

Identification of a protein responsible for the synthesis of archaeal membrane-spanning GDGT lipids

Zhirui Zeng ^{1,2}✉, Huahui Chen¹, Huan Yang³, Yufei Chen¹, Wei Yang ¹, Xi Feng¹, Hongye Pei³ & Paula V. Welander ⁴✉

Glycerol dibiphytanyl glycerol tetraethers (GDGTs) are archaeal monolayer membrane lipids that can provide a competitive advantage in extreme environments. Here, we identify a radical SAM protein, tetraether synthase (Tes), that participates in the synthesis of GDGTs. Attempts to generate a *tes*-deleted mutant in *Sulfolobus acidocaldarius* were unsuccessful, suggesting that the gene is essential in this organism. Heterologous expression of *tes* homologues leads to production of GDGT and structurally related lipids in the methanogen *Methanococcus maripaludis* (which otherwise does not synthesize GDGTs and lacks a *tes* homolog, but produces a putative GDGT precursor, archaeol). *Tes* homologues are encoded in the genomes of many archaea, as well as in some bacteria, in which they might be involved in the synthesis of bacterial branched glycerol dialkyl glycerol tetraethers.

¹Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China. ²Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China. ³State Key Laboratory of Biogeology and Environmental Geology, Hubei Key Laboratory of Critical Zone Evolution, School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China. ⁴Department of Earth System Science, Stanford University, Stanford, CA 94305, USA. ✉email: zengzr@sustech.edu.cn; welander@stanford.edu

One of diagnostic features that distinguishes archaea from bacteria and eukaryotes is the chemical composition of their membrane lipids. Archaeal membranes are mainly composed of two types of isoprenoid ether lipids: glycerol diphytanyl diethers (archaeol) and glycerol dibiphytanyl glycerol tetraethers (GDGTs)¹. Archaeol forms a bilayer membrane that is similar to bacterial and eukaryotic fatty acid bilayers, while GDGTs are membrane-spanning lipids that result in a monolayer membrane². Various archaeal groups contain different proportions of archaeol and GDGTs in their membranes. For instance, haloarchaea and some methanogens are found to produce only archaeol, while thermophilic crenarchaea have GDGTs as their dominant membrane lipids³. From a physicochemical aspect, the formation of monolayers results in greater membrane rigidity and lower permeability which might provide a structural advantage to archaeal cells in extreme environments⁴. Culture experiments indicate that archaea elevate the proportion of tetraether over diether lipids (GDGT/archaeol ratio) under high growth temperature to warrant membrane thermal stability^{5,6}. This observation is also in agreement with liposome studies showing that GDGT-based liposomal membranes are more stable than archaeol-based membranes⁷. In addition, liposomal transport studies show that GDGTs significantly reduce the membrane permeability of water, proton, ammonia, urea, and glycerol compared to archaeol^{7,8}. This low membrane permeability is proposed to provide an energetic advantage for archaea in chronic energy stress environments².

Some archaea can further modify the GDGT structure by introducing cyclopentane and cyclohexane rings, hydroxyl or methyl moieties, altering the polar head groups, or cross-linking the two biphytanyl chains³. Such structure variability contributes to the overall complexity of the archaea lipidome, allowing GDGT profiles to be taxonomic markers for archaeal community analyses⁹. Moreover, GDGT modification generally provides additional rigidity or flexibility to adapt to environmental fluctuations, and these lipid modifications, in turn, can reflect environmental conditions^{10–13}. Because fossilized GDGTs can be well preserved in sediments over a 100-million years old, these GDGT patterns (GDGT structures and their contents) can serve as paleoproxies that record environmental change over large time scales^{14,15}.

Identification of the GDGT biosynthesis pathway is fundamental to the study of the physiological function of GDGTs as well as to the interpretation of GDGT-based biomarkers. Currently, the biochemical mechanism and the protein(s) required in the key step of core GDGT formation, from diether to tetraether, remain unknown¹⁶. In this study, we utilized a comparative genomics approach coupled with gene heterologous expression and lipid analyses to identify a GDGT synthase protein (tetraether synthase, Tes). Our results demonstrate that heterologous expression of Tes candidates in *Methanococcus maripaludis*, which does not produce GDGTs, resulted in the formation of GDGT, glycerol trialkyl glycerol tetraether (GTGT), and macrocyclic archaeol. Further, bioinformatics analyses demonstrated that Tes homologs are widely distributed in archaeal phyla, and are also present in some bacterial groups, possibly shedding light into the biosynthesis of the mysterious bacterial branched glycerol dialkyl glycerol tetraethers (brGDGTs).

Results

Identification of a GDGT synthesis protein in methanogens.

We hypothesized that a potential tetraether synthase (Tes) that catalyzes the condensation of archaeol to GDGT would be a radical S-adenosylmethionine (SAM) protein, because radical SAM proteins are capable of catalyzing challenging C–C bond

formation¹⁷. Previous studies have shown that this class of proteins are required for GDGT modifications such as cyclization and calditol head group formation in *Sulfolobus acidocaldarius*^{18,19}, and are likely involved in the methylation of GDGT-related lipids in methanogens²⁰. In addition, C–C bond formation of the bacterial membrane-spanning lipid diabolic acid has been proposed to occur through a radical mechanism²¹. We searched the *S. acidocaldarius* genome for all annotated radical SAM proteins (Pfam: PF04055) and, among the 18 proteins identified, Saci_0703 was the most promising candidate. Firstly, homologs of this protein (e -value $< 1e^{-50}$; identity $> 30\%$) are only present in the genomes of archaea that produce GDGTs and absent from those that do not. Secondly, the gene that encodes the Saci_0703 homolog in *Methanosarcina acetivorans* C2A (MA_1486) is localized next to a ferredoxin (MA_1485) required to activate geranylgeranyl reductase (GGR; MA_1484)²², which catalyzes the saturation of ether lipid double bonds to form archaeol²³, a putative precursor of GDGT. Thirdly, the *saci_0703* gene is highly expressed under all growth conditions in *S. acidocaldarius* (Supplementary Table 1). As GDGTs are the major membrane lipids present in *S. acidocaldarius*²⁴, high expression of any GDGT synthesis gene would be necessary to meet this need.

We first attempted to identify the function of Saci_0703 through gene deletion analysis in *S. acidocaldarius*. However, we were unable to generate a *saci_0703* gene deletion mutant suggesting that this gene may be essential for growth. We next attempted heterologous expression of *saci_0703* in the methanogen *Methanococcus maripaludis*, which does not synthesize GDGTs, does not contain a Saci_0703 homolog, and produces the putative GDGT precursor archaeol²⁵. Yet, the expression of *saci_0703* in *M. maripaludis* did not produce any GDGTs, perhaps due to the codon bias of expressing a Crenarchaeota gene in a Euryarchaeota host or to a lack of expression of a protein from a thermophile in a mesophile. We decided to express Saci_0703 homologs from more closely related organisms to the *M. maripaludis* host and searched for homologs in 322 methanogen genomes in the Joint Genome Institute (JGI) Integrated Microbial Genomes (IMG) database. We found that 172 genomes contain one Saci_0703 homolog, 22 genomes contain two copies, and 128 genomes do not contain any homologs (e -value $< 1e^{-50}$; identity $> 30\%$). We chose to express one homolog from *Methanococcus aeolicus* Nankai-3 (Maeo_0574), and two homologs from *M. acetivorans* C2A (MA_1486 and MA_1114).

The methanogen genes encoding each hypothesized Tes homolog were expressed individually on a self-replicating plasmid in *M. maripaludis*. Expression of *maeo_0574* or *ma_1486* alone resulted in the production of GDGT (Fig. 1), indicating that the Tes protein is necessary to synthesize GDGTs. In addition to GDGT production, expression of these two Tes homologs in *M. maripaludis* also generated glycerol trialkyl glycerol tetraether (GTGT) and macrocyclic archaeol (Fig. 2a and Supplementary Fig. 2). However, expression of *ma_1114*, the second Tes homolog in the *M. acetivorans* genome, did not result in GDGT production but did produce trace amounts of macrocyclic archaeol. Further lipid quantification showed that expression of *maeo_0574* or *ma_1486* produced tetraether lipids accounting for approximately 1.42% and 0.98% of total membrane lipids, respectively. The strain expressing *ma_1114* only produced macrocyclic archaeol at approximately 0.03% of total membrane lipids (Fig. 3). It is currently unclear if the differences observed in GDGT production by the various Tes homologs tested is due to functional or biochemical variations between homologs or results from differences in expression level of each protein in the heterologous system.

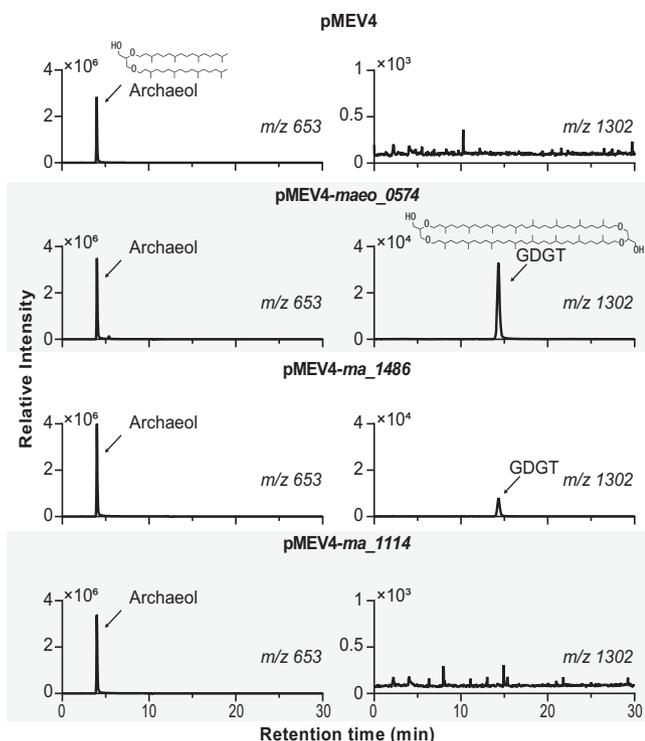


Fig. 1 Tes (Maeo_0574 or MA_1486) is sufficient for GDGT synthesis.

LC-MS extracted ion chromatograms of acid hydrolyzed lipid extracts of *M. maripaludis* wild type (WT) with empty plasmid pMEV4 or *ma_1114* heterologous expression showing only production of archaeol, and *M. maripaludis* WT with *maeo_0574* or *ma_1486* heterologous expression producing GDGT. Mass spectra are shown in Supplementary Fig. 1. Source data are provided as a Source Data file.

The structure of GTGT is like that of GDGT, except one of the C₄₀-biphytanyl chains is replaced by two C₂₀-phytanyl chains (Fig. 2). GTGT had been proposed as an intermediate in GDGT biosynthesis¹, a consequence of incomplete side chain condensation. The co-occurrence of GTGT and GDGT when we heterologously expressed Tes suggests that this GTGT intermediate hypothesis may be correct (Fig. 2b). The production of macrocyclic archaeol was unexpected, as this lipid has only been detected in a few archaeal species^{26–28}. To confirm the structure of the macrocyclic archaeol, we purified this compound from *M. maripaludis* expressing *maeo_0574*. We subsequently cleaved the ether groups through an ether cleavage reaction and analyzed the resulting products through gas chromatography-mass spectrometry (GC-MS) (Supplementary Fig. 1). The detection of biphytane with C₄₀ chains rather than the C₂₀-phytane, indicates macrocyclic archaeol was indeed synthesized by Tes. We propose that macrocyclic archaeol is a side product of the heterologous expression of Tes homologs in *M. maripaludis* (Fig. 2a). Together, the biosynthesis of GDGT, GTGT, and macrocyclic archaeol suggest that Tes can introduce head-to-head linkages between two phytanyl chains of archaeol.

We next wanted to determine if Tes had any potential transmembrane helices as it could provide some insight into whether this was a membrane-bound protein and where the Tes condensation reaction may be occurring in the cell. We first used Phyre2²⁹ to predict whether four Tes homologs, Saci_0703, Maeo_0574, MA_1486, and MA_1114, had any transmembrane helices. Phyre2 predicted that the Saci_0703 and MA_1486 homologs contain one transmembrane helix but Maeo_0574 and MA_1114 do not contain any. We also searched for

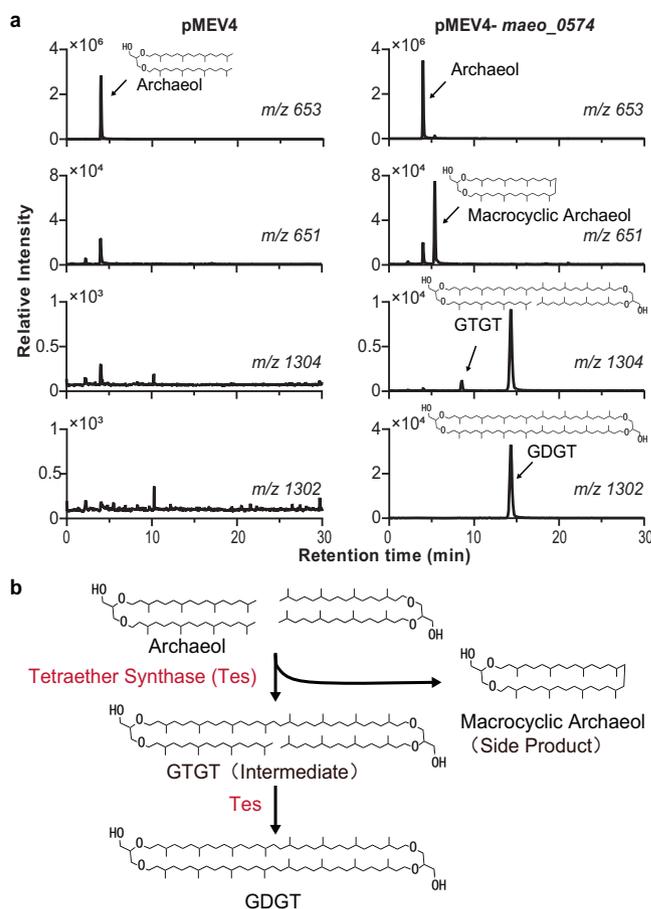


Fig. 2 GDGT biosynthesis pathway, intermediate, and side product. **a**

LC-MS extracted ion chromatograms of acid hydrolyzed extracts from heterologous expression strains showing that Tes (*maeo_0574*) can generate not only GDGT, but also the potential biosynthetic intermediate GTGT, and the side product macrocyclic archaeol. **b** Proposed GDGT biosynthesis pathway based on the Tes heterologous expression products. The chromatograms of all Tes homologs expression strains are shown in Supplementary Fig. 2, and mass spectra are shown in Supplementary Fig. 1. Notice that the double bonds on archaeol may be required for the condensation reaction (see discussion). Source data are provided as a Source Data file.

transmembrane helices through TMHMM 2.0³⁰, which predicted that none of these proteins had transmembrane helices. Thus, it remains unclear whether these proteins are integral membrane proteins, membrane-associated proteins, or cytoplasmic proteins. Given their role in modifying key membrane lipids in archaea, we hypothesize that Tes homologs are at least membrane associated but this, of course, requires biochemical and structural characterization.

Distribution of Tes in microbial genomes. Understanding the predominant biological source of GDGTs in modern environments is one of the fundamental constraints necessary for the proper interpretation of GDGT-based proxies, such as the TEX₈₆ (tetraether index of 86 carbons) sea-surface temperature proxy¹⁵. Our identification of the Tes protein allows us to better understand the distribution of GDGT producers through the analyses of genomic and metagenomic databases. To do so, we performed a Basic Local Alignment Search Tool (BLASTP) search (*e*-value < 1e⁻⁵⁰; percent identity >30%) of protein sequences to identify Tes homologs in the National Center for Biotechnology

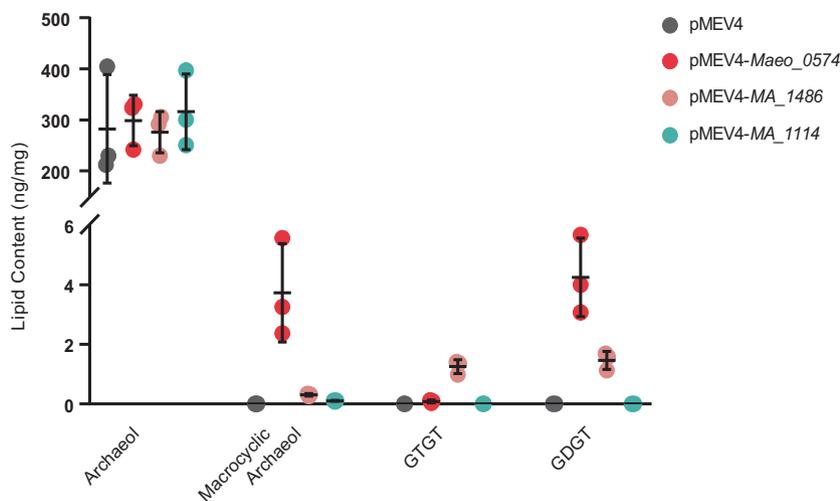


Fig. 3 Quantification of lipids formation during Tes expression in *M. maripaludis*. The abundance of core GDGT-related lipids (ng/mg dry cell mass) from heterologous expression of Tes genes *maeo_0574*, *ma_1486*, or *ma_1114*, showing *maeo_0574* is more efficient to synthesize GDGT and related lipids than the other homologs. The abundance of core lipids was calculated relative to the peak area of the synthetic C46 GTGT internal standard (5). Results are the mean ± standard error from three biological replicates. Source data are provided as a Source Data file.

Information (NCBI) database. Using the *M. acetivorans* MA_1486 homolog as a search query, we found that Tes homologs are widely distributed in all three archaeal superphyla (Asgard, TACK, and DPANN) and the Euryarchaeota. Only 3 (Huberarchaeota, Parvarchaeota, and Pacearchaeota) out of 27 archaeal phyla³¹ do not contain a Tes homolog (Fig. 4 and Supplementary Figs. 3–5 and Supplementary Table 2) suggesting these archaea do not produce GDGTs.

More relevant to the TEX₈₆ paleotemperature proxy, we searched for Tes homologs in the metagenome-assembled genomes (MAGs) of Marine Group II (MG-II) and Marine Group III (MG-III) Euryarchaeota. The major producers of GDGTs in marine water columns have been suggested to be the Marine Group I (MG-I) Thaumarchaeota – providing a critical source constraint for TEX₈₆ paleoproxies¹⁹. However, 16S rRNA analyses of various open ocean settings have identified the MG-II Euryarchaeota as another potential source of GDGTs in the upper water column which could skew the TEX₈₆ signal³². Our bioinformatic search for Tes homologs in MG-II archaea revealed that only 1 MG-II MAG (NCBI Genome ID: GCA_002457195.1) out of 231 MG-II MAGs was found to have a Tes homolog. Further, we did not identify any Tes homologs in Marine Group III (MG-III) MAGs, another group of uncultured marine Euryarchaeota generally considered low-abundance members of deep mesopelagic and bathypelagic communities³³. These data suggest that both MG-II and MG-III do not produce any GDGTs and are most likely not a significant contributor to open ocean GDGT pools.

Interestingly, 20 Halobacteria genomes, for instance, *Halobellus clavatus* and *Halorientalis regularis*, were found to have Tes homologs. However, previously published lipid analyses of these strains did not find any GDGT production³⁴ so the function of these Halobacterial Tes homologs remains unclear. We also

observed Tes homologs in the MAGs from the DPANN archaea which were unexpected as most archaea belonging to this group are thought to be symbionts whose genomes lack lipid biosynthesis genes³⁵. The one cultivated DPANN species, *Nanoarchaeum equitans*, has been found to contain GDGTs but these were thought to be acquired from its Crenarchaeota host, *Ignicoccus hospitalis*³⁶. In support of this, we find that the *N. equitans* genome does not harbor a Tes homolog while the *I. hospitalis* genome does. Given this pattern, it is possible that Tes homologs identified in DPANN MAGs may belong to their symbiotic host.

Finally, we also identified 456 Tes homologs in the Bacterial domain such as the Acidobacteria and various Proteobacteria (Fig. 4 and Supplementary Fig. 4). To date, no bacteria have been shown to produce GDGTs so it would seem unlikely that these bacterial Tes homologs are condensing diethers to GDGT. Rather we hypothesize that these bacterial homologs may be responsible for the synthesis of membrane-spanning lipids in bacteria such as the ether-linked brGDGTs³⁷. Geochemists first identified brGDGTs in soils and peats³⁸ but they have also been found in a variety of terrestrial and marine environments^{39,40}. Bacterial brGDGTs are unique lipids in that they harbor a mixture of archaeal and bacterial properties. Although the ether linkages to glycerol along with the monolayer structure and cyclopentane moieties are reminiscent of archaeal membranes, the alkyl moieties of brGDGTs are not isoprenoidal but rather composed of branched carbon chains suggesting a bacterial source^{38,41}. Acidobacteria are thought to be one source of brGDGTs and have been shown to produce small amounts of brGDGTs with no rings⁴². A recent study confirmed that the Acidobacterial strain, *Edaphobacter aggregans*, produces brGDGTs when grown under conditions of oxygen limitation⁴³. We found that *E. aggregans* contains one Tes homolog (identity = 32%, *e*-value = 4e⁻⁶¹) in

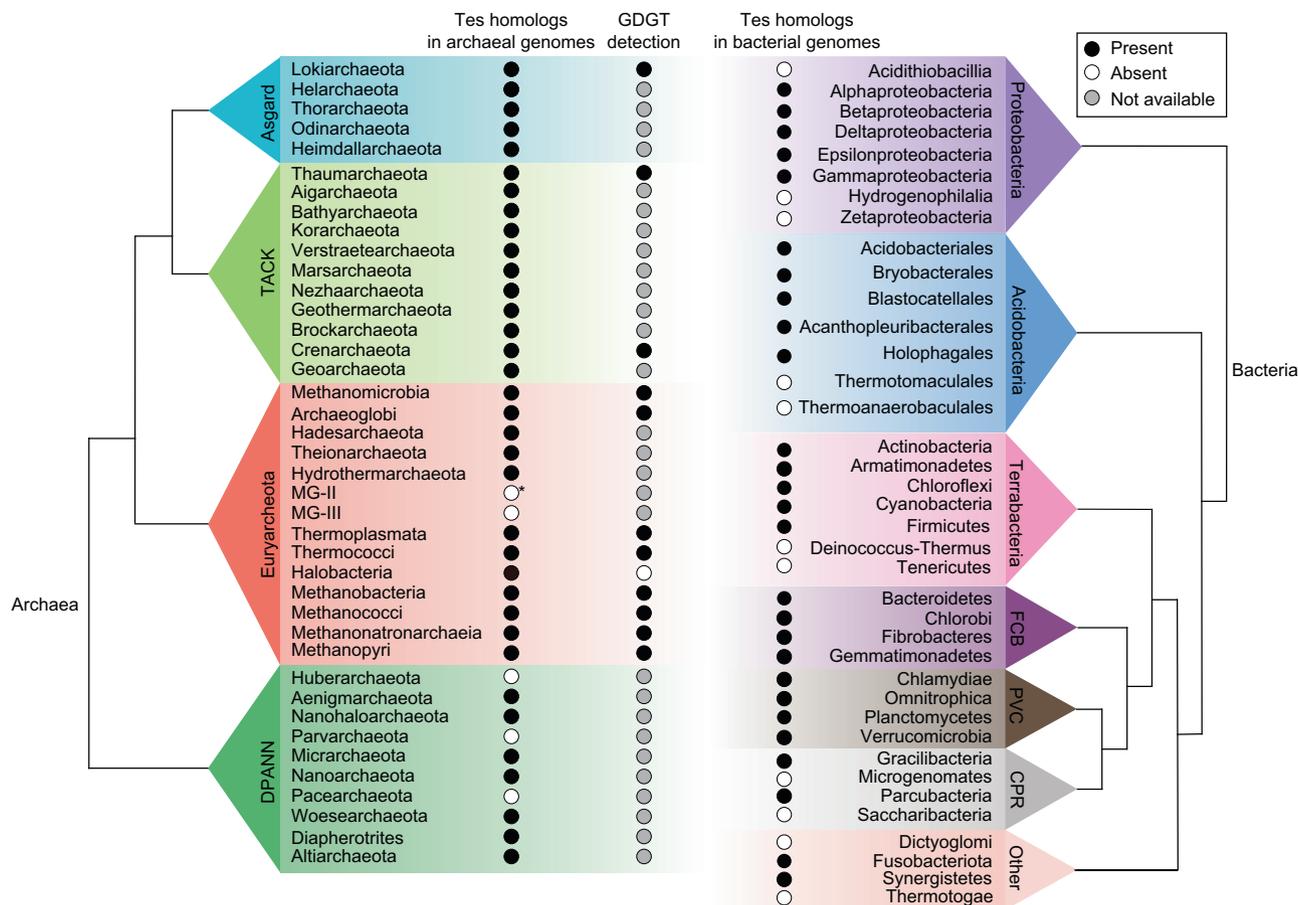


Fig. 4 Tes homologs are found in diverse archaeal and bacterial phyla. These data were generated through BLASTP searches of Tes homologs (e -value $< 1e^{-50}$, identity $> 20\%$) in the NCBI database and searching published studies demonstrating GDGT production from pure cultures^{3,5,9,10,36,69-71}. At least one strain or one genome in the phylum, class or order containing the Tes homolog and GDGTs (at least GDGT-0) is marked with a black circle. Otherwise, archaea or bacteria without any Tes homolog or GDGTs are marked with white circles. Uncultured archaeal strains or with cultured strains not yet tested for GDGT production, are marked with gray circles. Only one putative Tes homolog (e -value = $6e^{-53}$, identity = 28%) was identified from 231 MG-II MAGs (*). Phylogenetic lineages of archaea and bacteria are modified from references^{31,72,73}. Proteobacteria and acidobacteria are divided into class level and order level, respectively. Terrabacteria, FCB (Fibrobacteres-Chlorobi-Bacteroidetes), PVC (Planctomycetes-Verrucomicrobia-Chlamydiae) and CPR (Candidate Phyla Radiation) are bacterial superphyla^{72,73} and only several phyla are shown in each group. Bacterial phyla that do not belong to these groups are assigned “Other” in the figure. Details and references are listed in Supplementary Table 2.

its genome, and that the top two bacterial Tes homologs with a sequence identity of 42% (e -value = $1e^{-149}$) and 40% (e -value = $1e^{-139}$) are from Acidobacterial species *Candidatus Koribacter versatilis* Ellin345 and *Candidatus Solibacter usitatus* Ellin6076, respectively, suggesting that these proteins may be involved in brGDGTs biosynthesis.

Discussion

Archaeal GDGT membrane lipids have been studied for decades^{2,3}, yet the full biosynthetic pathway for synthesizing these membrane-spanning tetraethers has remained unclear. Here, we identify a radical SAM protein, Tes, required for the conversion of two diethers to one tetraether, unraveling a long-standing puzzle of GDGT biosynthesis. The discovery of Tes is significant for understanding the biochemical mechanisms behind GDGT formation as well as having implications for our interpretation of GDGT-based paleotemperature proxies.

The mechanistic details of how the unusual C-C bond formation occurs between two diethers to generate a tetraether have been difficult to ascertain. One area of uncertainty has been whether the diether precursor for GDGT biosynthesis is fully saturated or unsaturated¹⁶. Feeding studies with deuterium-labeled lipid precursors have suggested that the presence of a

Δ^{14-15} double bond in the proposed digeranylgeranyl glycerol phosphate (DGGGP) precursor was crucial for the formation of GDGTs⁴⁴ while other studies have suggested that the diether precursor is fully saturated⁴⁵. While the results we present here cannot distinguish whether the diether precursor is fully saturated or not, we can glean some insight into the potential mechanism of GDGT formation through the protein domain features of Tes. This protein contains two CxxxCxxC motifs¹⁷, the characteristic signature of all radical SAM proteins, suggesting Tes could bind two iron-sulfur clusters and may harbor two distinct active sites with separate catalytic functions. One site could potentially catalyze the desaturation of a fully saturated diether precursor⁴⁶ or mediate double bond migration to generate a terminal olefin⁴⁷. The second site could then generate a C-C bond between sp^2 and sp^3 carbon center^{48,49}. The coordination of two active sites could trigger the head-to-head condensation of two diether lipids to form one GDGT. The discovery of Tes now allows for biochemical and structural characterization that could elucidate this potentially novel radical mechanism for C-C formation.

In addition to the question of whether the Tes substrate is saturated or unsaturated, it is also unclear if the configuration of the glycerol head groups is relevant to the Tes mechanism. There are two glycerol configurations of GDGTs in nature, the parallel

and antiparallel GDGTs^{50,51}. Gräther and Arigoni⁵⁰ showed that some archaeal strains, such as *Methanothermobacter marburgensis* (DSM 2133), *Thermoplasma acidophilum*, and *Saccharolobus solfataricus* P1, produce approximate 1:1 ratio of these two configurations. We found that each of these genomes contains only one Tes homolog. Therefore, it is very likely that one Tes protein is able to generate both parallel and antiparallel GDGTs, and at least some Tes homologs have no preference for the orientation of the glycerol headgroups during head-to-head condensation.

Further studies of Tes could also provide insight into the interplay between the various GDGT biosynthesis proteins discovered thus far. Tes is the third radical SAM protein, in addition to calditol synthase (Cds)¹⁸ and the GDGT ring synthase (Grs)¹⁹, found to be involved in GDGT biosynthesis. Archaea seem to employ radical mechanisms readily in the modification of their membrane lipids and we hypothesize that other GDGT derivatives, such as the H-shaped GDGTs or GDGTs with an unusual cyclohexane ring³, could also be generated through radical mechanisms catalyzed by radical SAM proteins. In addition, it has been proposed that GDGT formation and modification may have a sequential component that allows the cell to balance its membrane structural needs with the need for redox balance under chronic energy stress conditions often brought on by the extreme environments many archaea inhabit^{2,48}. Thus, a biosynthetic scheme where the cell balances condensation by Tes, Grs cyclization, and saturation by the geranylgeranyl reductase (GGR) could serve a role in the cell's effort to maintain physiological homeostasis⁴⁸. Understanding how these proteins are coordinated and how the entire mechanism of GDGT formation and modification is regulated could provide significant insight into the physiological and environmental factors that influence GDGT formation. The discovery of Tes now allows for direct testing of these hypotheses.

The discovery of Tes also aids in our efforts to better constrain GDGT-based paleotemperature proxies. The TEX₈₆ proxy was the first proxy to be based on archaeal lipids and was established as a ratio of the relative abundance of specific cyclized GDGTs in surface sediments that correlated with annual mean sea surface temperature (SST)¹⁵. While the connection between GDGT cyclization and temperature continues to be debated, elucidating the GDGT biosynthesis pathway has allowed for a better grasp of the sources of GDGTs in open oceans. We have previously demonstrated that the lack of GDGT ring synthase (Grs) homologs in MG-II genomic datasets suggests that these archaea, which have not been cultured, did not produce any cyclized GDGTs¹⁹. Here, we demonstrate that the MG-II Euryarchaeota also do not contain a Tes homolog suggesting that they do not produce any GDGTs at all. These bioinformatic analyses further confirm that the dominant source of GDGTs in the open ocean are the MG-I Thaumarchaeota, although culturing and lipid analyses of MG-II Euryarchaeota are necessary to unequivocally confirm that these archaea do not produce any GDGTs.

Finally, the identification of Tes homologs in many bacterial genomes provides potential protein candidates for the formation of membrane-spanning brGDGT lipids in bacteria (Fig. 4 and Supplementary Fig. 4). As mentioned above, brGDGTs were first discovered in a variety of environments and were proposed to be of bacterial origin due to the chemical nature of their alkyl chains and to their linkage to glycerol-3-phosphate rather than glycerol-1-phosphate, as is observed in the archaeal GDGTs³⁷. Additionally, geochemical studies have found a statistical relationship between brGDGTs distribution in sediments and changes in pH and temperature which has resulted in the development of brGDGT-based paleoproxies that help reconstruct past environmental pH and temperatures⁵². However, it has been difficult to

properly constrain brGDGT paleoproxies because bacteria that produce significant amounts of these lipids, and the proteins required for their biosynthesis, have not been well identified⁴². It has been proposed that the synthesis of brGDGTs occurs through the head-to-head condensation of two iso-C15:0 branched phospholipids³⁷. Branched chained fatty acids (BCFAs) such as iso-C15:0 are produced by a variety of bacteria and are characterized by a single methyl group at the penultimate (iso) or antepenultimate (anteiso) carbon⁵³. Given that the tails of the bacterial iso-branched fatty acids and the archaeal isoprenoid lipids are structurally identical, it is possible that these bacterial Tes homologs are involved in the formation of brGDGTs. Follow-up studies demonstrating the biochemical activity of these bacterial Tes homologs can be carried out to verify if they are true brGDGTs biosynthetic enzymes and, if confirmed, can provide insight into the potential sources of these lipids in a variety of ecosystems and allow for a more robust interpretation of brGDGT paleoproxies.

Methods

Microbial strains, media, and growth conditions. Strains used in this study are listed in Supplementary Table 1. *M. maripaludis* S001⁵⁴ and its derivatives were grown in McF medium under a gas phase of N₂/CO₂ (80:20) at 37 °C⁵⁵, and the solid medium was added 1.5% agar. Puromycin was added, when needed, at a final concentration of 2.5 µg/ml. Lysogeny broth (LB) was used to culture *Escherichia coli* DH10B at 37 °C with shaking at 200 rpm and supplemented with 100 µg/ml ampicillin when needed. For growth on solid medium, LB was solidified with 1.5% agar.

S. acidocaldarius MW001 were grown at pH 3.5 and 75 °C in Brock medium supplemented with 0.1% NZ-Amine, 0.2% sucrose, and 10 µg/ml uracil⁵⁶. For growth on solid medium, above medium was solidified with 0.6% gelrite.

Plasmid construction. Plasmids and primers used in this study are listed in Supplementary Tables 2–3. Primers were synthesized from GENEWIZ company (Suzhou, China). The DNeasy Blood and Tissue Kit (Qiagen) was used to isolate genomic DNA. According to the manufacturer's protocol, PCR was performed using Taq DNA polymerase or Phusion high-fidelity DNA Polymerase (New England Biolabs). The FastPure Plasmid Mini Kit (Vazyme) was used for the isolation of *E. coli* plasmid DNA. DNA fragments were purified by the FastPure Gel DNA Extraction Mini kit (Vazyme) during cloning procedures. The ClonExpress One Step Cloning Kit (Vazyme) was used for the construction of the plasmids via stepwise Gibson assembly. DNA sequences were confirmed by sequencing at GENEWIZ (Suzhou, China).

To construct the *saci_0703* deletion plasmids, 600 bp of the *saci_0703* upstream and downstream regions were PCR amplified from *Sulfolobus acidocaldarius* MW001 genomic DNA with primers Z034F/R and Z035F/R, respectively, and inserted into the *NcoI* and *BamHI* restriction sites of pSVA407⁵⁶ via the stepwise Gibson assembly to yield plasmid pSVA407-*saci_0703*UD.

To construct expression plasmids in *Methanococcus maripaludis* S001, *saci_0703*, *ma_1486* and *ma_1114* were amplified from *S. acidocaldarius* MW001 and *Methanosarcina acetivorans* C2A genomic DNA by PCR respectively, and inserted into the *AfeI* and *NotI* sites of expression plasmid pMEV4⁵⁷ via the stepwise Gibson assembly, to yield expression plasmids pMEV4-*saci_0703*, pMEV4-*ma_1486* and pMEV4-*ma_1114*. The *maeo_0574* DNA fragment was synthesized by GENEWIZ company and cloned into the plasmid pMEV4 by the same procedure.

Transformation of plasmids into host cells. Electrocompetent *E. coli* DH10B cells were transformed by electroporation using a Micro-Pulsar Electroporator (BioRad) with the program Ec1 (1.8 kv, 1 pulse for a 0.1-cm cuvette) and the transformants were selected on 100 µg/ml ampicillin. The constructed expression plasmids were transformed into *M. maripaludis* S001 using the polyethylene glycol (PEG)-mediated transformation approach⁵⁸ and selected by plating on puromycin (2.5 µg/ml) McF agar medium, and confirmed via PCR amplification using primers P1F/R. Transformation of deletion plasmid pSVA407-*saci_0703*UD into *S. acidocaldarius* MW001 was performed by electroporation using a Micro-Pulsar Electroporator (BioRad) with the program Ec1 (1.8 kv, 1 pulse for a 0.1 cm cuvette) and the transformants were selected on Brock medium plate supplemented with 0.1% NZ-Amine and 0.2% sucrose 19.

Lipid extraction and analyses. Cultures of *M. maripaludis* S001 transformants (100 mL) were incubated to stationary phase at 37 °C, harvested by centrifugation at 10,000 × g for 10 min, and pellets were stored at –80 °C before extraction. Frozen cell pellets were freeze-dried, and ~30 mg of dried mass was weighed and ground with a laboratory spatula. An internal standard, C46 GTGT (118 ng)⁵⁹, was

added to the ground pellets, before acid hydrolysis in 40 mL capped glass vials with 5 mL 10% hydrochloric acid (HCl) in methanol (MeOH) (vol/vol) at 70 °C overnight. In all, 10 mL of dichloromethane (DCM) and 10 mL of nanopure water were added to the acid-treated mixture, and lipids in the mixture was extracted with DCM for three times in a 50-mL Teflon tube. The samples were centrifuged for 10 min at 2800 × g to separate the aqueous and organic phases. The bottom organic phase layer was transferred to a collection glass vial. The organic phases were combined and filtered through a 0.45-µm PTFE filter and dried under N₂ gas.

Lipid analysis was performed on an Agilent 1200 series high-performance liquid chromatography-atmospheric pressure chemical ionization-6460A triple quadrupole mass spectrometer (HPLC-APCI-MS²). Compound separation was achieved with an Alltech Prevail Cyano column (150 mm × 2.1 mm, 3 µm) at a constant flow rate of 0.2 mL/min throughout. The elution gradient followed Schouten et al.⁶⁰ with some modifications. Lipids were eluted isocratically for the first 5 min with 90% A and 10% B, where A = *n*-hexane and B = *n*-hexane/isopropanol (9:1, v/v), followed by the linear gradient: 90/10 A/B to 82/18 A/B from 5 to 45 min. Finally, 100% B (10 min) was used to wash the column, and then 90/10 A/B to equilibrate the column. Archaeal ether lipids were scanned in a single ion monitoring (SIM) mode, targeting *m/z* 653, 651, 1304, 1302, 1300, 1298, 1296, 1294, 1292. APCI settings were as follows: nebulizer pressure 60 psi, vaporizer temperature 400 °C, drying gas flow 6 L/min and temperature 200 °C, capillary voltage 3500 V, corona current 5 µA (~3200 V). Data were analyzed by Agilent MassHunter software (B.07.00 version).

To obtain the MS² spectra and the fragmentation features of targeted compounds, several samples containing macrocyclic archaeol and GDGT-0 were selected and analyzed with an Agilent 1260 series HPLC system coupled to an Agilent 6530B quadrupole time-of-flight (qTOF) mass spectrometer through an APCI interface. The elution gradient followed Yang et al.⁶¹. Separation of archaeal ether lipids was performed with two silica columns (150 mm × 2.1 mm, 1.9 µm, Thermo Finnigan; USA) maintained at 40 °C. Archaeal ether lipids were eluted isocratically for the first 5 min with 84% A and 16% B, where A = *n*-hexane and B = ethyl acetate (EtOA). Linear gradients, i.e., 84/16 A/B to 82/18 A/B from 5 to 30 min, and 82/18 A/B to 100% B from 30 to 31 min were used. 100% B was held for 4 min and then changed to 16% B in 1 min. Finally, 84/16 A/B was held for 14 min. The APCI settings were as follows: drying gas temperature 200 °C, vaporizer temperature 400 °C, the flow rate of drying gas 6 L/min, nebulizer pressure 60 psi, capillary voltage 3500 V, and corona current 5 µA. To generate the MS² spectra, the collision energy for archaeol, GDGT-0, macrocyclic archaeol, and GTGT-0 were 0, 35, 15, and 35, respectively. The qTOF settings were capillary voltage 3.5 kV, corona current 5 µA, fragmentor voltage 200 V, skimmer voltage 60 V, and octopole voltage 750 V. In the auto MS/MS scanning mode, the mass of [M + H]⁺ ions scanned in MS¹ ranged from *m/z* 100 to 1700. In MS² experiments, the scanned mass of product ions ranged from *m/z* 50 to 1700.

Macrocyclic archaeol was purified by a semi-preparative HPLC on an Agilent series 1290 HPLC equipped with a fraction collector. Compound separation was performed with an Alltech Prevail Cyano column (150 mm × 2.1 mm, 3 µm) and the same elution gradient used on HPLC-QQQ-MS. The fraction containing purified macrocyclic archaeol was transferred into a vial and dried with N₂ gas. Boron tribromide (BBr₃) (0.5 mL, 1 M in DCM; Sigma-Aldrich, USA) was added and the vial was sealed. The vial was heated at 60 °C for 2 h, and residual solvents were removed in a gentle stream of N₂ gas. The bromides were then reduced to alkanes by superhydride (0.5 mL, lithium triethylborohydride (LiEt₃BH), in DCM, Sigma-Aldrich, USA) in a sealed vial with an Argon atmosphere at 60 °C for 2 h. Water was added to quench the reaction. The resulting alkanes were extracted for six times with *n*-hexane. The alkanes were analyzed by a Thermo Finnigan Trace 1300 gas chromatography coupled to an ISQ 7000 mass spectrometer (GC-MS) equipped with a DB-5MS silica capillary column (60 m × 0.25 mm × 0.25 µm) with helium as carrier gas. The initial oven temperature is 70 °C, followed by an increase to 210 °C at 10 °C/min and an increase to 310 °C at 3 °C/min. The final temperature was held at 310 °C for 26 min. The mass spectrometer conditions were as follows: EI energy 70 eV and the auxiliary temperature of GC and MS was 310 °C. Biphytane was identified based on the mass spectra and previous studies^{62,63}.

Bioinformatics analyses. Tetraether synthase (Tes) homologs were identified using BLASTP searches (*e*-value < 1e⁻⁵⁰, identity >20%) of the archaeal database on JGI IMG (<https://img.jgi.doe.gov>) with the Tes (MA_1486) sequence of *Methanosarcina acetivorans* C2A as the search query. A total of 2073 archaeal genomes were searched including 889 isolate genomes, 792 metagenome acquired genomes (MAGs), and 392 single-cell acquired genomes (SAGs). We also searched Tes homologs with 13,919 finished and draft bacterial genomes in the JGI IMG database using the same query and parameters above. Total 1104 archaeal and 456 bacterial Tes homologs were retrieved from multiple phyla, respectively.

Sequences greater than 400 amino acids (aa) were selected and clustered at 90% similarity through CD-HIT version 4.7⁶⁴, resulting in unique 426 archaeal and 185 bacterial Tes homologs from the initial genome searches, respectively. The outgroup for archaeal Tes phylogenetic tree consisted of 41 archaeal GDGT ring synthase (Grs) homologs and the outgroup for that of bacteria consisted of 5 archaeal Tes homologs. Protein sequences were aligned via MUSCLE version 3.8.31⁶⁵ and automatically trimmed via trimAl version 1.4⁶⁶ with the outgroup included. Maximum likelihood trees were built by RAxML version 8.2.12⁶⁷ using

the CAT model of rate heterogeneity with the LG substitution matrix and 1000 bootstrap iterations. Interactive tree of life (iTOL) was used for tree visualization and annotation (<https://itol.embl.de/>)⁶⁸.

BLASTP searches of Tes homologs were also carried out in the Non-redundant protein sequences (nr) database on NCBI (<https://www.ncbi.nlm.nih.gov/>) with the same query described above. Parameters were set as follows: word size = 6, BLOSUM 62, Expect threshold = 10. Low complexity regions were filtered and lower case letters were masked. The specific MAGs datasets of Marine Group II (taxid: 133814), Marine Group III (taxid:347538), Acidobacteriota (taxid:57723), Proteobacteria (taxid:1224), Terrabacteria group (taxid:1783272), FCB group (Fibrobacteres—Chlorobi—Bacteroidetes, taxid:1783270), PVC group (Planctomycetes—Verrucomicrobia—Chlamydiae, taxid:1783257), CPR group (Candidate Phyla Radiation, taxid:1618338—1618340; Patescibacteria group, taxid:1783273) were assigned respectively during the BLASTP searches to identify the Tes homologs (*e*-value < 1e⁻⁵⁰, identity >20%, length >400 aa) in different archaeal and bacterial communities. The results were summarized in Fig. 4 and Supplementary Table 2.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data supporting the findings from this study are available in this article and its Supplementary Information file. Source data are provided with this paper. Gene and protein sequences were obtained from gene database on JGI IMG (<https://img.jgi.doe.gov>) and on NCBI (<https://www.ncbi.nlm.nih.gov/>). Source data are provided with this paper.

Received: 8 December 2021; Accepted: 7 March 2022;

Published online: 22 March 2022

References

- Koga, Y. & Morii, H. Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol. Mol. Biol. Rev.* **71**, 97–120 (2007).
- Valentine, D. L. Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* **5**, 316–323 (2007).
- Schouten, S., Hopmans, E. C. & Damsté, J. S. S. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review. *Org. Geochem.* **54**, 19–61 (2013).
- Yosuke, K. Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* **2012**, 789652 (2012).
- Lai, D., Springstead, J. R. & Monbouquette, H. G. Effect of growth temperature on ether lipid biochemistry in *Archaeoglobus fulgidus*. *Extremophiles* **12**, 271–278 (2008).
- Sprott, G. D., Meloche, M. & Richards, J. C. Proportions of diether, macrocyclic diether, and tetraether lipids in *Methanococcus jannaschii* grown at different temperatures. *J. Bacteriol.* **173**, 3907–3910 (1991).
- Patel, G. B., Agnew, B. J., Deschatelets, L., Fleming, L. P. & Sprott, G. D. In vitro assessment of archaeosome stability for developing oral delivery systems. *Int. J. Pharm.* **194**, 39–49 (2000).
- Mathai, J. C., Sprott, G. D. & Zeidel, M. L. Molecular mechanisms of water and solute transport across archaeobacterial lipid membranes. *J. Biol. Chem.* **276**, 27266–27271 (2001).
- Elling, F. J. et al. Chemotaxonomic characterisation of the thaumarchaeal lipidome. *Environ. Microbiol.* **19**, 2618–2700 (2017).
- Jayme, F. B. et al. Influence of growth phase, pH, and temperature on the abundance and composition of tetraether lipids in the thermoacidophile *Picrophilus torridus*. *Front. Microbiol.* **7**, 1–5 (2016).
- Sara, J. et al. The effects of temperature and growth phase on the lipidomes of *Sulfolobus islandicus* and *Sulfolobus tokodaii*. *Life* **5**, 1539–1566 (2015).
- Elling, F. J., Konneke, M., Musmann, M., Greve, A. & Hinrichs, K. U. Influence of temperature, pH, and salinity on membrane lipid composition and TEX₈₆ of marine planktonic thaumarchaeal isolates. *Geochim. Cosmochim. Acta* **171**, 238–255 (2015).
- Zhou, A., Weber, Y., Chiu, B. K., Elling, F. J. & Leavitt, W. D. Energy flux controls tetraether lipid cyclization in *Sulfolobus acidocaldarius*. *Environ. Microbiol.* **22**, 343–353 (2020).
- Pearson, A. & Ingalls, A. E. Assessing the use of archaeal lipids as marine environmental proxies. *Annu. Rev. Earth Planet. Sci.* **41**, 359–384 (2013).
- Schouten, S., Hopmans, E. C., Schefe, E. & Sinninghe-Damste, J. S. Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures? *Earth Planet. Sci. Lett.* **204**, 265–274 (2002).
- Jain & Samta. Biosynthesis of archaeal membrane ether lipids. *Front. Microbiol.* **5**, 641 (2014).

17. Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-adenosylmethionine enzymes. *Chem. Rev.* **114**, 4229–4317 (2014).
18. Zeng, Z., Liu, X. L., Wei, J. H., Summons, R. E. & Welander, P. V. Caldito-linked membrane lipids are required for acid tolerance in *Sulfolobus acidocaldarius*. *Proc. Natl Acad. Sci. USA* **115**, 12932–12937 (2018).
19. Zeng, Z., Liu, X. L., Farley, K. R., Wei, J. H. & Welander, P. V. GDGT cyclization proteins identify the dominant archaeal sources of tetraether lipids in the ocean. *Proc. Natl Acad. Sci. USA* **116**, 22505–22511 (2019).
20. Coffinet, S. et al. Evidence for enzymatic backbone methylation of the main membrane lipids in the archaeon *Methanomassiliicoccus luminyensis*. *Appl. Environ. Microbiol.* **88**, e02154–21 (2021).
21. Fitz, W. & Arigoni, D. Biosynthesis of 15,16-dimethyltriacontanedioic acid (diabolic acid) from [16-²H₃]- and [14-²H₂]-palmitic acids. *J. Chem. Soc. Chem. Commun.* **20**, 1533–1534 (1992).
22. Isobe, K. et al. Geranylgeranyl reductase and ferredoxin from *Methanosarcina acetivorans* are required for the synthesis of fully reduced archaeal membrane lipid in *Escherichia coli* Cells. *J. Bacteriol.* **196**, 417–423 (2014).
23. Xu, Q. et al. Insights into substrate specificity of geranylgeranyl reductases revealed by the structure of digeranylgeranyl glycerophospholipid reductase, an essential enzyme in the biosynthesis of archaeal membrane lipids. *J. Mol. Biol.* **404**, 403–417 (2010).
24. Quehenberger, J., Pittenauer, E., Allmaier, G. & Spadiut, O. The influence of the specific growth rate on the lipid composition of *Sulfolobus acidocaldarius*. *Extremophiles* **24**, 413–420 (2020).
25. Koga, Y., Morii, H., Akagawa-Matsushita, M. & Ohga, M. Correlation of polar lipid composition with 16S rRNA phylogeny in Methanogens. Further analysis of lipid component parts. *Biosci. Biotechnol. Biochem.* **62**, 230–236 (1998).
26. Comita, P. B. & Gagosian, R. B. Membrane lipid from deep-sea hydrothermal vent methanogen: a new macrocyclic glycerol diether. *Science* **222**, 1329–1331 (1983).
27. Baumann, L. M. F. et al. Intact polar lipid and core lipid inventory of the hydrothermal vent methanogens *Methanocaldococcus villosus* and *Methanothermococcus okinawensis*. *Org. Geochem.* **126**, 33–42 (2018).
28. Comita, P. B., Gagosian, R. B., Pang, H. & Costello, C. E. Structural elucidation of a unique macrocyclic membrane lipid from a new, extremely thermophilic, deep-sea hydrothermal vent archaeobacterium, *Methanococcus jannaschii*. *J. Biol. Chem.* **259**, 15234–15241 (1984).
29. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. N. P. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
30. Moller, S., Croning, M. & Apweiler, R. J. B. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**, 646–653 (2001).
31. Baker, B. J., Anda, V. D., Seitz, K. W., Dombrowski, N. & Lloyd, K. G. Diversity, ecology and evolution of Archaea. *Nat. Microbiol.* **5**, 887–900 (2020).
32. Lincoln, S. A., Wai, B., Eppley, J. M., Church, M. J. & DeLong, E. F. Planktonic Euryarchaeota are a significant source of archaeal tetraether lipids in the ocean. *Proc. Natl Acad. Sci. USA* **111**, 9858–9863 (2014).
33. Haro-Moreno, J. M., Rodriguez-Valera, F., López-García, P., Moreira, D. & Martín-Cuadrado, A. B. New insights into marine group III Euryarchaeota, from dark to light. *ISME J.* **11**, 1102–1117 (2017).
34. Cui, H. L., Yang, X., Gao, X. & Xu, X. W. *Halobellus clavatus* gen. nov., sp. nov. and *Halorientalis regularis* gen. nov., sp. nov., two new members of the family Halobacteriaceae. *Int. J. Syst. Evol. Microbiol.* **61**, 2682–2689 (2011).
35. Castelle, C. J. et al. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat. Rev. Microbiol.* **16**, 629–645 (2018).
36. Jahn, U., Summons, R., Sturt, H., Grosjean, E. & Huber, H. Composition of the lipids of *Nanoarchaeum equitans* and their origin from its host *Ignicoccus* sp. strain KIN4/I. *Arch. Microbiol.* **182**, 404–413 (2004).
37. Weijers, J., Schouten, S., Hopmans, E. C., Geenevasen, J. & Damsté, J. S. S. Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits. *Environ. Microbiol.* **8**, 648–657 (2006).
38. Damste, J. S. S., Hopmans, E. C., Pancost, R. D., Schouten, S. & Geenevasen, J. A. J. Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments. *Chem. Commun.* **17**, 1683–1684 (2000).
39. Peterse, F., Kim, J. H., Schouten, S., Kristensen, D. K. & Damste, J. Constraints on the application of the MBT/CBT palaeothermometer at high latitude environments (Svalbard, Norway). *Org. Geochem.* **40**, 692–699 (2009).
40. Tierney, J. E. & Russell, J. M. Distributions of branched GDGTs in a tropical lake system: Implications for lacustrine application of the MBT/CBT paleoproxy. *Org. Geochem.* **40**, 1032–1036 (2009).
41. Schouten, S., Hopmans, E. C., Pancost, R. D. & Damste, J. S. Widespread occurrence of structurally diverse tetraether membrane lipids: Evidence for the ubiquitous presence of low-temperature relatives of hyperthermophiles. *Proc. Natl Acad. Sci. USA* **97**, 14421–14426 (2000).
42. Sinninghe, D. et al. An overview of the occurrence of ether- and ester- linked iso-diabolic acid membrane lipids in microbial cultures of the Acidobacteria: Implications for brGDGT paleoproxies for temperature and pH. *Org. Geochem.* **124**, 63–76 (2018).
43. Halamka, T. A. et al. Oxygen limitation can trigger the production of branched GDGTs in culture. *Geochem. Perspect. Lett.* **19**, 36–39 (2021).
44. Eguchi, T., Nishimura, Y. & Kakinuma, K. Importance of the isopropylidene terminal of geranylgeranyl group for the formation of tetraether lipid in methanogenic archaea. *Tetrahedron Lett.* **44**, 3275–3279 (2003).
45. 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).
46. Ji, X., Li, Y., Jia, Y., Wei, D. & Qi, Z. Mechanistic insights into the radical S-adenosyl-1-methionine enzyme NosL from a substrate analogue and the shunt products. *Angew. Chem. Int. Ed.* **128**, 3395–3398 (2016).
47. Eguchi, T., Takyo, H., Morita, M., Kakinuma, K. & Koga, Y. Unusual double-bond migration as a plausible key reaction in the biosynthesis of the isoprenoidal membrane lipids of methanogenic archaea. *Chem. Commun.* 1545–1546 (2000).
48. Pearson, A. Resolving a piece of the archaeal lipid puzzle. *Proc. Natl Acad. Sci. USA* **116**, 22423–22425 (2019).
49. Yokoyama, K. & Lilla, E. A. C–C bond forming radical SAM enzymes involved in the construction of carbon skeletons of cofactors and natural products. *Nat. Prod. Rep.* **35**, 660–694 (2018).
50. Gräther, O. & Arigoni, D. Detection of regioisomeric macrocyclic tetraethers in the lipids of *Methanobacterium thermoautotrophicum* and other archaeal organisms. *J. Chem. Soc.* **4**, 405–406 (1995).
51. Liu, X. L., Dar, B., Cb, B. & Res, C. Glycerol configurations of environmental GDGTs investigated using a selective sn 2 ether cleavage protocol. *Org. Geochem.* **128**, 57–62 (2019).
52. Weijers, J., Schouten, S., Donker, J., Hopmans, E. C. & Damsté, J. S. S. Environmental controls on bacterial tetraether membrane lipid distribution in soils. *Geochim. Cosmochim. Acta* **71**, 703–713 (2007).
53. Diomande, S. E., Christophe, N. T., Guinebretiere, M. H., Broussolle, V. & Brillard, J. Role of fatty acids in *Bacillus* environmental adaptation. *Front. Microbiol.* **6**, 813 (2015).
54. Walters, A. D., Smith, S. E. & Chong, J. Shuttle vector system for *Methanococcus maripaludis* with improved transformation efficiency. *Appl. Environ. Microbiol.* **77**, 2549–2551 (2011).
55. Sarmiento, F., Leigh, J. A. & Whitman, W. B. Genetic systems for hydrogenotrophic methanogens. *Methods Enzymol.* **494**, 43–73 (2011).
56. Michaela, W. et al. Versatile genetic tool box for the Crenarchaeote *Sulfolobus acidocaldarius*. *Front. Microbiol.* **3**, 1–12 (2012).
57. Zhe, L., Jain, R., Smith, P., Fetchko, T. & Whitman, W. B. Engineering the autotroph *Methanococcus maripaludis* for geraniol production. *ACS Synth. Biol.* **5**, 577–581 (2016).
58. 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).
59. Huguet, C., Hopmans, E. C., Febo-Ayala, W., Thompson, D. H. & Schouten, S. An improved method to determine the absolute abundance of glycerol dibiphytanyl glycerol tetraether lipids. *Org. Geochem.* **37**, 1036–1041 (2006).
60. Schouten, S., Huguet, C., Hopmans, E. C., Kienhuis, M. & Damsté, J. S. S. Analytical methodology for TEX₈₆ paleothermometry by high-performance liquid chromatography/atmospheric pressure chemical ionization-mass spectrometry. *Anal. Chem.* **79**, 2940–2944 (2007).
61. Yang et al. The 6-methyl branched tetraethers significantly affect the performance of the methylation index (MBT) in soils from an altitudinal transect at Mount Shennongjia. *Org. Geochem.* **82**, 42–53 (2015).
62. Delong, E. F. et al. Dibiphytanyl ether lipids in nonthermophilic Crenarchaeotes. *Appl. Environ. Microbiol.* **64**, 1133–1138 (1998).
63. Liu, X. L., Summons, R. E. & Hinrichs, K. U. Extending the known range of glycerol ether lipids in the environment: structural assignments based on tandem mass spectral fragmentation patterns. *Rapid Commun. Mass Spectrom.* **26**, 2295–2302 (2012).
64. Godzik, L. A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
65. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
66. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
67. Alexandros, S. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
68. Ivica, L. & Peer, B. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
69. Imachi, H., Nobu, M. K., Nakahara, N., Morono, Y. & Takai, K. Isolation of an archaeon at the prokaryote–eukaryote interface. *Nature* **577**, 519–525 (2020).

70. Bale, N. J. et al. New insights into the polar lipid composition of extremely halo(alkali)philic Euryarchaea from hypersaline lakes. *Front. Microbiol.* **10**, 377 (2019).
71. Meador, T. B. et al. Thermococcus kodakarensis modulates its polar membrane lipids and elemental composition according to growth stage and phosphate availability. *Front. Microbiol.* **5**, 10 (2014).
72. Hug, L. A. et al. A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016).
73. Rinke, C. et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437 (2013).

Acknowledgements

We thank Prof. Xiuzhu Dong, Dr. Lei Yue, and Prof. William B. Whitman for the gift of *M. maripaludis* S001 and pMEV4 plasmid expression system. We thank Dr. Yongxin Chen and Huan Fang in Oil and Gas research center, Northwest Institute of Eco-environment and resources, Chinese Academy of Sciences for providing access to HPLC-Q-TOF-MS/MS and helping with mass spectra analysis. We also thank Prof. Kenichi Yokoyama for helpful discussions on potential radical SAM enzyme mechanisms. Support for this study was provided by the National Natural Science Foundation of China to Z.Z. (No. 32170041 and 92051112), by the Simons Collaboration on the Origins of Life to P.V.W. (No. 511568), and by the National Science Foundation to P.V.W. (No. 1752564).

Author contributions

Z.Z. and P.V.W. conceived the project; H.C., Z.Z., and X.F. performed archaea culture and genetic experiments; H.C., H.Y., and H.P. conducted lipid analysis; Y.C. and Z.Z. conducted phylogenetic analysis; Z.Z., H.C., H.Y., Y.C., W.Y., and P.V.W. interpreted results; Z.Z. and P.V.W. wrote the manuscript with contributions from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-29264-x>.

Correspondence and requests for materials should be addressed to Zhirui Zeng or Paula V. Welander.

Peer review information *Nature Communications* thanks Hisashi Hemmi, Kai-Uwe Hinrichs and the other, anonymous, reviewer for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022