A distinct pathway for tetrahymanol synthesis in bacteria

Amy B. Banta[1], Jeremy H. Wei[1], and Paula V. Welander[2]

Department of Earth System Science, Stanford University, Stanford, CA 94305

Edited by John M. Hayes, Woods Hole Oceanographic Institution, Berkeley, CA, and approved September 25, 2015 (received for review June 11, 2015)

Tetrahymanol is a polycyclic triterpenoid lipid first discovered in the ciliate *Tetrahymena pyriformis* whose potential diagenetic product, gammacerane, is often used as a biomarker for water column stratification in ancient ecosystems. Bacteria are also a potential source of tetrahymanol, but neither the distribution of this lipid in extant bacteria nor the significance of bacterial tetrahymanol synthesis for interpreting gammacerane biosignatures is known. Here we couple comparative genomics with genetic and lipid analyses to link a protein of unknown function to tetrahymanol synthesis in bacteria. This tetrahymanol synthase (Ths) is found in a variety of bacterial genomes, including aerobic methanotrophs, nitrite-oxidizers, and sulfate-reducers, and in a subset of aquatic and terrestrial metagenomes. Thus, the potential to produce tetrahymanol is more widespread in the bacterial domain than previously thought. However, Ths is not encoded in any eukaryotic genomes, nor is it homologous to eukaryotic squalene-tetrahymanol cyclase, which catalyzes the cyclization of squalene directly to tetrahymanol. Rather, heterologous expression studies suggest that bacteria couple the cyclization of squalene to a hopene molecule by squalene-hopene cyclase with a subsequent Ths-dependent ring expansion to form tetrahymanol. Thus, bacteria and eukaryotes have evolved distinct biochemical mechanisms for producing tetrahymanol.

gammacerane | tetrahymanol | biomarkers | methanotrophs | sterols

Sterols are cyclic triterpenoid lipids that are ubiquitous and essential in eukaryotes and which play a key role in various cellular functions including maintaining membrane integrity, stress tolerance, and phagocytosis (1–3). However, not all eukaryotes are capable of producing sterols de novo (4). *Tetrahymena pyriformis*, a ciliated protozoan commonly found in aquatic environments, is a sterol auxotroph that must obtain these lipids through its diet (5). However, *T. pyriformis* is also capable of thriving in anoxic ecosystems where sterol biosynthesis is restricted due to a lack of oxygen. Under these sterol-starvation conditions, *T. pyriformis* produces the polycyclic triterpenoid tetrahymanol which is thought to function as a sterol surrogate (Fig. 1*A*) (6, 7).

Tetrahymanol was first discovered in *T. pyriformis* (8) and has subsequently been detected in other eukaryotes including numerous marine and freshwater ciliates, an anaerobic free-living protist, an anaerobic rumen fungus, and a fern plant (9–11). Tetrahymanol has also been detected directly in freshwater and marine sediments (12, 13), and it is recognized as a biological precursor of gammacerane (Fig. 1*A*), a polycyclic hydrocarbon detected in sedimentary rocks dating as far back as the late Proterozoic (~850 Mya) (14–16). Based on the common occurrence of tetrahymanol in stratified aquatic environments today, the gammacerane index—a relative measurement of the prevalence of gammacerane in a sample—is used as an indicator for water column stratification (possibly linked to hypersalinity) during source rock deposition (12, 13, 17, 18).

Tetrahymanol is also a minor component of the lipidome of two α-Proteobacteria belonging to the Bradyrhizobiaceae family, *Rhodopseudomonas palustris* and *Bradyrhizobium japonicum* (19–21), and it is uncertain how significant this bacterial production of tetrahymanol is for interpreting gammacerane biosignatures. Although a direct correlation between sterol starvation and tetrahymanol production has been established in ciliates (6), the physiological role of tetrahymanol in bacteria is unknown. Recent studies have highlighted increased tetrahymanol production in *R. palustris* TIE-1 under certain physiological conditions (e.g., photoautotrophic growth) and also when cellular hopanoid lipid profiles are altered in gene deletion mutants (22, 23), but the physiological significance of these changes is not known. Further, the biochemical mechanism of tetrahymanol synthesis in bacteria is unclear. In ciliates, squalene-tetrahymanol cyclase (Stc) catalyzes the cyclization of squalene directly to tetrahymanol (24), but neither of the two known bacterial tetrahymanol producers harbor a copy of Stc (10, 24). *R. palustris* and *B. japonicum* do possess an evolutionarily related cyclase, squalene-hopene cyclase (Shc), whose main function is the cyclization of squalene to the hopanoid diploptene (25). It has been proposed that bacterial Shc in tetrahymanol-producing bacteria is a low-fidelity enzyme (26) that could conditionally produce tetrahymanol in addition to hopanoids. In support of this, deletion of *shc* in *R. palustris* and *B. japonicum* eliminates both hopanoid and tetrahymanol production (27, 28). However, Shc alone cannot convert squalene to tetrahymanol in vitro (29, 30), suggesting that another protein could be involved in its biosynthesis.

Consequently, our lack of understanding of tetrahymanol function and biosynthesis in bacteria has led to the assumption that tetrahymanol is not widespread in the bacterial domain. However, our knowledge of the diversity of extant tetrahymanol producers relies on direct observation of a lipid that may be conditionally produced in organisms that are not well studied. Therefore, in this study, we set out to identify bacterial tetrahymanol biosynthesis genes that could function as genetic markers that would allow us to assess the potential for tetrahymanol synthesis in the bacterial domain in a culture-independent manner.

Significance

Sterols and sterol-like lipids have critical physiological roles in both eukaryotes and bacteria. These cyclic lipids are also well preserved in ancient rocks and can function as microbial molecular fossils. Gammacerane is one such lipid that is used as an indicator of water column stratification deep in time. However, a proper interpretation of gammacerane biosignatures requires a full understanding of the biosynthesis and taxonomic distribution of one of its potential precursors, tetrahymanol, in modern microbes. Here we establish that bacterial tetrahymanol producers are more diverse than previously thought and demonstrate that bacteria have evolved a biochemical mechanism distinct from eukaryotes for synthesizing tetrahymanol.

Author contributions: A.B.B., J.H.W., and P.V.W. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]A.B.B. and J.H.W. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: welander@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1511482112/-/DCSupplemental.

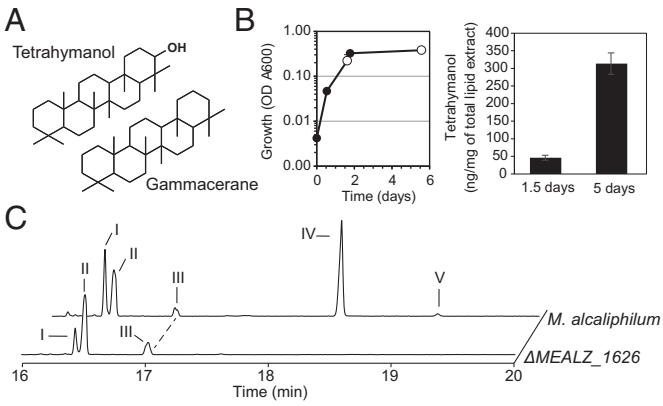

Fig. 1. Tetrahymanol biosynthesis requires a hypothetical protein of unknown function. (*A*) Structure of tetrahymanol and its proposed diagenetic product gammacerane. (*B*) *Left* shows growth of *M. alcaliphilum* over 5 d and represents one growth curve done in triplicate (error bars represent biological triplicate samples and may be smaller than markers). Open circles represent time points when cells were harvested for lipid analysis. *Right* shows eightfold increase in tetrahymanol in stationary phase cells versus late exponential phase cells. (*C*) GC-MS extracted ion chromatograms (*m/z* 191) demonstrating loss of tetrahymanol (IV) and 3-methyltetrahymanol (V) after deletion of locus MEALZ_1626 in *M. alcaliphilum*. Other lipids identified: I, 4-dimethylsterol; II, hopenes; and III, 3-methylhopenes.

Results

Tetrahymanol Production in a γ-Proteobacterium. *Methylomicrobium alcaliphilum* is an obligate aerobic methanotroph isolated from sediments of a hypersaline alkaline lake (31) whose genome encodes homologs of several hopanoid and sterol biosynthesis genes (32). Lipid analysis of *M. alcaliphilum* revealed tetrahymanol and a putative 3-methyltetrahymanol in addition to an array of aminohopanoids, 3-methylaminohopanoids, and 4-methylsterols typically found in γ-Proteobacterial methanotrophs (Figs. S1 and S2) (33, 34). The presence of tetrahymanol was unexpected because *M. alcaliphilum* is neither physiologically nor phylogenetically close to the Bradyrhizobiaceae family, the only bacteria known to synthesize tetrahymanol. *M. alcaliphilum* not only produces significantly more tetrahymanol than what is typically observed in the α-Proteobacterial species (20, 21, 30) but also varies the amount of tetrahymanol under different culture conditions (Fig. 1*B*). Specifically, we observed an eightfold increase in tetrahymanol in stationary phase cells (day 5 of growth) compared with late exponential phase cells (day 1.5 of growth). Given the higher levels of tetrahymanol production in *M. alcaliphilum*, this organism seemed ideal for pursuing studies to uncover potential tetrahymanol biosynthesis proteins.

Identification of a Tetrahymanol Synthase. To identify putative bacterial tetrahymanol biosynthesis proteins, we used comparative genomics to search for protein-encoding genes present in *M. alcaliphilum*, *R. palustris*, and *B. japonicum* but absent in bacteria that we have experimentally verified to produce hopanoids but not tetrahymanol (*Materials and Methods*). The resultant 31 *M. alcaliphilum* genes have no homologs (maximum e-value 1e-5) in tetrahymanol producing ciliates and other eukaryotes. However, one gene, encoding a hypothetical protein with no identifiable motifs (locus tag: MEALZ_1626), occurs in all *Methylomicrobium*, *Bradyrhizobium*, and *Rhodopseudomonas* genomes and exclusively in bacterial genomes also containing squalene-hopene cyclase (maximum e-value 1e-50). Although the genomic context of this protein does not demonstrate a link to lipid biosynthesis, its distribution pattern in bacterial genomes led us to hypothesize that MEALZ_1626 may play a role in tetrahymanol production. Deletion of this locus in *M. alcaliphilum* resulted in a loss of tetrahymanol but not hopanoids (Fig. 1*C*). Further, placing the MEALZ_1626 gene in a permissible site on the chromosome of the deletion strain restored production of tetrahymanol (Fig. S3*A*), indicating that this gene, renamed *ths* for tetrahymanol synthase, encodes a protein that is specifically required to produce tetrahymanol.

Homologs of *M. alcaliphilum* Ths are present in 104 bacterial genomes in three different phyla (Table S1), including several α-Proteobacteria (*Rhodopseudomonas*, *Bradyrhizobium*, *Nitrobacter*, *Afipia*, *Agromonas*, and *Rhodovulum* species), aerobic methanotrophic γ-Proteobacteria (*Methylomonas* and *Methylomicrobium* species), and sulfate-reducing δ-Proteobacteria (*Desulfovibrio* species). The occurrence of *ths* in the sulfate-reducing *Desulfovibrio* was surprising; several strains of these anaerobic bacteria synthesize hopanoids, but there are no reports of tetrahymanol production (35). However, lipid analysis verified that *Desulfovibrio inopinatus* is capable of producing trace amounts of tetrahymanol (Fig. S4).

An unrooted maximum likelihood tree of Ths amino acid sequences from cultured organisms (Fig. 2*A* and Figs. S5*A* and S6) indicates that there are two divergent lineages of Ths: one comprises the α-Proteobacteria, and a second comprises the γ-Proteobacterial methanotrophs and δ-Proteobacterial sulfate-reducing bacteria. To verify that Ths homologs in the α-Proteobacteria are truly involved in tetrahymanol synthesis, we deleted the *R. palustris* TIE-1 *ths* homolog (locus tag: Rpal_0860, 28% identity and 43% similarity to *M. alcaliphilum* Ths). Tetrahymanol production was lost in the *R. palustris* Δ*ths* strain (Fig. 2*B*). Expression of *ths* from a plasmid restored tetrahymanol biosynthesis in the mutant, verifying that this protein is a bona fide tetrahymanol synthase in the α-Proteobacteria phyla as well (Fig. S3*B*).

To better understand the types of ecosystems tetrahymanol-producing bacteria might inhabit, we searched for homologs of Ths in environmental metagenomes (2,606 total JGI environmental metagenomes; e-value cutoff: 1e-05). We identified 472 potential Ths homologs in 35 different metagenomes (Fig. 3 and Table S2) including 25 soil, 8 freshwater, and 2 marine metagenomes. An unrooted maximum likelihood tree constructed with a subset of Ths metagenome and genome sequences (Fig. S5*B*)

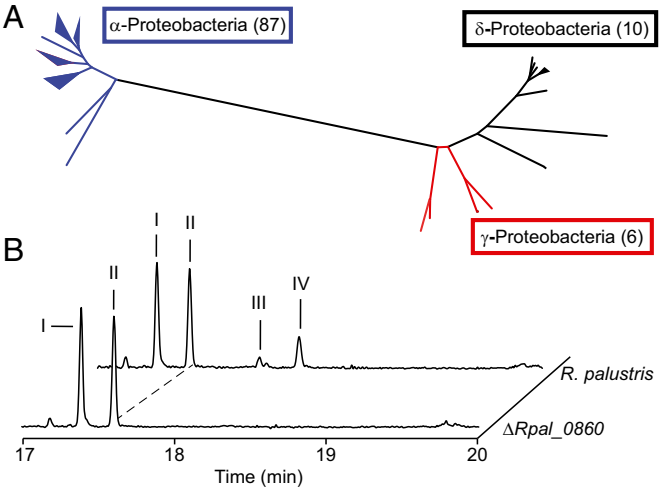

Fig. 2. The MEALZ_1626 homolog from α-Proteobacteria is a bona fide tetrahymanol synthase. (*A*) Unrooted maximum likelihood phylogenetic tree of putative tetrahymanol synthase protein sequences demonstrating two potentially divergent lineages. The numbers following each phylum name indicate the number of sequences in each phylum. The full phylogenetic tree and the sequence alignment are provided in *Supporting Information*. (*B*) GC-MS extracted ion chromatograms (*m/z* 191) demonstrating loss of tetrahymanol (IV) and 2-methyltetrahymanol (III) after deletion of locus Rpal_0860 in *R. palustris*. Other lipids identified: I, 2-methyldiplopterol, and II, diplopterol.

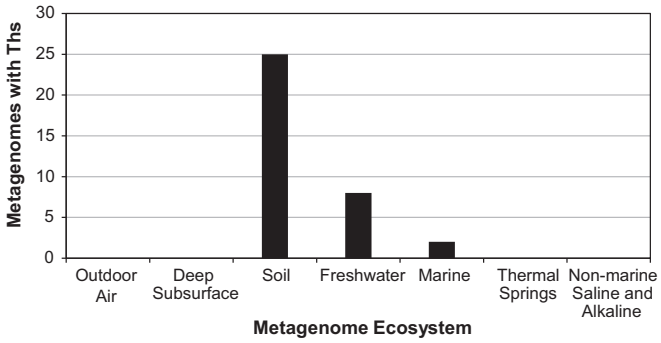

Fig. 3. Tetrahymanol synthase homologs found in metagenomes are restricted to freshwater, soil, and marine environments. All Joint Genome Institute (JGI) environmental metagenome databases were queried for homologs to the *M. alcaliphilum* Ths (e-value cutoff: 1e-05). A total of 472 Ths sequences were identified in 35 different metagenomes. Each bar represents the number of metagenomes with Ths sequences in the indicated ecosystem category.

demonstrates that the majority of soil metagenome Ths homologs clustered within the α-Proteobacterial clade. This clustering suggests that these bacteria are a potential source of tetrahymanol in terrestrial environments. A few freshwater lake sequences also clustered within the α-Proteobacteria, but those associated with methanotrophy (e.g., Lake Washington samples) were more closely related to the γ-Proteobacterial methanotrophs (Fig. S5*B*). Overall, the distribution of Ths in metagenomes seems to reflect the environmental distribution of the Ths-containing cultured bacteria as shown in Table S1.

Tetrahymanol Synthase Mechanism. Our genetic analyses indicate that *ths* is required for tetrahymanol biosynthesis in bacteria. As mentioned above, the squalene-hopene cyclase is also required, but it is unclear if any additional characteristics of Shc or the cellular environment in tetrahymanol-producing bacteria are necessary. To test this, we expressed *M. alcaliphilum ths* in two hopanoid-producing strains that typically do not synthesize tetrahymanol, the γ-Proteobacterium *Methylococcus capsulatus* Bath (36) and the β-Proteobacterium *Burkholderia phytofirmans* PsJN (37, 38). Both *ths*$^+$ strains produced tetrahymanol (Fig. 4), demonstrating that *ths* is sufficient to induce tetrahymanol biosynthesis in hopanoid-producing bacteria and that the occurrence of both *ths* and *shc* in a genome is a strong indicator of tetrahymanol production.

However, it still remained unclear how Shc and Ths together catalyzed the conversion of squalene to tetrahymanol. A mechanism in which Ths demethylates squalene at the C-23 position to generate C(23)-norsqualene before cyclization by Shc (Fig. 5*A*) is plausible based on in vitro studies (39). However, this mechanism would require addition of the C-23 methyl group after cyclization to form tetrahymanol. A more straightforward scenario would require Shc to first cyclize squalene to a hopene, such as diploptene (27, 28), and then Ths would convert this hopene to tetrahymanol via a ring expansion (Fig. 5*A*). To test these two possibilities, we first engineered *Escherichia coli* to produce high levels of squalene by introducing a plasmid that expresses the *M. alcaliphilum* squalene synthase (*sqs*) gene as well as a second plasmid that increases overall isoprenoid synthesis (40, 41). This strain, also expressing *M. alcaliphilum ths* from a third plasmid, did not demethylate or otherwise modify squalene, nor did it produce tetrahymanol, indicating that Ths was not directly using squalene as a substrate (Fig. 5*B* and Fig. S7). We next modified this *E. coli* strain to produce hopenes by introducing a plasmid that expresses both *M. alcaliphilum sqs* and *shc* and found that expression of *M. alcaliphilum ths* in this system resulted in tetrahymanol production (Fig. 5*B*). Taken together, these data strongly suggest that Ths is converting a hopene to tetrahymanol rather than modifying squalene before cyclization.

Discussion

The discovery of tetrahymanol synthase demonstrates the effectiveness of combining comparative genomics and classical genetics with organic geochemistry to link gene identity to novel protein functions. There is a critical need for these types of studies given the rapid rate at which acquisition of genomic data are currently outpacing the assignment of hypothetical protein function. The tetrahymanol synthase case is particularly compelling because this protein has no known motifs or characterized homologs, nor does the genomic context provide any clues to its involvement in lipid synthesis. Thus, without the framework provided by lipid analysis and comparative genomics it would have been difficult to link the *ths* open reading frame with tetrahymanol biosynthesis.

Our findings also indicate that tetrahymanol biosynthesis has experienced convergent evolution with eukaryotes and bacteria developing distinct mechanisms for producing the

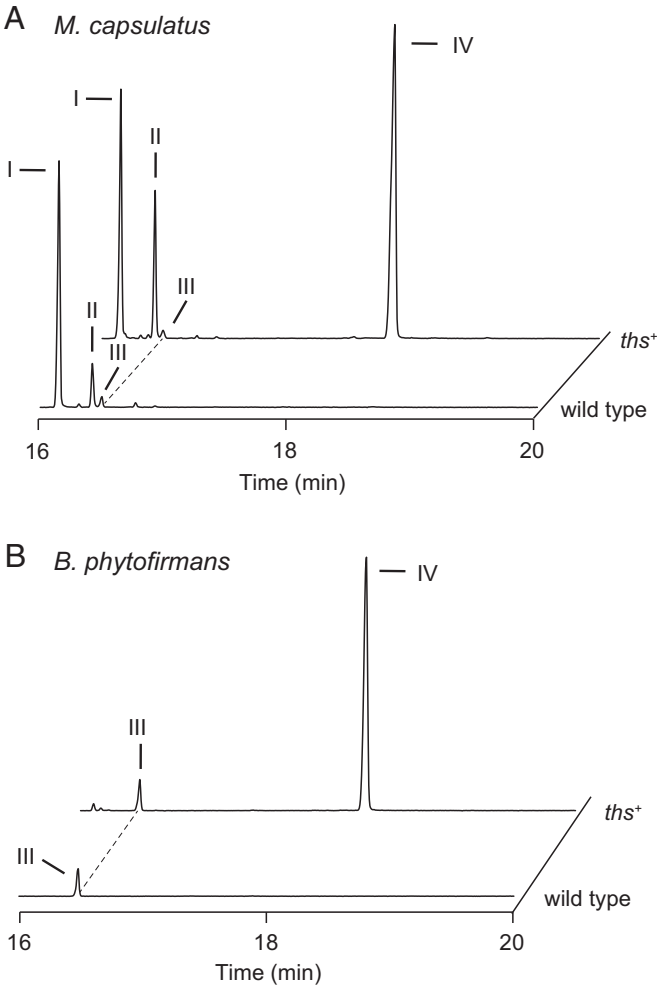

Fig. 4. Heterologous expression of *M. alcaliphilum* tetrahymanol synthase (*ths*) induces tetrahymanol production in *Methylococcus capsulatus* and *Burkholderia phytofirmans*. GC-MS combined extracted ion chromatograms (*m/z* 191, 456, and 442) of lipids extracted from (*A*) *M. capsulatus* or (*B*) *B. phytofirmans* cells expressing tetrahymanol synthase (*ths*) from a plasmid (pSRK-*ths*) compared with the empty vector control (pSRKGm). Lipids identified: I, 4-methylsterol (*m/z* 442); II, 4-dimethylsterol (*m/z* 456); III, hopenes (*m/z* 191); and IV, tetrahymanol (*m/z* 191).

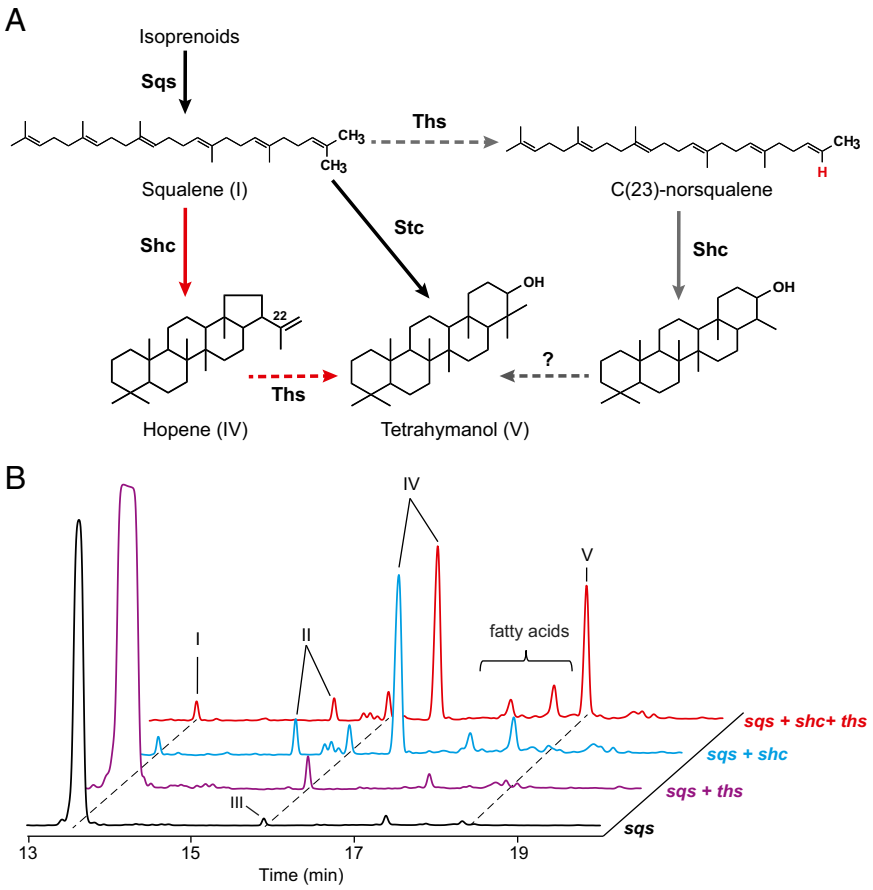

Fig. 5. Tetrahymanol synthase functions on hopenes rather than squalene. (*A*) Two potential bacterial tetrahymanol biosynthetic pathways. In gray, Ths first demethylates squalene to form C-(23)-norsqualene, which is then converted to a demethyl derivative of tetrahymanol by Shc. This pathway would require the addition of a methyl group at C-23 to form tetrahymanol. In red, Shc cyclizes squalene to diploptene, and Ths then expands the E ring to form tetrahymanol. Conversion of isoprenoids to squalene is catalyzed by Sqs, and conversion of squalene to tetrahymanol in eukaryotes is catalyzed by Stc (black arrows). Solid arrows indicate reactions that have been demonstrated in vitro (24, 30, 39, 40), and dashed arrows are proposed steps. (*B*) GC-MS combined extracted ion chromatograms (*m/z* 69, 191, and 215) of *E. coli* strains engineered to produce squalene or hopenes with and without coexpression of *M. alcaliphilum ths*. Production of tetrahymanol (V) is observed when *ths* is expressed in the strain synthesizing hopenes (red). Sqs, squalene synthase; Shc, squalene-hopene cyclase; Ths, tetrahymanol synthase; and Stc, squalene-tetrahymanol cyclase. Lipids identified: I, squalene (*m/z* 69); II, hop-17 (21)-ene (*m/z* 191); III, cholestanol standard (*m/z* 215); IV, hopenes (*m/z* 191); and V, tetrahymanol (*m/z* 191). Full total ion chromatograms are shown in *Supporting Information*.

same molecule. Studies of the biochemical mechanism involved in catalyzing the conversion of squalene to tetrahymanol by Ths and Shc will prove useful in better constraining the evolutionary history of cyclic triterpenoid biosynthesis. Our data suggest that Ths is functioning by expanding the E ring of a hopene after cyclization of squalene by Shc. Although we cannot yet distinguish which hopene is converted to tetrahymanol, a ring expansion of diploptene through the protonation of the C-22 double bond to propagate a carbocation seems most likely as this mechanism is similar to what occurs in the formation of the C ring during cyclization of squalene to hopene by Shc (25). However, because Ths has no identifiable motifs that could be responsible for protonation of the diploptene double bond, further studies of the structure and function of Ths may reveal novel protein folds or biochemical mechanisms.

Our analysis of Ths distribution in genomes and metagenomes demonstrates that the potential for tetrahymanol production is more widespread in the bacterial domain than previously thought. We find that Ths homologs are present in a variety of α-Proteobacteria typically found in soil environments associated with plants. Further, the majority of metagenomic Ths homologs we identify in this study tend to cluster within the α-Proteobacteria clades. Thus, tetrahymanol-producing α-Proteobacteria are potentially a significant source of tetrahymanol in terrestrial systems. The low abundance of Ths homologs in marine metagenomes, despite the presence of *ths* in a few marine bacterial isolates (Table S1), suggests that bacterial tetrahymanol production may not be common in modern marine ecosystems. However, this restricted distribution may be a result of biases in the available genomic and metagenomic databases rather than a true accounting of tetrahymanol-producing bacteria in these ecosystems. As more metagenomes become publicly available, we may see an increase in environmental Ths sequences. Nonetheless, the genomic and metagenomic data we have now suggest that marine ciliates may be the predominant depositional source of tetrahymanol in marine environments (42), whereas bacterial sources may be significant in terrestrial or lacustrine ecosystems.

The production of tetrahymanol in aerobic methanotrophs and sulfate-reducing bacteria is also notable. First, not all aerobic methanotroph or sulfate-reducing bacterial genomes have a Ths homolog, indicating that only a subset of these taxa have the potential to produce tetrahymanol. Physiological studies aimed at understanding why only certain methanotroph or sulfate-reducing bacterial species produce tetrahymanol may demonstrate a correlation between this lipid and specific environmental niches

or physiological conditions. Second, recent studies have documented the occurrence of aerobic methanotrophs in the suboxic zone of stratified marine and freshwater water bodies, whereas sulfate-reducing bacteria are often found in the anoxic sediments of these systems (43–48). Thus, tetrahymanol synthesis in bacteria may be functionally linked to these types of stratified environments. This is significant because the gammacerane index is used primarily as an indicator of stratified conditions in ancient ecosystems (17, 18). The identification of Ths enables physiological studies to directly test a potential link between environmental conditions and tetrahymanol in bacteria. We can now characterize levels of *ths* expression as well as phenotypes of *ths* gene deletion mutants in cultured strains under environmental conditions relevant to water stratification (e.g., high salinity, suboxic, or sulfidic). In addition, *ths* can be used as a genetic marker to probe for the capability of bacterial tetrahymanol biosynthesis in stratified environments and also for environmental transcriptomic studies to assess how changes in these ecosystems affect *ths* expression. If a functional link can be established in bacteria between tetrahymanol and a response to specific environmental parameters relevant to water stratification, then gammacerane would continue to function as a robust indicator of water stratification conditions regardless of whether the depositional source was bacterial or eukaryotic.

Materials and Methods

Bacterial Strains, Media, and Growth Conditions. Strains used in this study are listed in Table S3. *Escherichia coli* was cultured in lysogeny broth (LB) or terrific broth (TB) at 37 °C. *Methylococcus capsulatus* Bath was cultured in nitrate mineral salts (NMS) medium (49) supplemented with 5 μM $CuSO_4$ and 100 μM ferric citrate at 37 °C. *Methylomicrobium alcaliphilum* 20Z was cultured at 30 °C in modified high salt NMS medium (HS-NMS, 1.5% NaCl, pH 9) as detailed in ref. 50. *Rhodopseudomonas palustris* TIE-1 was cultured in YPS-MOPS pH 7 at 30 °C (28). *Burkholderia phytofirmans* PsJN was cultured in 0.2X tryptic soy broth (TSB) at 30 °C. Methanotroph cultures were sealed in serum vials without removing the ambient air and provided ultrapure methane at 60 kPa over ambient pressure. All liquid cultures were incubated at the appropriate temperature with shaking at 225 RPM. For growth on solid medium, LB, NMS, HS-NMS, YPS-MOPS, or TSB was solidified with 1.5% agar. Media was supplemented, if necessary, with gentamicin (10 μg/mL for *E. coli*, *B. phytofirmans*, and methanotrophs; 800 μg/mL on plates and 400 μg/mL in liquid for *R. palustris*), carbenicillan (100 μg/mL), chloramphenicol (20 μg/mL), 600 μM diaminopimelic acid (DAP), or sucrose (5% for *M. alcaliphilum*; 10% for *R. palustris*). For conjugal transfer of plasmids, HS-NMS plates were supplemented with Difco nutrient broth and modified to reduce the salt to 0.2% and the pH to 8 (50). Methanotroph plates were incubated in Vacu-Quik Jars (Almore International, Inc.) filled with ultrapure methane at 20 kPa over ambient pressure. Two 40-mL cultures of *Desulfovibrio inopinatus* grown in DSMZ 196-13781 medium plus 14 mM lactate were generously supplied by the Bosak Laboratory (MIT).

Molecular Cloning Techniques. Plasmids and oligonucleotides used in this study are listed in Table S3. Details of molecular cloning techniques are described in *Supporting Information*.

Construction of Tetrahymanol Synthase Mutants. A homologous recombination-based selection/counterselection method previously used in *M. capsulatus* was used to delete the MEALZ_1626 locus in *M. alcaliphilum* (51). The deletion plasmid pJW002 was transferred into *M. alcaliphilum* via conjugation using the donor strain *E. coli* BW29427, a dap auxotroph, as described in refs. 50 and 51. Merodiploids with pJW002 integrated into the *M. alcaliphilum* chromosome were selected on HS-NMS plates containing gentamicin. Single gentamicin-resistant colonies were cultured without antibiotic for 3 d to allow excision of pJW002 from the chromosome and then plated on HS-NMS supplemented with 5% (wt/vol) sucrose to select for the loss of the *sacB* gene. MEALZ_1626 deletion mutants were identified by screening sucrose-resistant colonies by PCR with primers internal to the MEALZ_1626 and flanking the boundaries of the construct (Table S3). The same selection/counterselection method described above was used to delete Rpal_0860 in *R. palustris* TIE-1 as described in ref. 28 using plasmid pVW135 (Table S3).

Heterologous Expression. *M. capsulatus* Bath strains harboring pSRKGm or pSRKGm-*ths* (pABB238) were cultured in 50 mL NMS supplemented with gentamicin, and expression was induced at midexponential phase with 100 μM IPTG for 16 h. *B. phytofirmans* PsJN strains harboring pSRK-*lac*UV5 (pABB251) or pSRK-lacUV5-*ths* (pABB261) were cultured on 0.2X TSB plates supplemented with gentamicin and 100 μM IPTG for 48 h after which colonies were harvested from two plates for lipid extraction. *E. coli* DH10B strains harboring three plasmids, pTrc99a, pTrc-*sqs* (pABB303), or pTrc-*sqs-shc* (pABB305); pSRK-*lac*UV5 (pABB251) or pSRK-lacUV5-*ths* (pABB261); and pJBEI2997 (Addgene plasmid 35151) (41), were cultured in 5 mL TB supplemented with chloramphenicol, carbenicillan, and gentamicin until midexponential phase. Expression was induced with 500 μM IPTG for 69 h.

Lipid Analysis. Cultures were harvested by centrifugation at 4,500 × *g* at 4 °C for 10 min (50 mL of methanotrophs, 20 mL of *R. palustris*, 40 mL of *D. inopinatus*, or 5 mL of *E. coli*), and cell pellets were stored at −20 °C before lipid extraction. Cells were resuspended in 2 mL of deionized water and transferred to a solvent washed Teflon centrifuge tube. Five milliliters of methanol and 2.5 mL of dichloromethane were added, and the cell mixture was sonicated for 1 h. Ten milliliters of deionized water and 10 mL of dichloromethane were added to samples after sonication, mixed, and incubated at −20 °C overnight. Samples were centrifuged for 10 min at 2,800 × *g*, and the organic layer was transferred to a 40-mL baked glass vial. The total lipid extract was evaporated under N_2 and derivatized to acetate esters by incubating in 100 μL of 1:1 acetic anhydride:pyridine for 1 h at 70 °C. Samples were dried under N_2 after derivatization and resuspended in 100–200 μL of dichloromethane before analysis.

C-30 hopanoids, methylsterols, and tetrahymanol were analyzed via high-temperature gas chromatography-mass spectrometry (GC-MS) (52). Tetrahymanol was quantified by comparison with a cholestanol standard added to samples before derivatization and was normalized to the largest hopene peak detected in the sample. Lipid extracts were separated on an Agilent 7890B Series GC with helium as the carrier gas at a constant flow of 1.2 mL/min and programmed as follows: 100 °C for 2 min, ramp 15 °C/min to 320 °C, and hold 28 min. The first analysis of *M. alcaliphilum* extracts was done on a DB5-HT column (30 m × 0.25 mm i.d. × 0.1 μm film thickness) (Fig. S1). All subsequent analyses of *M. alcaliphilum* and other bacterial extracts were done on a DB17-HT column (30 m × 0.25 mm i.d. × 0.125 μm film thickness) to allow for better separation of the desmethyl and 2-methylhopanoids in *R. palustris* extracts. Two microliters of the sample were injected into a Gerstel-programmable temperature vaporization (PTV) injector, operated in splitless mode at 320 °C. The GC was coupled to a 5977A Series MSD with the source at 230 °C and operated at 70 eV in EI mode scanning from 50 to 850 Da in 0.5 s. Aminohopanoids were detected via liquid chromatography-mass spectrometry at the Vincent Coates Foundation Mass Spectrometry Laboratory, Stanford University (mass-spec.stanford.edu). All lipids were identified based on their retention time and comparison with previously published spectra (33, 52, 53).

Bioinformatics Analysis. The JGI Integrated Microbial Genomes (IMG) phylogenetic profiler (img.jgi.doe.gov) was used to identify *M. alcaliphilum* 20Z proteins with homologs found in the genomes of *R. palustris* TIE-1 and *B. japonicum* USDA110 but not found in *M. capsulatus* Bath, *Gluconacetobacter diazotrophicus* Pal5, *Methylosarcina lacus*, *Methylomonas methanica*, or *Burkholderia phytofirmans* PsJN. Tetrahymanol synthase homologs were detected in the IMG genomic and metagenomics databases by BLASTP (54). Geneious (Biomatters Limited) was used to align protein sequences from genomes via MUSCLE (55). Metagenomic Ths protein sequences were aligned to genomic Ths sequences via MAFFT (www.ebi.ac.uk/Tools/msa/mafft/) (56), and redundancy in the alignments was reduced through the Decrease Redundancy Program (web.expasy.org/decrease_redundancy/). Sequence similarities from MUSCLE-aligned sequences were rendered using ESPript 3.0 (espript.ibcp.fr/ESPript/ESPript/) (57). Large gaps from metagenomics sequence alignments were removed via GBLOCKS (molevol.cmima.csic.es/castresana/Gblocks_server.html) (58). Maximum likelihood trees were constructed by PhyML (59) using the LG+gamma model, four gamma rate categories, 10 random starting trees, NNI branch swapping, and substitution parameters estimated from the data. Ths trees were generated and edited by importing the resulting PhyML trees into iTOL (itol.embl.de/) (60).

ACKNOWLEDGMENTS. We thank Prof. T. Bosak and Dr. S. Zaarur for kindly providing us with cultures of *Desulfovibrio inopinatus* for lipid analysis and Profs. W. Metcalf, W. van der Donk, and C. D. Poulter for helpful discussions regarding isoprenoid biochemistry. We also thank Prof. A. Sessions, members of the P.V.W. Laboratory, and two anonymous reviewers for constructive comments on the manuscript and Dr. P. B. Welander for assistance with figure design. This work was supported by a grant from the National Science Foundation (EAR -1418831 to P.V.W.).

1. Castoreno AB, et al. (2005) Transcriptional regulation of phagocytosis-induced membrane biogenesis by sterol regulatory element binding proteins. *Proc Natl Acad Sci USA* 102(37):13129–13134.
2. Nes WD, Janssen GG, Crumley FG, Kalinowska M, Akihisa T (1993) The structural requirements of sterols for membrane function in *Saccharomyces cerevisiae*. *Arch Biochem Biophys* 300(2):724–733.
3. Parks LW, Casey WM (1995) Physiological implications of sterol biosynthesis in yeast. *Annu Rev Microbiol* 49:95–116.
4. Hieb WF, Rothstein M (1968) Sterol requirement for reproduction of a free-living nematode. *Science* 160(3829):778–780.
5. Tomazic ML, Poklepovich TJ, Nudel CB, Nusblat AD (2014) Incomplete sterols and hopanoids pathways in ciliates: Gene loss and acquisition during evolution as a source of biosynthetic genes. *Mol Phylogenet Evol* 74:122–134.
6. Conner RL, Landrey JR, Burns CH, Mallory FB (1968) Cholesterol inhibition of pentacyclic triterpenoid biosynthesis in *Tetrahymena pyriformis*. *J Protozool* 15(3):600–605.
7. Raederstorff D, Rohmer M (1988) Polyterpenoids as cholesterol and tetrahymanol surrogates in the ciliate *Tetrahymena pyriformis*. *Biochim Biophys Acta* 960(2):190–199.
8. Mallory FB, Conner RL, Gordon JT (1963) Isolation of a pentacyclic triterpenoid alcohol from a protozoan. *J Am Chem Soc* 85(9):1362–1363.
9. Kemp P, Lander DJ, Orpin CG (1984) The lipids of the rumen fungus *Piromonas communis*. *J Gen Microbiol* 130(1):27–37.
10. Takishita K, et al. (2012) Lateral transfer of tetrahymanol-synthesizing genes has allowed multiple diverse eukaryote lineages to independently adapt to environments without oxygen. *Biol Direct* 7:5.
11. Zander JM, Caspi E, Pandey GN, Mitra CR (1969) Presence of tetrahymanol in *Oleandra wallichii*. *Phytochemistry* 8(11):2265–2267.
12. Ten Haven HL, Rohmer M, Rullkotter J, Bisseret P (1989) Tetrahymanol, the most likely precursor of gammacerane, occurs ubiquitously in marine sediments. *Geochim Cosmochim Acta* 53(11):3073–3079.
13. Venkatesan MI (1989) Tetrahymanol—Its widespread occurrence and geochemical significance. *Geochim Cosmochim Acta* 53(11):3095–3101.
14. Cao CQ, et al. (2009) Biogeochemical evidence for euxinic oceans and ecological disturbance presaging the end-Permian mass extinction event. *Earth Planet Sci Lett* 281(3-4):188–201.
15. Sepulveda J, et al. (2009) Molecular isotopic evidence of environmental and ecological changes across the Cenomanian-Turonian boundary in the Levant Platform of central Jordan. *Org Geochem* 40(5):553–568.
16. Summons RE, et al. (1988) Distinctive hydrocarbon biomarkers from fossiliferous sediment of the Late Proterozoic Walcott Member, Chuar Group, Grand Canyon, Arizona. *Geochim Cosmochim Acta* 52(11):2625–2637.
17. Sinninghe Damste JS, et al. (1995) Evidence for gammacerane as an indicator of water column stratification. *Geochim Cosmochim Acta* 59(9):1895–1900.
18. Peters KE, Waters CC, Moldowan JM (2007) *Biomarkers and Isotopes in Petroleum Exploration and Earth History*, The Biomarker Guide (Cambridge Univ Press, Cambridge, UK), Vol 2.
19. Kannenberg EL, Perzl M, Hartner T (1995) The occurrence of hopanoid lipids in *Bradyrhizobium* bacteria. *FEMS Microbiol Lett* 127(3):255–261.
20. Kleemann G, et al. (1990) Tetrahymanol from the phototrophic bacterium *Rhodopseudomonas palustris*: First report of a gammacerane triterpene from a prokaryote. *J Gen Microbiol* 136(12):2551–2553.
21. Bravo JM, Perzl M, Härtner T, Kannenberg EL, Rohmer M (2001) Novel methylated triterpenoids of the gammacerane series from the nitrogen-fixing bacterium *Bradyrhizobium japonicum* USDA 110. *Eur J Biochem* 268(5):1323–1331.
22. Eickhoff M, Birgel D, Talbot HM, Peckmann J, Kappler A (2013) Oxidation of Fe(II) leads to increased C-2 methylation of pentacyclic triterpenoids in the anoxygenic phototrophic bacterium *Rhodopseudomonas palustris* strain TIE-1. *Geobiology* 11(3): 268–278.
23. Neubauer C, et al. (2015) Lipid remodeling in *Rhodopseudomonas palustris* TIE-1 upon loss of hopanoids and hopanoid methylation. *Geobiology* 13(5):443–453, 10.1111/gbi.12143.
24. Saar J, Kader JC, Poralla K, Ourisson G (1991) Purification and some properties of the squalene-tetrahymanol cyclase from *Tetrahymena thermophila*. *Biochim Biophys Acta* 1075(1):93–101.
25. Fischer WW, Pearson A (2007) Hypotheses for the origin and early evolution of triterpenoid cyclases. *Geobiology* 5(1):19–34.
26. Siedenburg G, Jendrossek D (2011) Squalene-hopene cyclases. *Appl Environ Microbiol* 77(12):3905–3915.
27. Silipo A, et al. (2014) Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nat Commun* 5:5106.
28. Welander PV, et al. (2009) Hopanoids play a role in membrane integrity and pH homeostasis in *Rhodopseudomonas palustris* TIE-1. *J Bacteriol* 191(19):6145–6156.
29. Perzl M, Müller P, Poralla K, Kannenberg EL (1997) Squalene-hopene cyclase from *Bradyrhizobium japonicum*: Cloning, expression, sequence analysis and comparison to other triterpenoid cyclases. *Microbiology* 143(Pt 4):1235–1242.
30. Kleemann G, Kellner R, Poralla K (1994) Purification and properties of the squalene-hopene cyclase from *Rhodopseudomonas palustris*, a purple non-sulfur bacterium producing hopanoids and tetrahymanol. *Biochim Biophys Acta* 1210(3):317–320.
31. Khmelenina VN, Kalyuzhnaya MG, Starostina NG, Suzina NE, Trotsenko YA (1997) Isolation and characterization of halotolerant alkaliphilic methanotrophic bacteria from Tuva Soda Lakes. *Curr Microbiol* 35(5):257–261.
32. Vuilleumier S, et al. (2012) Genome sequence of the haloalkaliphilic methanotrophic bacterium *Methylomicrobium alcaliphilum* 20Z. *J Bacteriol* 194(2):551–552.
33. Talbot HM, Watson DF, Murrell JC, Carter JF, Farrimond P (2001) Analysis of intact bacteriohopanepolyols from methanotrophic bacteria by reversed-phase high-performance liquid chromatography-atmospheric pressure chemical ionisation mass spectrometry. *J Chromatogr A* 921(2):175–185.
34. Volkman JK (2003) Sterols in microorganisms. *Appl Microbiol Biotechnol* 60(5): 495–506.
35. Blumenberg M, et al. (2006) Biosynthesis of hopanoids by sulfate-reducing bacteria (genus *Desulfovibrio*). *Environ Microbiol* 8(7):1220–1227.
36. Neunlist S, Rohmer M (1985) Novel hopanoids from the methylotrophic bacteria *Methylococcus capsulatus* and *Methylomonas methanica*. (22S)-35-amino-bacteriohopane-30,31,32,33,34-pentol and (22S)-35-amino-3 beta-methylbacteriohopane-30,31,32,33,34-pentol. *Biochem J* 231(3):635–639.
37. Cvejic JH, et al. (2000) Bacterial triterpenoids of the hopane series as biomarkers for the chemotaxonomy of *Burkholderia, Pseudomonas* and *Ralstonia* spp. *FEMS Microbiol Lett* 183(2):295–299.
38. Schmerk CL, et al. (2015) Elucidation of the *Burkholderia cenocepacia* hopanoid biosynthesis pathway uncovers functions for conserved proteins in hopanoid-producing bacteria. *Environ Microbiol* 17(3):735–750.
39. Hoshino T, Kondo T (1999) The cyclization mechanism of squalene in hopene biosynthesis: The terminal methyl groups are critical to the correct folding of this substrate both for the formation of the five-membered E-ring and for the initiation of the polycyclization reaction. *Chem Commun* 8:731–732.
40. Katabami A, et al. (2015) Production of squalene by squalene synthases and their truncated mutants in *Escherichia coli*. *J Biosci Bioeng* 119(2):165–171.
41. Peralta-Yahya PP, et al. (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nat Commun* 2:483.
42. Harvey HR, Mcmanus GB (1991) Marine ciliates as a widespread source of tetrahymanol and hopan-3-β-ol in sediments. *Geochim Cosmochim Acta* 55(11):3387–3390.
43. Chistoserdova L (2015) Methylotrophs in natural habitats: Current insights through metagenomics. *Appl Microbiol Biotechnol* 99(14):5763–5779.
44. Foti M, et al. (2007) Diversity, activity, and abundance of sulfate-reducing bacteria in saline and hypersaline soda lakes. *Appl Environ Microbiol* 73(7):2093–2100.
45. Karr EA, et al. (2005) Diversity and distribution of sulfate-reducing bacteria in permanently frozen Lake Fryxell, McMurdo Dry Valleys, Antarctica. *Appl Environ Microbiol* 71(10):6353–6359.
46. Oswald K, et al. (2015) Light-dependent aerobic methane oxidation reduces methane emissions from seasonally stratified lakes. *PLoS One* 10(7):e0132574.
47. Wakeham SG, et al. (2007) Microbial ecology of the stratified water column of the Black Sea as revealed by a comprehensive biomarker study. *Org Geochem* 38(12): 2070–2097.
48. Muyzer G, Stams AJ (2008) The ecology and biotechnology of sulphate-reducing bacteria. *Nat Rev Microbiol* 6(6):441–454.
49. Whittenbury R, Phillips KC, Wilkinson JF (1970) Enrichment, isolation and some properties of methane-utilizing bacteria. *J Gen Microbiol* 61(2):205–218.
50. Ojala DS, Beck DA, Kalyuzhnaya MG (2011) Genetic systems for moderately halo(alkali)philic bacteria of the genus *Methylomicrobium*. *Methods Enzymol* 495:99–118.
51. Welander PV, Summons RE (2012) Discovery, taxonomic distribution, and phenotypic characterization of a gene required for 3-methylhopanoid production. *Proc Natl Acad Sci USA* 109(32):12905–12910.
52. Sessions AL, et al. (2013) Identification and quantification of polyfunctionalized hopanoids by high temperature gas chromatography-mass spectrometry. *Org Geochem* 56:120–130.
53. Talbot HM, Squier AH, Keely BJ, Farrimond P (2003) Atmospheric pressure chemical ionisation reversed-phase liquid chromatography/ion trap mass spectrometry of intact bacteriohopanepolyols. *Rapid Commun Mass Spectrom* 17(7):728–737.
54. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389–3402.
55. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797.
56. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30(4):772–780.
57. Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 31(13):3320–3323.
58. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17(4):540–552.
59. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52(5):696–704.
60. Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* 23(1):127–128.
61. Li MZ, Elledge SJ (2012) SLIC: A method for sequence- and ligation-independent cloning. *Methods Mol Biol* 852:51–59.
62. Quandt J, Hynes MF (1993) Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. *Gene* 127(1):15–21.
63. Khan SR, Gaines J, Roop RM, 2nd, Farrand SK (2008) Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl Environ Microbiol* 74(16):5053–5062.
64. Amann E, Ochs B, Abel KJ (1988) Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* 69(2):301–315.
65. Marx CJ, Lidstrom ME (2002) Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria. *Biotechniques* 33(5):1062–1067.