. subtilis*, which enables it to use longer ($>C_{20}$) isoprenyl chains as substrates. Indeed, the secondary structure of the partial amino acid sequences (FIG. 3) showed that the G99 or A99 residue that is observed in most archaeal sequences (other than the euryarchaeotal orders Archaeoglobales, Halobacteriales and Methanomicrobiales) was included in an α-helix structure, as in the case of W99 of *A. fulgidus* (helix α4a (REF. 45)). Moreover, the amino acid sequence alignment of GGGP synthases (FIG. 3) also reveals the presence of a bulky tryptophan residue in the α5′ helix (as defined for *A. fulgidus*) in all the thaumarchaeotal sequences (data not shown), whereas there is a small amino acid residue (such as glycine or alanine) in the corresponding position in the other sequences. In fact, the analysis of protein secondary structure does not predict the existence of an α-helix in this position in the archaeal GGGP synthases other than in *A. fulgidus* and in PcrB of *B. subtilis* (FIG. 3). This amino acid change

Isopropylidene
An isopropyl moiety with a terminal double bond.

Allylic
A term used to describe a double bond at the terminal position of a carbon chain.

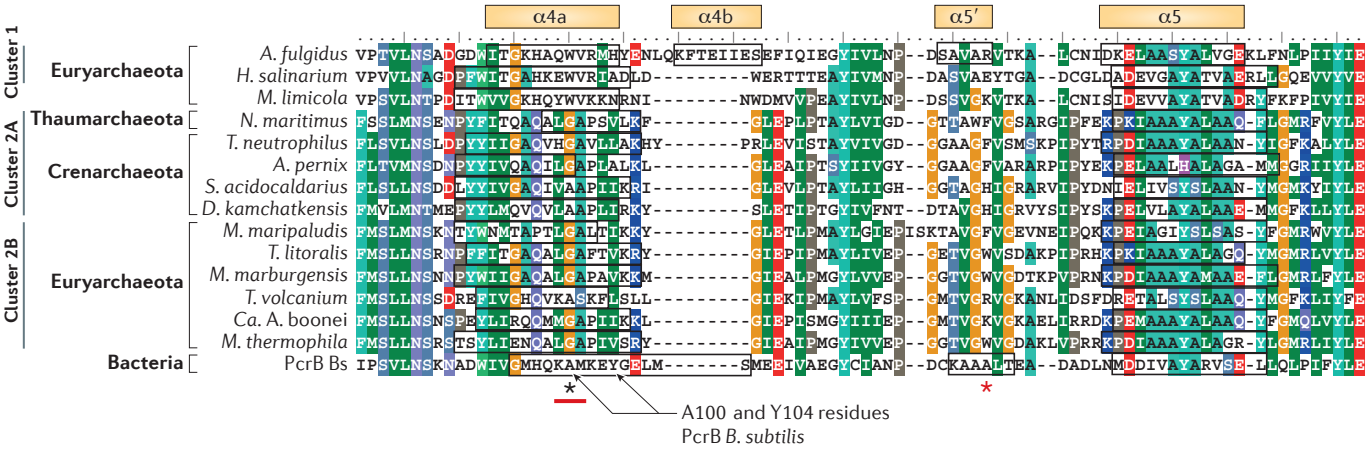


Figure 3 | **Partial GGGP synthase protein alignment.** Annotated putative geranylgeranylglyceryl phosphate (GGGP) synthases of representatives of different archaeal orders are aligned and compared with the GGGP synthase protein of *Archaeoglobus fulgidus*, for which the crystal structure has previously been determined[45]. The black asterisk indicates the position of the amino acid residue corresponding to the tryptophan 99 (W99) position of *A. fulgidus*. The red asterisk indicates the position that contains a tryptophan in the thaumarchaeotal sequences. The chain-length determination area is arbitrary and is indicated (red line) for clarification purposes. The location of α-helixes according to the *A. fulgidus* crystal structure[45] in the partial sequence is indicated above the alignment. The black boxes surrounding amino acid sequences in the alignment correspond to α-helix prediction by the Jpred 3 server[54]. Cluster 1, cluster 2A and cluster 2B correspond to the clusters indicated in FIG. 4. Sequences were aligned by MUSCLE (multiple sequence comparison by log-expectation)[53]. *A. pernix*, *Aeropyrum pernix*; *Ca.* A. boonei, *Candidatus* Aciduliprofundum boonei; *D. kamchatkensis*, *Desulfurococcus kamchatkensis*; *H. salinarium*, *Halobacterium salinarium*; *M. limicola*, *Methanoplanus limicola*; *M. marburgensis*, *Methanothermobacter marburgensis*; *M. maripaludis*, *Methanococcus maripaludis*; *M. thermophila*, *Methanosaeta thermophila*; *N. maritimus*, *Nitrosopumilus maritimus*; PcrB Bs, PcrB protein of *Bacillus subtilis* (NCBI Protein reference YP_007532597.1); *S. acidocaldarius*, *Sulfolobus acidocaldarius*; *T. litoralis*, *Thermococcus litoralis*; *T. neutrophilus*, *Thermoproteus neutrophilus*; *T. volcanicum*, *Thermoplasma volcanicum*.

in the thaumarchaeotal sequences would certainly affect the positioning of the isoprenyl substrate in the GGGP synthase TIM-barrel structure.

These key differences in the amino acid composition of GGGP synthases suggest that their structure and the amino acid interactions between the isoprenyl substrate and the TIM-barrel structure of the GGGP synthase are likely to be quite different from the enzyme that has been characterized in the euryarchaeon *A. fulgidus*. Our analysis of the amino acid sequence diversity of archaeal GGGP synthases strongly suggests that they have functional plasticity and can select substrates that are longer than GGPP.

The phylogeny of GGGP synthase reveals two main clusters (FIG. 4). Cluster 1 includes the GGGP synthase sequences of the euryarchaeotal orders Archaeoglobales, Methanomicrobiales and Halobacteriales, all of which contain the bulky tryptophan residue in the 'chain-length determination area' (FIG. 3). Sequences that are included in cluster 2 (which contain a small amino acid residue, such as glycine or alanine) can be further subdivided into cluster 2A, which includes the sequences of Thaumarchaeota and the Crenarchaeota, and cluster 2B, which includes the sequences of the remaining euryarchaeotal groups (FIG. 4). The large difference between the three euryarchaeotal orders in cluster 1 and the other Archaea (FIG. 4) has previously been related to the presence of an ancestral divergent type of GGGP synthase in Halobacteria[8].

DGGGP synthase

The next step in the proposed pathway consists of the formation of DGGGP from GGGP, which is catalysed by the DGGGP synthase (FIG. 1). DGGGP synthase is a member of the UbiA prenyltransferase family, which, apart from being involved in archaeal ether-lipid formation, also transfers prenyl groups to hydrophobic ring structures such as quinones, haemes, chlorophylls, vitamin E or shikonin[48].

We searched for putative DGGGP synthases in archaeal genomes on the basis of protein homology with the DGGGP synthase of the crenarchaeote *Sulfolobus solfataricus*, the function of which has previously been experimentally tested[48]. Putative DGGGP synthases clustered according to the main archaeal orders and were highly divergent (FIG. 5). Surprisingly, a previous study suggested a lack of homologues of DGGGP synthases in the two Thaumarchaeota genomes that had been released at that time[32]. To confirm this, we investigated the presence of putative DGGGP synthase homologues as well as other members of the UbiA superfamily in currently available thaumarchaeotal genomes. We annotated several putative protohaeme IX farnesyltransferases and other prenyltransferases in thaumarchaeotal genomes on the basis of homology with experimentally tested prenyltransferases[48] (FIG. 5). However, these homologues of the UbiA superfamily in Thaumarchaeota were highly divergent from previously annotated DGGGP synthases in other archaeal orders.

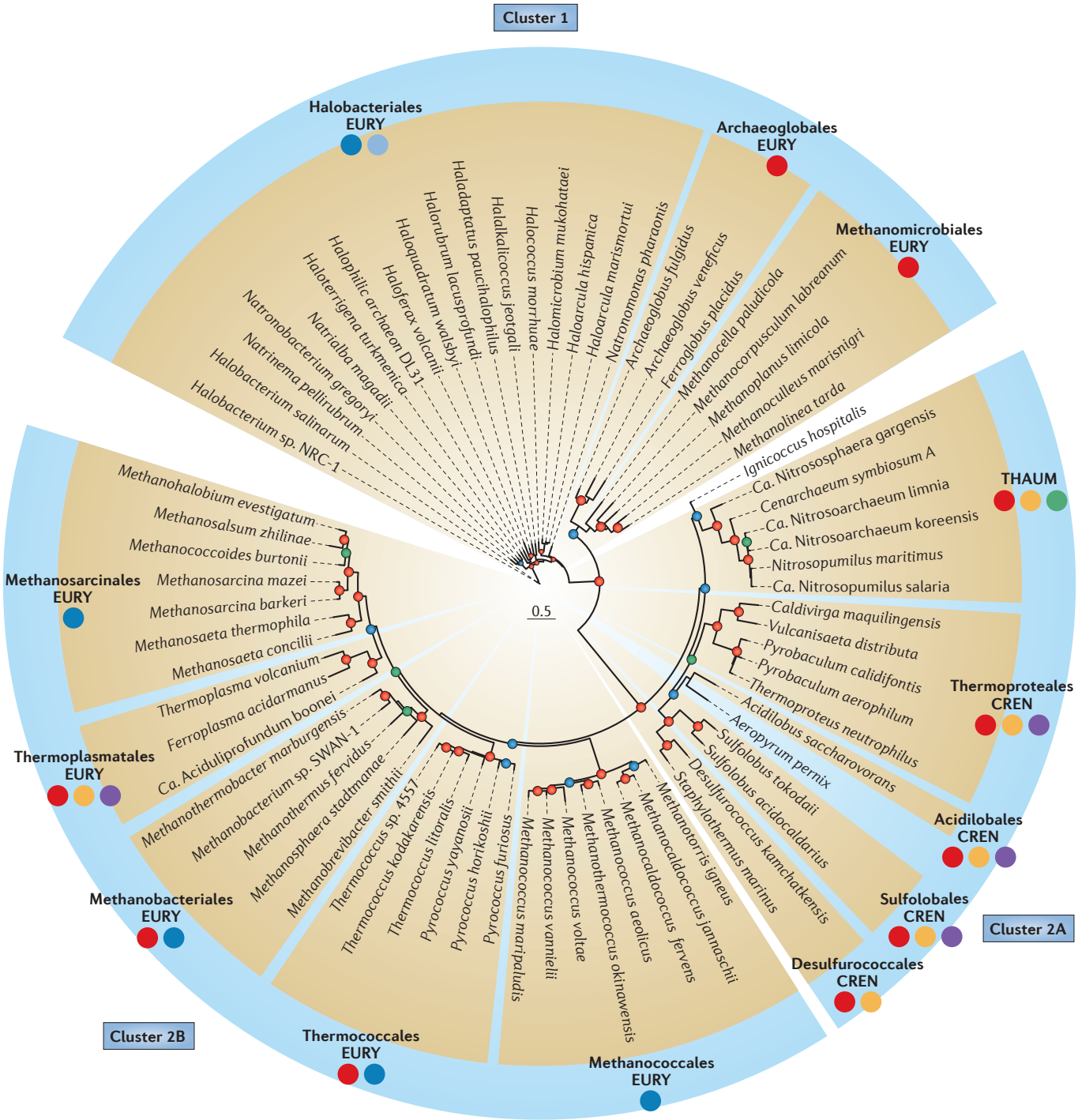


Figure 4 | **Maximum likelihood tree based on the protein sequences of archaeal putative GGGP synthases.** Cluster 1 consists of divergent putative geranylgeranylglyceryl phosphate (GGGP) synthases of the euryarchaeotal orders Halobacteriales, Archaeoglobales and Methanomicrobiales. Cluster 2 is subdivided into cluster 2A, which includes GGGP synthases of the Thaumarchaeota and the Crenarchaeota, and cluster 2B, which includes the remaining GGGP synthases of other euryarchaeotal groups. The scale bar represents number of substitutions per site. The coloured circles indicate the presence of the various membrane lipids (TABLE 1): archaeol, dark blue; extended archaeol, light blue; GDGT-0, red; GDGT-1–4, yellow; GDGT-5–8, purple; crenarchaeol, green. Sequences were aligned using MUSCLE (multiple sequence comparison by log-expectation)[53]. Alignment was trimmed in Gblocks 0.91b with relaxed parameters[55] and manually curated. The phylogenetic tree was computed by PHYML v3.0 (REF. 58) using the LG model plus gamma distribution and invariant site (LG+G+I) indicated by ProtTest 2.4 (REF. 56). Branch support was calculated with the approximate likelihood ratio test (aLRT) and indicated on the branches (colour code in the nodes: red (90%), blue (70%, <90%) and green (50%, <70%), less than 50% is not shown). Trees were edited in iTOL[57]. *Ca.*, *Candidatus*; CREN, Crenarchaeota; EURY, Euryarchaeota; THAUM, Thaumarchaeota.

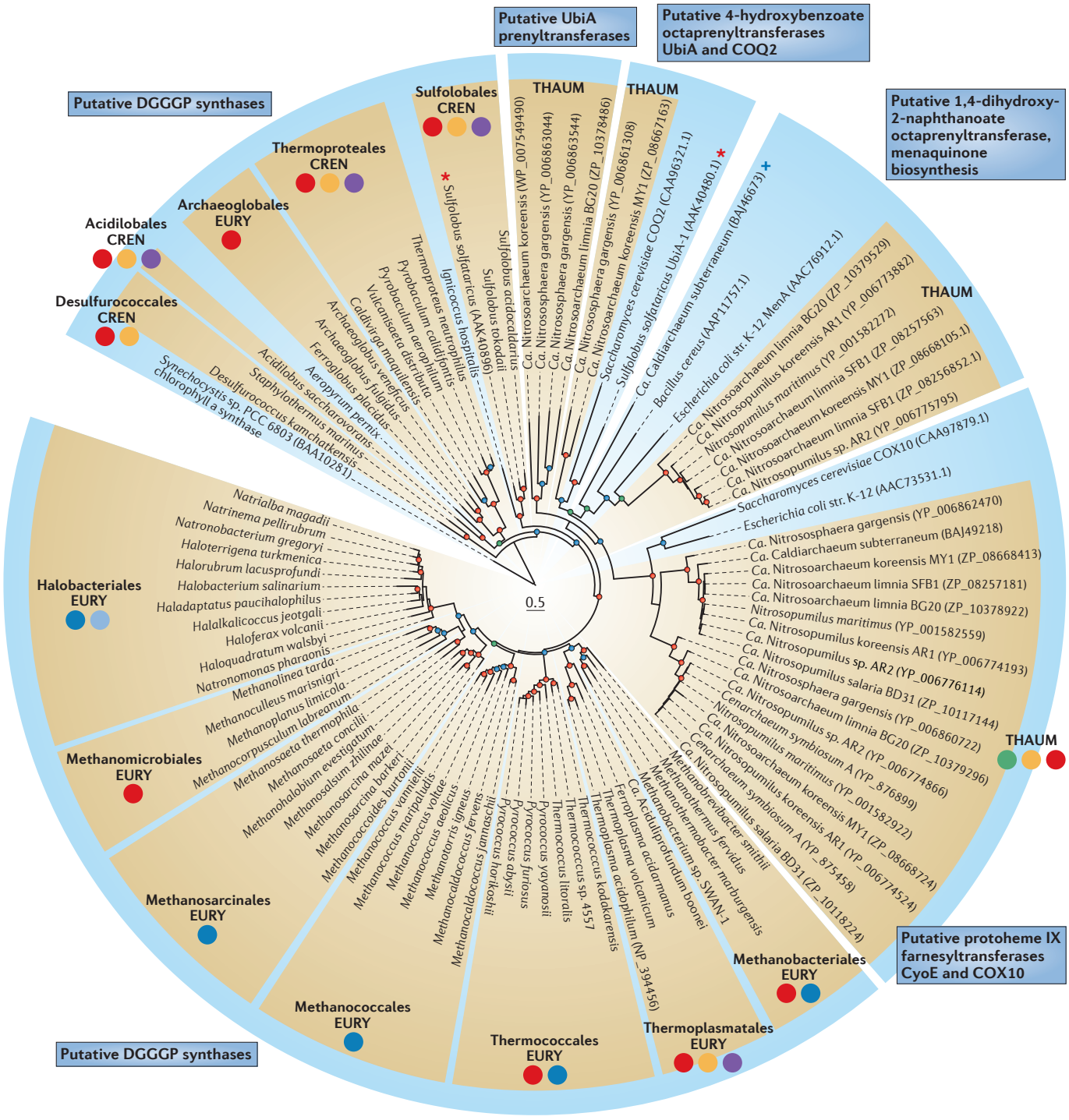
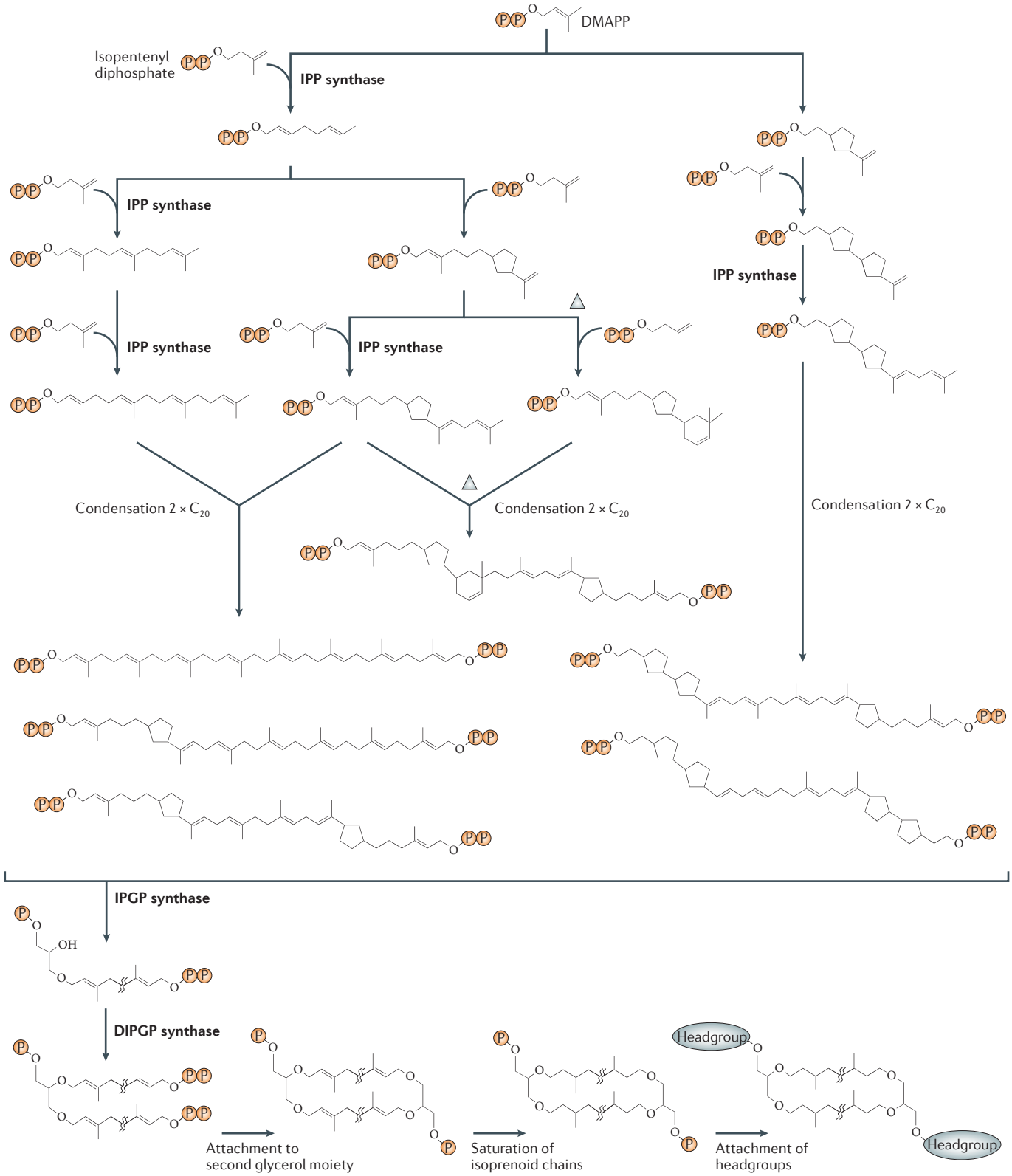


Figure 5 | **Maximum likelihood tree based on the protein sequences of putative archaeal DGGGP synthases and thaumarchaeotal UbiA prenyltransferases.** Putative Thaumarchaeota UbiA prenyltransferases were assigned on the basis of their homology with experimentally tested prenyltransferases[48]. Red asterisks indicate *Sulfolobus solfataricus* DGGGP synthase (NCBI Protein reference AAK40896), and *S. solfataricus* UbiA-1 (NCBI Protein reference AAK40480.1) previously tested[48]. Blue cross indicates *Ca.* Caldiarchaeum subterraneum Aigarchaeota phylum[21]. The scale bar represents number of substitutions per site. The coloured circles indicate the presence of the various membrane lipids (TABLE 1): archaeol, dark blue; extended archaeol, light blue; GDGT-0, red; GDGT-1–4, yellow; GDGT-5–8, purple; crenarchaeol, green. Sequences were aligned using MUSCLE (multiple sequence comparison by log-expectation)[53]. Alignment was trimmed in Gblocks 0.91b with relaxed parameters[50] and manually curated. Phylogenetic tree was computed by PHYML v3.0 (REF. 58) using the LG model plus gamma distribution and invariant site (LG+G+I) indicated by ProtTest 2.4 (REF. 56). Branch support was calculated with the approximate likelihood ratio test (aLRT) and indicated on the branches (colour code in the nodes: red (90%), blue (70%, <90%)and green (50%, <70%), less than 50% is not shown). Trees were edited in iTOL[57]. Ca., *Candidatus*; CREN, Crenarchaeota; EURY, Euryarchaeota; THAUM, Thaumarchaeota.



Biphytanyl
A molecule that is composed of two head-to-head-condensed phytanyl units (C_{40} isoprenoid).

The inability to clearly identify DGGGP synthases in thaumarchaeotal genomes suggests that DGGGP synthases in this phylum are highly divergent compared with others. Interestingly, Thaumarchaeota are the only archaea capable of biosynthesizing GDGTs that contain a cyclohexane moiety (known as crenarchaeol). Damsté *et al.*[5] showed that this additional cyclohexane ring led to a 'bulge' in one of the biphytanyl chains, which

◀ Figure 6 | **An alternative archaeal lipid biosynthesis scheme.** Isoprenyl diphosphate (IPP) synthases generate C_{20} isoprenoid units with or without ring moieties during their catalysis. Grey triangles indicate the introduction of the cyclohexane moiety in the precursor of crenarchaeol in Thaumarchaeota. The condensation of two C_{20} isoprenoid units produces various C_{40} substrates ('multiple keys') that are then used as substrates by isoprenylglyceryl phosphate (IPGP) synthase and di-isoprenylglyceryl phosphate (DIPGP) synthase ('multiple locks'), followed by the attachment of the second glycerol moiety, saturation of the isoprenoid chains and final attachment of the headgroups. Note that the hydrogenation step is indicated here after the assembly of the glycerol dibiphytanyl glycerol tetraether (GDGT) but could potentially also occur before attachment of IPGP to the glycerol moiety. The formation of the cyclohexane moiety in Thaumarchaeota is indicated here during the formation of the C_{20} isoprenoid but this leads to an intermediate that does not have a terminal double bond, potentially inhibiting head-to-head-coupling of C_{20} isoprenyl units . The cyclohexane ring could also be formed after the coupling of two C_{20} prenyl units to a C_{40} prenyl unit. Waved lines indicate condensation of the molecule for artistic purposes.

prevents the dense packing of the biphytanyl chains in the thaumarchaeotal GDGT membranes. It is possible that this 'bulky' biphytanyl chain can only be accommodated by a DGGGP synthase that is rather different from those that use regular biphytanyl chains as substrates.

An alternative pathway for ether-lipid biosynthesis

Our results, together with the sometimes contradictory circumstantial evidence on — for example — which substrates are used in the formation of GDGTs (REF 35 versus REFS 36,38), are difficult to reconcile with the current ether membrane-lipid biosynthetic pathway (FIG. 1). We therefore propose an alternative pathway that better explains our, and earlier[2], observations, but that circumvents some unresolved issues (that is, head-to-head condensation of saturated phytanyl chains and ring formation) in the currently proposed pathway. The new hypothetical pathway is based on a multiple-key, multiple-lock mechanism, in which multiple keys that have different structures, owing to the presence or absence of rings, must accommodate — and specifically interact at the molecular level — with different locks (that is, GGGP synthase and DGGGP synthase) (FIG. 6). The differences in the amino acid sequences of IPP synthases, GGGP synthases and DGGGP synthases indicate that there is greater functional plasticity than previously anticipated. One explanation could be that the rings are already present in the prenyl chains before they are coupled to the glycerol unit (for GGGP synthase and DGGGP synthase). The formation of ring structures at this early stage would avoid the need to form them by internal cyclization of saturated chains. This cyclization may potentially occur simultaneously with chain elongation using isopentenyl diphosphate (FIG. 6).

The presence of small amino acid residues in the chain-length determination area of archaeal GGGP synthases (FIG. 4) indicates that these synthases could accommodate prenyl substrates that are longer than C_{20}. This implies that the substrates of GGGP synthases could be C_{40} prenyl substrates that contain ring moieties. Thus, head-to-head condensation of two C_{20} isoprenyl molecules may occur before attachment to the glycerol unit. These C_{20} isoprenyl molecules contain an isopropylidene double bond, which is required for such condensation[37,38] (except for the unusual C_{20} isoprenyl unit with a cyclohexane moiety that has been hypothesized for Thaumarchaeota) (FIG. 6). This eliminates the need for an unusual (and experimentally poorly supported) condensation of the two saturated phytanyl chains of archaeol (FIG. 1). This proposed head-to-head condensation of two C_{20} isoprenyl molecules could potentially be catalysed by phytoene synthase, which converts two GGPP C_{20} moieties into phytoene (C_{40}) by tail-to-tail coupling in the second step of the biosynthesis of carotenoids[49]. Interestingly, homologues of phytoene synthase have been annotated in archaeal genomes (Supplementary information S2 (table)) of the orders Sulfolobales and Themoproteales of the Crenarchaeota phylum and in the orders Thermoplasmatales, Methanomicrobiales, Methanobacteriales, Methanosarcinales and Halobacteriales of the Euryarchaeota phylum, but not in any of the available genomes of the Thaumarchaeota phylum. The lack of homologues of phytoene synthase in Thaumarchaeota might not be surprising, as the C_{20} intermediate GGGP, which contains the cyclohexane moiety — as hypothesized in our pathway — does not have a terminal isopropylidene moiety (FIG. 6).

After the formation of GGGP, the second IPP unit is attached to the glycerol moiety. The potential presence of ring moieties before the reactions that are catalysed by GGGP synthase and DGGGP synthase would again explain the diversification of the DGGGP synthases that was observed in our study. The apparent lack of the DGGGP synthase-coding gene in genomes of the Thaumarchaeota phylum could be explained by the presence of a more divergent DGGGP synthase that can accommodate the bulky unique cyclohexane moiety of the biphytanyl chain.

Finally, a second glycerol moiety is attached, which is followed by saturation of the isoprenyl chains and the attachment of the headgroup. Considering the alternative pathway that is presented in this Analysis article, we propose to rename the GGGP and DGGGP synthases isoprenylglyceryl phosphate (IPGP) synthase and di-isoprenylglyceryl phosphate (DIPGP) synthase, respectively, in order to reflect the more general nature of these enzymes and their flexibility with respect to the chain length of their substrate (FIG. 6).

The proposed pathway is consistent with our analysis of the sequences of key enzymes of the pathway as well as with most of the experimental evidence for the GDGT biosynthetic steps. Furthermore, the isoprenoid glycerol dialkanol diethers (which are compounds that have C_{40} isoprenoid chains and ring moieties, but attached only to one glycerol group), that were recently detected in archaeal cultures[50,51], as well as the biphytane diols that were detected in the environment[52], are all products of potential intermediates that fit well with our proposed biosynthetic pathway. Clearly, the steps that are proposed in our hypothetical biosynthetic scheme require experimental verification using archaeal cultures, specifically of the Thaumarchaeota phylum. Such results, together with further genomic data mining, will shed further light on the unique membrane lipid pathway of the Archaea.

Phytoene
A C_{40} intermediate in the biosynthesis of carotenoids; it is produced from two molecules of geranylgeranyl pyrophosphate (GGPP) by the action of the enzyme phytoene synthase.

1. Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl Acad. Sci. USA* **74**, 5088–5090 (1977).
2. Koga, Y. & Morii, H. Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol. Mol. Biol. Rev.* **71**, 97–120 (2007).
An extensive review of the different steps in the archaeal ether lipid biosynthetic pathway and the experimental data that support this.
3. Kates, M. Biology of halophilic bacteria, Part, II. Membrane lipids of extreme halophiles: biosynthesis, function and evolutionary significance. *Experientia* **49**, 1027–1036 (1993).
4. Thompson, D. H. *et al.* Tetraether bolaform amphiphiles as models of archae-bacterial membrane lipids: Raman spectroscopy, ^{31}P NMR, X-ray scattering, and electron microscopy. *J. Am. Chem. Soc.* **114**, 9035–9042 (1992).
5. Sinninghe Damsté, J. S., Hopmans, E. C., Schouten, S., van Duin, A. C. T. & Geenevasen, J. A. J. Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota. *J. Lipid Res.* **43**, 1641–1651 (2002).
The first description of crenarchaeol as a characteristic lipid of marine Thaumarchaeota.
6. Jarrell, K. F. *et al.* Major players on the microbial stage: why archaea are important. *Microbiology* **157**, 919–936 (2011).
7. Wuchter, C. *et al.* Archaeal nitrification in the ocean. *Proc. Natl Acad. Sci. USA* **103**, 12317–12322 (2006).
8. Boucher, Y., Kamekura, M. & Doolittle, W. F. Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Mol. Microbiol.* **52**, 515–527 (2004).
9. Lombard, J., Lopez-Garcia, P. & Moreira, D. Phylogenomic investigation of phospholipid synthesis on archaea. *Archaea* **2012**, 630910 (2012).
10. Schouten, S., Hopmans, E. C. & Sinninghe Damsté, J. S. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review. *Org. Geochem.* **54**, 19–61 (2013).
A recent review on the sources, analysis methods and applications of archaeal ether lipids in organic geochemistry.
11. 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).
12. Gribaldo, S. & Brochier-Armanet, C. The origin and evolution of Archaea: a state of the art. *Phil. Trans. R. Soc. B* **361**, 1007–1022 (2006).
13. Stetter, K. O., Fiala, G., Huber, G., Huber, R. & Segerer, A. Hyperthemophilic microorganisms. *FEMS Microbiol. Rev.* **75**, 117–124 (1990).
14. Madsen, E. L. *Environmental Microbiology: From Genomes to Biogeochemistry.* (Wiley-Blackwell, 2008).
15. Nelson-Sathi, S. *et al.* Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc. Natl Acad. Sci. USA* http://dx.doi.org/10.1073/pnas.1209119109 (2012).
16. Ruepp, A. *et al.* The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature* **407**, 508–5139 (2000).
17. Brochier-Armanet, C., Forterre, P. & Gribaldo, S. Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.* **14**, 274–281 (2011).
18. Elkins, J. G. *et al.* A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl Acad. Sci. USA* **105**, 8102–8107 (2008).
19. Huber, H., Hohn, M. J., Stetter, K. O. & Rachel, R. The phylum Nanoarchaeota: present knowledge and future perspectives of a unique form of life. *Res. Microbiol.* **154**, 165–171 (2003).
20. Brochier-Armanet, C., Boussau, B., Gribaldo, S. & Forterre, P. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Rev. Microbiol.* **6**, 245–252 (2008).
Analysis article proposing the classification of Thaumarchaeota as a separate phylum of the Archaea.
21. Nunoura, T. *et al.* Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res.* **39**, 3204–3223 (2011).
22. Pester, M., Schleper, C. & Wagner, M. The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* **14**, 300–306 (2011).
23. Koga, Y. & Morii, H. Recent advances in structural research on ether lipids from Archaea including comparative and physiological aspects. *Biosci. Biotechnol. Biochem.* **69**, 2019–2034 (2005).
24. Koga, Y., Akagawa-Matsushita, M., Ohga, M. & Nishihara, M. Taxonomic significance of the distribution of component parts of polar ether lipids in methanogens. *Syst. Appl. Microbiol.* **16**, 342–351 (1993).
25. Langworthy, T. A. in *The Bacteria* (eds Woese, C. R. & Wolfe, R. S.) 459–497 (Academic Press,1985).
26. Uda, I., Sugai, A., Itoh, Y. H. & Itoh, T. Variation in molecular species of polar lipids from *Thermoplasma acidophilum* depends on growth temperature. *Lipids* **36**, 103–105 (2001).
27. Chong, P. L. G. Archaebacterial bipolar tetraether lipids: physico-chemical and membrane properties. *Chem. Phys. Lip.* **163**, 253–265 (2010).
28. Macalady, J. L. *et al.* Tetraether-linked membrane monolayers in *Ferroplasma* spp: a key to survival in acid. *Extremophiles* **8**, 411–419 (2004).
29. Pitcher, A. *et al.* Core and intact polar glycerol dibiphytanyl glycerol tetraether lipids of ammonia-oxidizing archaea enriched from marine and estuarine sediments. *Appl. Environ. Microbiol.* **77**, 3468–3477 (2011).
30. Reysenbach, A. L. *et al.* Isolation of a ubiquitous obligate thermoacidophilic archaeon from deep-sea hydrothermal vents. *Nature* **442**, 444–447 (2006).
31. De Rosa, M. *et al.* An asymmetric archaebacterial diether lipid from alkaliphilic halophiles. *J. Gen. Microbiol.* **128**, 343–348 (1982).
32. Matsumi, R., Atomi, H., Driessen, A. J. & van der Oost, J. Isoprenoid biosynthesis in archaea — biochemical and evolutionary implications. *Res. Microbiol.* **162**, 39–52 (2011).
33. Lai, D. *Isoprenoid Ether Lipid Biosynthesis in the Extremophile,* Archaeoglobus fulgidus. Thesis, Univ. California Los Angeles (2009).
34. Langworthy, T. A. Turnover of di-*O*-phytanylglycerol in *Thermoplasma*. *Rev. Infect. Dis.* **4**, S266 (1982).
35. Nemoto, N., Shida, Y., Shimada, H., Oshima, T. & Yamagishi, A. Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon *Thermoplasma acidophilum*. *Extremophiles* **7**, 235–243 (2003).
36. Poulter, C. D., Aoki, T. & Daniels, L. Biosynthesis of isoprenoid membranes in the methanogenic archaebacterium *Methanospirillum hungatei*. *J. Am. Chem. Soc.* **110**, 2620–2624 (1988).
37. Eguchi, T., Takyo, H., Morita, M., Kakinuma, K. & Koga, Y. Unusual double-bond migration as plausible key reaction on the synthesis of the isoprenoid membrane lipids of methanogenic archaea. *J. Chem. Soc. Chem. Commun.* **2000**, 1545–1546 (2000).
38. Eguchi, T. Nishimura, Y. & Kakinuma, K. Importance of the isopropylene terminal of geranylgeranyl group for the formation of tetraether lipid in methanogenic archaea. *Tetrahedron Lett.* **44**, 3275–3279 (2003).
39. Wang, K. C. & Ohnuma, S.-I. Isoprenyl diphosphate synthases. *Biochim. Biophys. Acta.* **1529**, 33–48 (2000).
A review of the classification, catalytic mechanism and chain-length determination of IPP synthases.
40. Wang, K. & Ohnuma, S. Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. *Trends Biochem. Sci.* **24**, 445–451 (1999).
41. Chen, A. & Poulter, C. D. Purification and characterization of farnesyl diphosphate/ geranylgeranyl diphosphate synthase. A thermostable bifunctional enzyme from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* **268**, 11002–11007 (1993).
42. Ogawa, T., Yoshimura, T. & Hemmi, H. Geranylfarnesyl diphosphate synthase from *Methanosarcina mazei*: different role, different evolution. *Biochem. Biophys. Res. Commun.* **393**, 16e20 (2010).
43. Peters, K. E., Wlaters, C. C. & Moldowan, J. M. *The biomarker guide: Biomarkers and Isotopes in Petroleum Exploration and Earth History* 2nd edn (Cambridge University Press, 2007).
44. Zhang, D. & Poulter, C. D. Biosynthesis of archaebacterial ether lipids. Formation of ether linkages by prenyltransferases. *J. Am. Chem. Soc.* **115**, 1270–1277 (1993).
45. Payadeh, J., Fujihashi, M., Gillon, W. & Pai, E. F. The crystal structure of (*S*)-3-*O*-geranylgeranylglyceryl phosphate synthase reveals an ancient fold for an ancient enzyme. *J. Biol. Chem.* **281**, 6070–6078 (2006).
First report of the crystal structure of an archaeal GGGP synthase.
46. Guldan, H., Matysik, F.-M., Bocola, M., Sterner, R. & Babinger, P. Functional assignment of an enzyme that catalyzes the synthesis of an Archaea-type ether lipid in Bacteria. *Angew. Chem. Int. Ed.* **50**, 8188–8191 (2011).
47. Ren, F. *et al.* Insights into TIM-barrel prenyl transferase mechanisms: crystal structures of PcrB from *Bacillus subtilis* and *Staphylococcus aureus*. *ChemBioChem*. **14**, 195–199 (2013).
48. Hemmi, H., Shibuya, K., Takahashi, Y., Nakayama, T. & Nishino, T. J. (*S*)-2,3-Di-*O*-geranylgeranylglyceryl phosphate synthase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. Molecular cloning and characterization of a membrane-intrinsic prenyltransferase involved in the biosynthesis of archaeal ether-linked membrane lipids. *Biol. Chem.* **279**, 50197–50203 (2004).
49. Klassen, J. L. Phylogenetic and evolutionary patterns in microbial carotenoid biosynthesis are revealed by comparative genomics. *PLoS ONE* **5**, e11257 (2010).
50. Knappy, C. S. & Keely, B. J. Novel glycerol dialkanol triols in sediments: transformation products of glycerol diphytanyl glycerol tetraether lipids or biosynthetic intermediates? *Chem. Commun.* **48**, 841–843 (2012).
51. Liu, X. -L., Lipp, J. S., Schröder, J. M., Summons, R. E. & Hinrichs, K. -U. Isoprenoid glycerol dialkanol diethers: a series of novel archaeal lipids in marine sediments. *Org. Geochem.* **43**, 50–55 (2012).
52. Schouten, S., Hoefs, M. J. L., Koopmans, M. P., Bosch, H.-J. & Sinninghe Damsté, J. S. Structural characterization, ocurrence and fate of archaeal ether-bound acyclic and cyclic biphytanes and corresponding diols in sediments. *Org. Geochem.* **29**, 1305–1319 (1998).
53. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
54. Cole, C., Barber, J. D. & Barton, G. J. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* **35**, W197–W201 (2008).
55. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).
56. Abascal, F., Zardoya, R. & Posada, D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105 (2005).
57. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).
58. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).

Competing interests statement
The authors declare no competing interests.

DATABASES

NCBI Protein: http://www.ncbi.nlm.nih.gov/protein
YP_007532597.1
AAK40896
AAK40480.1

SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF