C-4 sterol demethylation enzymes distinguish bacterial and eukaryotic sterol synthesis

Alysha K. Lee[a], Amy B. Banta[a], Jeremy H. Wei[a], David J. Kiemle[b], Ju Feng[b], José-Luis Giner[b,1], and Paula V. Welander[a,1]

[a]Department of Earth System Science, Stanford University, Stanford, CA 94305; and [b]Department of Chemistry, State University of New York–Environmental Science and Forestry, Syracuse, NY 13210

Edited by Andrew H. Knoll, Harvard University, Cambridge, MA, and approved May 4, 2018 (received for review February 15, 2018)

Sterols are essential eukaryotic lipids that are required for a variety of physiological roles. The diagenetic products of sterol lipids, sterane hydrocarbons, are preserved in ancient sedimentary rocks and are utilized as geological biomarkers, indicating the presence of both eukaryotes and oxic environments throughout Earth's history. However, a few bacterial species are also known to produce sterols, bringing into question the significance of bacterial sterol synthesis for our interpretation of sterane biomarkers. Recent studies suggest that bacterial sterol synthesis may be distinct from what is observed in eukaryotes. In particular, phylogenomic analyses of sterol-producing bacteria have failed to identify homologs of several key eukaryotic sterol synthesis enzymes, most notably those required for demethylation at the C-4 position. In this study, we identified two genes of previously unknown function in the aerobic methanotrophic γ-Proteobacterium *Methylococcus capsulatus* that encode sterol demethylase proteins (Sdm). We show that a Rieske-type oxygenase (SdmA) and an NAD(P)-dependent reductase (SdmB) are responsible for converting 4,4-dimethylsterols to 4α-methylsterols. Identification of intermediate products synthesized during heterologous expression of SdmA-SdmB along with ^{13}C-labeling studies support a sterol C-4 demethylation mechanism distinct from that of eukaryotes. SdmA-SdmB homologs were identified in several other sterol-producing bacterial genomes but not in any eukaryotic genomes, indicating that these proteins are unrelated to the eukaryotic C-4 sterol demethylase enzymes. These findings reveal a separate pathway for sterol synthesis exclusive to bacteria and show that demethylation of sterols evolved at least twice—once in bacteria and once in eukaryotes.

sterols | biomarkers | cyclic triterpenoids | *Methylococcus capsulatus*

Sterols, such as cholesterol, are tetracyclic isoprenoid lipids that play critical roles in a variety of eukaryotic physiological processes, including intra- and intercellular signaling, stress tolerance, maintenance of cell membrane integrity, and human disease (1–3). Sterols can also be preserved as sterane hydrocarbons in sedimentary rocks, where they function as biomarkers or "molecular fossils," indicating the presence of eukaryotes and/or oxic conditions deep in time (4). Studies that incorporate sterane biomarkers with other geological proxies have provided significant insight into ancient climatic events, mass extinctions, and various evolutionary transitions throughout Earth's history (5, 6).

Given the essential roles that sterols play in eukaryotic physiology and their utility as geological biomarkers, understanding the synthesis, evolution, and function is of great importance. This is particularly critical when attempting to constrain and assess the provenance of sterane biomarkers encountered in the rock record. However, one aspect of sterol synthesis that has not been studied as extensively is the production by bacterial species. Lipid analyses have shown that 19 bacterial strains from diverse phyla produce sterols (7–15). Phylogenetic analyses of oxidosqualene cyclase (Osc), the enzyme required for the initial cyclization of oxidosqualene to lanosterol or cycloartenol, suggest that bacterial sterol production may be widespread (9, 16). These studies also point to a complex evolutionary history of sterol synthesis in bacteria, in which acquisition of sterol biosynthetic genes via horizontal gene transfer from ancient eukaryotes is evident in some bacterial lineages but not in others (13, 17–19). It is also unclear what functional role sterols have in bacterial cells and what significance bacterial sterol production could have on our interpretation of sterane biomarkers.

In a previous study, we showed that some bacterial species produce biochemically modified sterols, but the eukaryotic enzymes required for these modifications are not encoded in these bacterial genomes (9). In particular, several γ-Proteobacterial aerobic methanotrophs remove one C-4 methyl group to produce 4α-methylsterols, while some δ-Proteobacterial myxobacteria remove both C-4 methyl groups (7–9). In plants, vertebrates, and fungi, C-4 demethylation is an O_2-dependent mechanism catalyzed by three enzymes: a nonheme iron-dependent C-4 sterol methyl oxidase (ERG25/SMO), an NAD(P)-dependent 3β-hydroxysteroid dehydrogenase/C-4 decarboxylase (ERG26/3β-HSD/D), and an NADPH-dependent 3-ketosteroid reductase (ERG27/3-SR) (*SI Appendix*, Fig. S1) (20–22). In yeast and vertebrates, the same three enzymes remove both C-4 methyl groups iteratively, while plants have two distinct ERG25 homologs, SMO1 and SMO2, that oxidize 4,4-dimethylsterols and 4α-methylsterols, respectively (*SI Appendix*, Fig. S1) (23). However, homologs of ERG25, ERG26, and ERG27 are not present in the genomes of aerobic methanotrophs capable of C-4 sterol

Significance

Sterols are essential eukaryotic lipids that can be preserved as steroids in sedimentary rocks for billions of years. Because eukaryotes are the predominant modern day producers of these lipids, fossilized sterols are used as geological biomarkers for the presence of specific eukaryotes in ancient environments. Sterol lipids are also produced by a few bacteria, but the biosynthesis and function of sterols in bacteria are not as well-understood. In this study, we used a combination of bioinformatics and lipid analyses to identify bacterial sterol synthesis proteins. Our results indicate that bacteria have evolved distinct aspects of the sterol synthesis pathway independent of eukaryotes and show that exploring sterol physiology in bacteria can provide insight into this geologically relevant pathway.

Author contributions: A.K.L., A.B.B., J.H.W., J.-L.G., and P.V.W. designed research; A.K.L., A.B.B., J.H.W., J.-L.G., and P.V.W. performed research; D.J.K., J.F., and J.-L.G. contributed new reagents/analytic tools; A.K.L., A.B.B., J.H.W., J.-L.G., and P.V.W. analyzed data; and A.K.L., A.B.B., J.H.W., J.-L.G., and P.V.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: jlginer@syr.edu or welander@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1802930115/-/DCSupplemental.

Published online May 21, 2018.

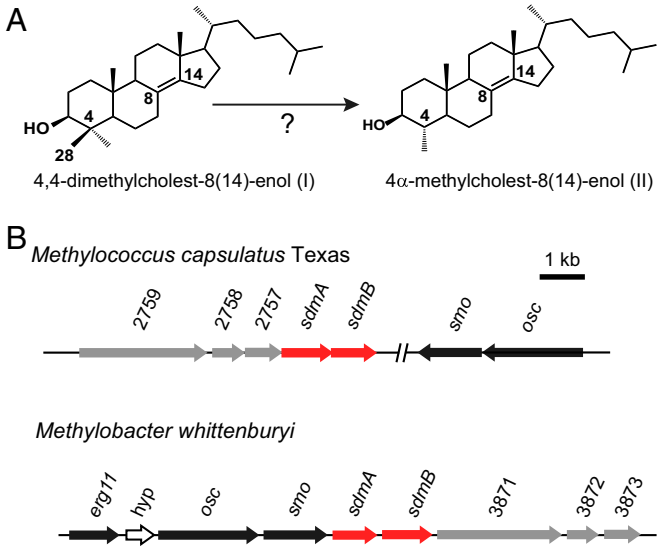

Fig. 1. Identification of putative sterol methylase genes in *M. capsulatus* Texas. (*A*) Two sterols produced by *M. capsulatus*. The enzymes required for demethylation of 4,4-dimethylcholest-8 (14)-enol to 4α-methylcholest-8 (14)-enol in bacteria are not known. Homologs of the enzymes that catalyze this step of sterol synthesis in eukaryotes are not found in bacterial genomes. (*B*) Genomic context of two putative C-4 demethylase genes, *sdmA* and *sdmB* (red arrows), in two aerobic methanotrophs. In *M. whittenburyi*, *sdmA* and *sdmB* are within a putative operon that contains other sterol biosynthesis genes (black arrows). Gray arrows represent homologs in *M. capsulatus* and *M. whittenburyi*, which tend to co-occur with *sdmA* and *sdmB* but are not thought to be involved in the synthesis of sterols. *erg11*, sterol C-14 demethylase; hyp, hypothetical protein; *sdmA*, *M. capsulatus* Texas locus tag H156DRAFT_2756; *sdmB*, *M. capsulatus* Texas locus tag H156DRAFT_2755.

demethylation (9). In the myxobacterium *Plesiocystis pacifica*, phylogenomic analyses identified putative homologs of ERG25 and ERG26, but the sequence identities are low, and no ERG27 candidate was found (18). Therefore, we hypothesized that bacteria demethylate sterols at the C-4 position through a distinct mechanism. Here, we identify the enzymes necessary for C-4 sterol demethylation in the aerobic methanotroph *Methylococcus capsulatus*. These proteins are phylogenetically and biochemically distinct from eukaryotic C-4 demethylation enzymes and are present in all bacteria that produce C-4 demethylated sterols. Our findings indicate that C-4 sterol demethylation in bacteria has evolved independently of the eukaryotic pathway.

Results

Identification of C-4 Sterol Demethylase Proteins in *M. capsulatus*. We used the Joint Genome Institute (JGI) Integrated Microbial Genomes (IMG) Phylogenetic Profiler to identify *M. capsulatus* proteins that have homologs in three other C-4 demethylating methanotrophs (*Methylobacter whittenburyi*, *Methylobacter luteus*, and *Methylosarcina lacus*) but are absent from bacteria that do not demethylate their sterols (*Methylomicrobium alcaliphilum*) or do not produce sterols at all (*Methylomonas methanica*, *Methylobacterium extorquens*, and *Escherichia coli*). We found 17 candidate genes unique to the C-4 demethylating methanotrophs, several of which are annotated as hypothetical proteins with no known function (*SI Appendix*, Table S1). Five of these genes form a putative operon that, in *M. whittenburyi*, is localized downstream of three sterol biosynthesis genes (Fig. 1). We utilized InterProScan (24) to identify putative protein motifs in these five candidate proteins. One candidate, SdmA, contains a Rieske-type oxygenase motif, and a second, SdmB, has an NAD(P) binding Rossmann-fold domain. Given that the removal of the C-4 methyl group in *M. capsulatus* could occur through an oxidative demethylation, we hypothesized that these two candidate proteins may be involved.

Deletion of *M. capsulatus sdmA* and *sdmB* resulted in complete loss of sterol C-4 demethylation (Fig. 2 and *SI Appendix*, Fig. S2). Expression of SdmA-SdmB together, but not individually, from a separate chromosomal location restored C-4 demethylation in this double mutant, indicating that both proteins are necessary for this reaction. To establish whether SdmA-SdmB is

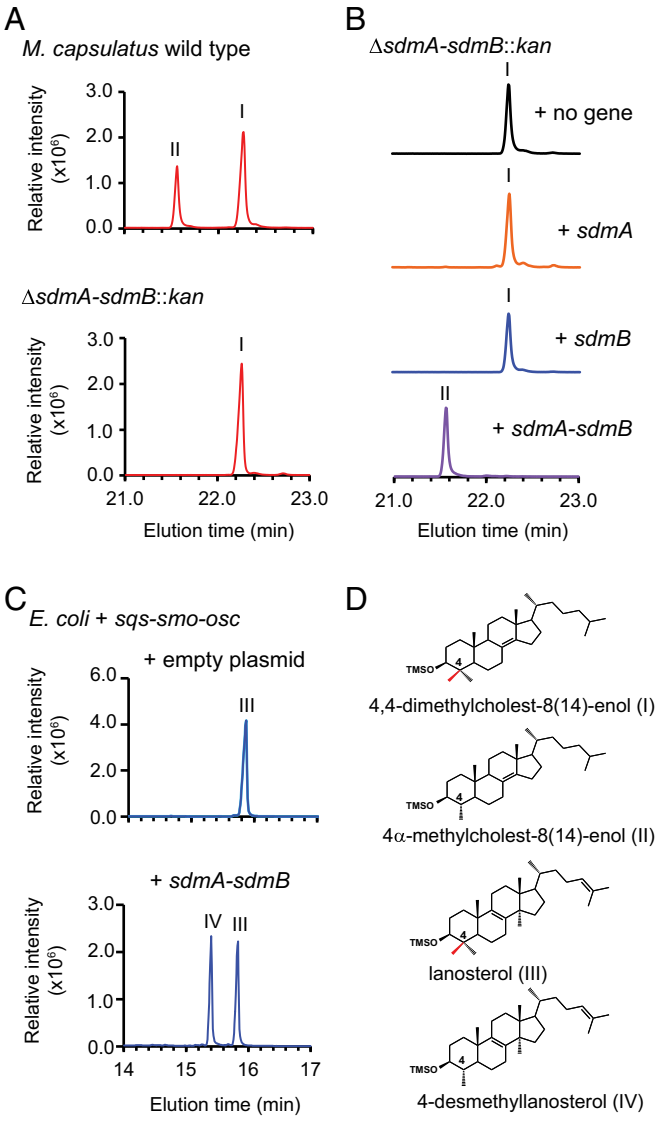

Fig. 2. Sterol demethylase proteins SdmA and SdmB are necessary and sufficient for C-4 demethylation. (*A*) Total ion chromatogram of alcohol-soluble lipid fractions isolated from the *M. capsulatus* wild-type strain (*Upper*) and the *M. capsulatus sdmA-sdmB* deletion mutant (*Lower*). The wild type produces both 4,4-dimethylcholest-8 (14)-enol (I) and 4α-methylcholest-8 (14)-enol (II), while the mutant only produces 4,4-dimethylcholest-8 (14)-enol (I). (*B*) Complementation of the *M. capsulatus sdmA-sdmB* deletion mutant shows that both *sdmA* and *sdmB* are required to restore C-4 demethylation. (*C*) Total ion chromatograms of alcohol-soluble lipid fractions from lanosterol (III) producing *E. coli* harboring an empty plasmid (*Upper*) or a plasmid with *sdmA-sdmB* (*Lower*). Expression for SdmA-SdmB results in the production of 4-desmethyllanosterol (IV), showing that these proteins are sufficient for C-4 demethylation. (*D*) Structure of lipids identified in *A*–*C*. Mass spectra of identified lipids are shown in *SI Appendix*, Fig. S2. All lipids were derivatized with trimethylsilyl.

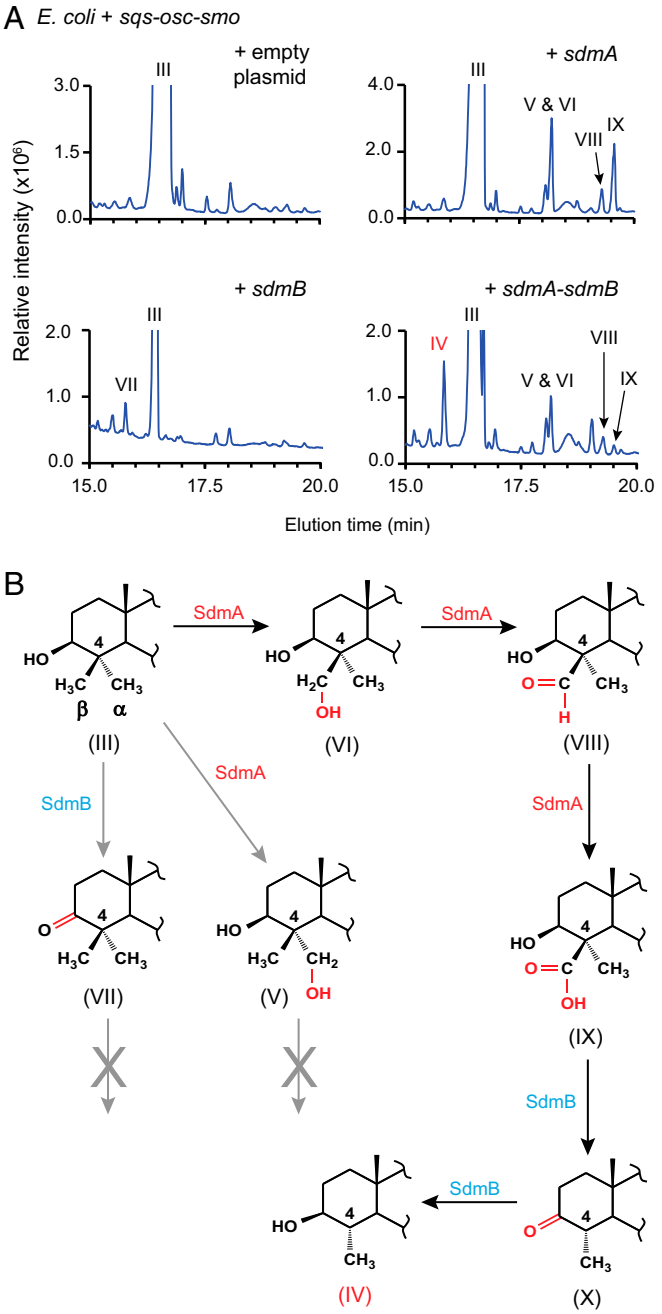

Fig. 3. C-4 sterol demethylation intermediates suggest a distinct bacterial demethylation mechanism. (*A*) Total ion chromatograms of TLEs from *E. coli* expressing *sdmA* and *sdmB* together and separately showing the production of C-4 demethylation intermediates (V, VI, VII, VIII, and IX) in addition to the final 4-desmethyllanosterol product (IV; red). Lipids were derivatized with trimethylsilyl. Mass spectra and full structures of sterols are shown in *SI Appendix*, Fig. S2. All structures were confirmed by NMR. (*B*) Proposed C-4 demethylation mechanism in aerobic methanotrophs based on reaction intermediates observed in *A* and *B*. SdmA converts the 4β-methyl group into first a hydroxyl (VI), then an aldehyde (VIII), and finally, a carboxylic acid (IX). SdmB is proposed to oxidize the C-3 hydroxyl group of the 4β-carboxylic acid, which undergoes decarboxylation to give 4-desmethyllanosterone (X). SdmB also functions as a ketoreductase that converts the C-3 ketone back to a hydroxyl group after C-4 demethylation (IV). SdmB also catalyzes the oxidation of lanosterol to lanosterone (VII) in a side reaction not related to C-4 demethylation (gray arrows). The 4α-28-hydroxy (V) intermediate is also a potential side product generated by a promiscuous SdmA in a side reaction (gray arrows); 4-desmethyllanosterone (X) was only observed in the NMR experiments. III, lanosterol; IV, 4-desmethyllanosterol (4α,14α-dimethylzymosterol); V, 4α-28-hydroxylanosterol; VI, 4β-28-hydroxylanosterol; VII, lanosterone (4α,14α-dimethylzymosterone); VIII, 4β-28-oxolanosterol; IX, 4β-28-carboxylanosterol (4β-3β-hydroxylanosta-8,24-dien-28-oic acid).

sufficient for sterol C-4 demethylation, we heterologously expressed both proteins in an *E. coli* strain producing lanosterol, which has two methyl groups at C-4 and one at C-14 (10). Coexpression of both proteins in *E. coli* resulted in removal of one C-4 methyl from lanosterol to generate 4-desmethyllanosterol [4α,14α-dimethylzymosterol (IV)], indicating that no other *M. capsulatus* proteins are needed for this reaction (Fig. 2).

Sterol Intermediates Reveal Bacterial C-4 Demethylation Mechanism. In addition to showing that SdmA-SdmB is sufficient for C-4 demethylation, our heterologous expression experiments also revealed a series of probable biosynthetic intermediates of the sterol demethylase reaction that we identified by NMR (Fig. 3 and *SI Appendix*, Tables S2 and S3). Expression of SdmA alone results in the production of a 4β-carboxylic acid (IX), a 4β-aldehyde (VIII), and two 4-hydroxylmethyl compounds (V and VI). Based on these intermediates, we hypothesize that SdmA oxidizes the 4β-methyl group first to an alcohol (VI), then to an aldehyde (VIII), and finally, to the carboxylic acid (IX) (Fig. 3). Thus, these data indicate that bacterial C-4 demethylation proceeds via decarboxylation as it does in eukaryotes and does not involve deformylation as has been proposed for C-14 demethylation (25).

Expression of SdmB alone results in the production of only one product, lanosterone (VII), suggesting that SdmB oxidizes the C-3 hydroxyl group to a ketone much as the nonhomologous ERG26 does. Specifically, we propose that SdmB oxidizes the C-3 hydroxyl group of the 4β-carboxylic intermediate, which then undergoes decarboxylation with or without enzymatic assistance to yield 4-desmethyllanosterone (X) (Fig. 3). The C-4 demethylation reaction is completed by reduction of the C-3 ketone to yield the final sterol product, 4-desmethyllanosterol (IV). Many NAD(P)-dependent ketosteroid reductases also catalyze the reverse reaction (26), and therefore, SdmB may affect both the decarboxylative oxidation and the reduction steps at C-3. If so, this would explain why SdmA and SdmB are sufficient for C-4 demethylation and why an ERG27 homolog would not be required in bacteria for the final reduction of the C-3 ketone.

SdmA-SdmB Removes the 4β-Methyl Group Rather than the 4α-Methyl Group. Our structural analyses also indicate that the SdmA-SdmB pair removes the 4β-methyl rather than the 4α-methyl targeted by the eukaryotic C-4 demethylase enzymes (27). To confirm this, we expressed SdmA-SdmB in the presence of an isotopically labeled isoprenoid precursor, (1-^{13}C)-1-deoxy-D-xylulose (DOX) (28). The ^{13}C label was incorporated into lanosterol via the native *E. coli* methylerythritol phosphate isoprenoid pathway (29, 30) with specific labeling of the 4β-methyl but not the 4α-methyl (Table 1 and *SI Appendix*, Fig. S3 and Table S4). When expressing SdmA-SdmB under these conditions, the single remaining C-4 methyl of 4-desmethyllanosterol is unlabeled, confirming that the bacterial enzymes remove the 4β-methyl group. Although some of the 4α-alcohol (V) was produced during heterologous expression of *M. capsulatus* SdmA or SdmA-SdmB together, we did not observe further oxidation of the 4α-alcohol or retention of label in the DOX study, indicating that this is likely a side product of a promiscuous SdmA (Fig. 3). It should also be noted that small amounts of 4-desmethyllanosterol were detected in the experiment with SdmA alone (*SI Appendix*, Table S2), which suggests that an *E. coli* enzyme or SdmA itself may be capable of oxidizing the 3-hydroxy group to a ketone.

Table 1. *M. capsulatus sdmA-sdmB* removes the 4β-methyl group, not the 4α-methyl, of lanosterol

Sterol product	Position	^{1}H, ppm	^{13}C, ppm	Unlabeled average integral	^{13}C-labeled average integral	Enrichment, %
Lanosterol	4β-Methyl	0.813	15.30	1.193 ± 0.005	1.832 ± 0.004	0.536 ± 0.006
Lanosterol	4α-Methyl	1.002	27.90	1.288 ± 0.002	1.300 ± 0.004	0.009 ± 0.003
4-Desmethyllanosterol	4-Methyl	0.999	15.00	1.221 ± 0.030	1.133 ± 0.031	−0.072 ± 0.036

DOX was fed to *E. coli* expressing bacterial lanosterol synthesis genes and *M. capsulatus sdmA-sdmB* (*SI Appendix*, Fig. S3). All measurements were done in triplicate, and SDs were calculated from technical replicates. The integrals are referenced to the E-methyl group of the side chain (C26/27). Enrichment, % = (L − UL)/L, where L is the integral of methyl signal in labeled molecule and UL is the integral of methyl signal in unlabeled molecule. *SI Appendix*, Table S4 shows the full dataset showing incorporation of ^{13}C into other methyl positions of lanosterol and 4-desmethyllanosterol.

The preference for the 4β-methyl by SdmA-SdmB is noteworthy for two reasons. First, decarboxylation of the 4β-carboxylic acid is stereoelectronically favored over the 4α-carboxylic acid, indicating that the removal of the 4β-methyl in bacteria may be more favorable than the removal of the 4α-methyl in eukaryotes. This preference may be advantageous in the bacterial system, particularly as it is unclear whether SdmB functions like ERG26 in eukaryotes catalyzing both the decarboxylation at C-4 and the requisite oxidation of the 3-hydroxyl (23). Second, preferential removal of the 4β-methyl group by SdmA-SdmB explains why *M. capsulatus* does not fully demethylate sterols at C-4 like eukaryotes. In eukaryotes, ERG25 oxidizes the 4α-methyl, and subsequent decarboxylation by ERG26 occurs concomitantly with epimerization of the remaining 4β-methyl into the 4α position (22, 27). In this case, the remaining 4α-methyl is available for a second demethylation by the same set of eukaryotic enzymes (22). However, in bacteria, the preference for the 4β-methyl by SdmA-SdmB precludes a second round of demethylation.

SdmA-SdmB Is Phylogenetically Distinct from Eukaryotic Enzymes. To determine if SdmA-SdmB is found in other sterol-producing organisms, we searched for homologs in the JGI IMG genomic database (BLAST cutoff: 30% identity, e value $1e^{-50}$). We identified 56 putative *M. capsulatus* SdmA homologs and 352 SdmB homologs in bacterial genomes but none in either eukaryotic or archaeal genomes. The larger number of putative SdmB homologs retrieved possibly reflects the high sequence similarity of the NAD(P) binding Rossmann-fold domain found in this protein to other functionally distinct bacterial oxidoreductases (31). Maximum likelihood phylogenetic trees of each protein showed that homologs most similar to *M. capsulatus* SdmA (Fig. 4) or SdmB (*SI Appendix*, Fig. S4) clustered in a distinct clade composed of 23 bacteria. Because sequence homology alone may not be sufficient to infer functional similarity between two proteins, we searched the 23 genomes within the SdmA and SdmB clades for other sterol synthesis proteins. We found that these 23 genomes also encode an Osc homolog required for sterol synthesis (*SI Appendix*, Table S5). In addition, several of these bacteria are known to demethylate sterols at the C-4 position (8, 9, 12), indicating a strong correlation between the co-occurrence of Osc, SdmA, and SdmB in a bacterial genome and C-4 demethylation of sterols. Finally, SdmA homologs were identified in a variety of freshwater, marine, terrestrial, and host-associated metagenomes (BLAST cutoff: 20% identity, e value $1e^{-50}$) in which we also find Osc homologs (9), suggesting that this sterol C-4 demethylation mechanism may be more widespread in the bacterial domain than what is observed in cultured organisms (*SI Appendix*, Fig. S5).

Discussion

The results presented here show that sterol C-4 demethylation has evolved at least twice—once in bacteria and once in eukaryotes. Although C-4 sterol demethylation in eukaryotes and bacteria occurs through similar oxidative reactions, there are important distinctions between the two pathways. Our heterologous expression and structural analyses revealed that bacterial C-4 demethylation in *M. capsulatus* only requires two proteins rather than three and differs significantly in its stereochemical mechanism from the eukaryotic system. In addition, homologs of the bacterial SdmA-SdmB demethylase proteins are unrelated to the eukaryotic sterol demethylase proteins and are exclusive to bacterial genomes. It is also possible that additional examples of convergent evolution in the sterol biosynthesis pathway remain to be discovered, as enzymes catalyzing other key steps in bacterial sterol synthesis have yet to be identified (9, 18). For example, a C-14 reductase homolog is not present in several aerobic methanotroph genomes, and the enzyme responsible for generating the unusual nuclear C-8 (14) desaturation in *M. capsulatus* sterols is unknown (32) (*SI Appendix*, Fig. S6). Furthermore, several myxobacterial species have been shown to completely demethylate their sterols at the C-4 position (9, 11, 12). Here, we show that these myxobacterial species harbor SdmA-SdmB homologs, suggesting that either the myxobacterial SdmA-SdmB homologs are modified to demethylate twice or that a second set of proteins may be required to further demethylate at C-4 (*SI Appendix*, Fig. S6). Identification and characterization of these potentially new bacterial sterol biosynthesis enzymes could further underscore the mechanistic and evolutionary differences between the bacterial and eukaryotic sterol pathways.

The convergent evolution of sterol C-4 demethylation also suggests that this modification carries physiological significance in both eukaryotes and bacteria. Demethylation of both C-4 methyl groups is essential for many sterol functions in eukaryotes, and as a result, deletion of eukaryotic sterol demethylase genes in yeast, human cell lines, and plants is lethal (20, 33–35). However, deletion of SdmA-SdmB is not lethal in *M. capsulatus*, and it is unclear what functional role this specific modification and sterols in general serve in bacteria. Bacterial sterols have been proposed to affect membrane rigidity and permeability as they do in eukaryotes, but studies showing that sterols make up a small fraction of bacterial membrane lipids belie this structural role (36). Furthermore, eliminating sterol production in the myxobacterium *Stigmatella aurantiaca* did not result in a severe growth phenotype under normal laboratory conditions, indicating that bacterial sterols may have a more nuanced physiological role (8). In aerobic methanotrophs, the production of 4-methylsterols raises the possibility that C-4 demethylation may play a regulatory role as has been observed in some eukaryotes. C-4–methylated steroids are required for dauer larval formation in the nematode *Caenorhabditis elegans*, and accumulation of 4,4-dimethylsterols in fission yeast functions as an indicator of low oxygen conditions, resulting in transcriptional activation of hypoxia stress response factors (37, 38). Because C-4 demethylation by SdmA-SdmB requires oxygen, inhibition of C-4 demethylation under suboxic conditions and the subsequent accumulation of 4,4-dimethylsterols may also serve as an indicator of oxygen depletion in *M. capsulatus* (39). Understanding the role of SdmA-SdmB in

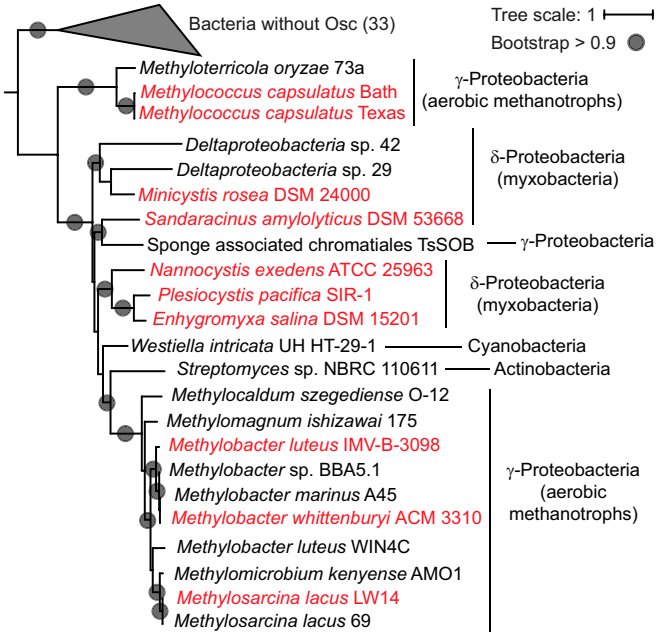

Fig. 4. Maximum phylogenetic tree of SdmA. The SdmA homologs from the 23 organisms shown cluster into one clade, and the genomes from each of these organisms also encode an Osc homolog. The 10 strains highlighted in red produce sterols that are demethylated at the C-4 positon at least once. Gray circles indicate branches that have bootstrap values greater than 0.9. The collapsed clade contains putative Rieske oxygenases from organisms that do not encode an Osc homolog in their genome and most likely do not produce sterols. The number in parentheses indicates the number of proteins in the collapsed clade. (Scale bar: one change per nucleotide site.)

demethylation provides us with a foundation to begin exploring such physiological hypotheses more directly.

Finally, while our discovery of SdmA-SdmB highlights the evolutionary divergence of sterol synthesis between bacteria and eukaryotes, it does not alter the current interpretation of sterane biomarkers in the rock record per se. The oldest known biomarker assemblages, found in the 1.6 billion-year-old Barney Creek Formation in northern Australia, contain aromatized 4-methylsteroids, and those are already attributed to aerobic methanotrophs, like *M. capsulatus* (40). Furthermore, the majority of sterane hydrocarbons used to infer the occurrence of eukaryotes deep in time are the diagenetic products of complex sterols with modified side chains, such as stigmasterol (5, 17). To date, no bacterium has been shown to produce sterols with side-chain alkylations, and therefore, these complex sterols remain unique to eukaryotes. However, our discovery of SdmA-SdmB emphasizes that our ability to predict how a bacterium modifies its sterols from genomic data is hindered by a lack of knowledge of the bacterial enzymes required to carry out these reactions—a bacterium may be capable of complex sterol production with enzymes not homologous to those found in eukaryotes. Furthermore, as we noted previously, bacteria may have other divergent sterol synthesis enzymes in addition to SdmA-SdmB, which could result in distinct isotopic signatures between eukaryotic and bacterial sterols. These isotopic signatures could potentially be traced in the geological record and would allow for the differentiation of eukaryotic and bacterial sterols directly. Thus, fully exploring the production of sterols by diverse bacteria, characterizing the bacterial proteins required for sterol synthesis, and studying the isotopic fractionation resulting from these synthesis pathways will not only advance our understating of sterol function and evolution but will also better inform interpretation of sterane biomarkers in the rock record.

Methods

Bacterial Culture. Strains used in this study are listed in *SI Appendix*, Table S6. *M. capsulatus* Texas (ATCC 19069) was cultured in Nitrate Minimal Salts medium (41) supplemented with 5 μM copper(II) sulfate, 30 μM ferric citrate, and if necessary, gentamycin (10 μg/mL) or kanamycin (30 μg/mL) at 37 °C in an atmosphere of air plus 1–5 psi methane either in serum vials sealed with butyl rubber stoppers and shaking at 225 rpm (Thermo Scientific, MaxQ8000) or on plates solidified with 1.5% agar in a sealed container (Vacu-Quik Jar; Almore International, Inc.). *E. coli* DH10B was cultured in LB or terrific broth (TB) at 30 °C or 37 °C with shaking at 225 rpm and supplemented, if necessary, with gentamycin (15 μg/mL), kanamycin (30 μg/mL), carbenicillin (100 μg/mL), and/or chloramphenicol (20 μg/mL).

Molecular Cloning. All plasmids and oligonucleotides used in this study are described in *SI Appendix*, Tables S7 and S8. Details of cloning and mutagenesis procedures are described in *SI Appendix*.

Deletion and Complementation of *M. capsulatus* Texas H156DRAFT2756-2755. *M. capsulatus* Texas Δ2756–2755::*kan* (PVW1823) was constructed by replacement of the H156DRAFT 2756–2755 coding sequence with a kanamycin resistance gene by electroporation of linear DNA followed by homologous recombination and selection. For complementation, deleted genes were expressed in trans from an isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible expression cassette located on the *M. capsulatus* chromosome at a location with convergent gene transcription following H156DRAFT_2433. Details are provided in *SI Appendix, Methods*.

Heterologous Expression in *E. coli*. All heterologous expression strains are shown in *SI Appendix*, Table S9. Lipid synthesis genes were overexpressed in *E. coli* from multiple compatible plasmids with either IPTG-inducible *lac* or arabinose-inducible *araBAD* promoters. Lanosterol was synthesized by *E. coli* DH10B overexpressing the mevalonate pathway from plasmid pJBEI2997 (42) as well as *M. alcaliphilum* squalene synthase (*sqs*), squalene monooxygenase (*smo*), and oxidosqualene cyclase (*osc*) from plasmid pTrc-*sqs*-synRBS-*osc*-synRBS-*smo* (pABB501). Additional genes of interest were expressed from IPTG-inducible plasmid pSRKGm-*lac*UV5-rbs5 (pABB492) and/or arabinose-inducible plasmid pBAD1031K (pABB466). *E. coli* strains were cultured at 37 °C, shaken in 4–20 mL TB supplemented with antibiotics (as necessary) until midexponential phase when expression was induced with 0–500 μM IPTG and/or 0.2% arabinose (as necessary) for 30–40 h at 30 °C, and shaken at 225 rpm before harvest of cells.

Lipid Extraction and GC-MS Analysis. Lipids were extracted from cells harvested from 4–50 mL bacterial culture using a modified Bligh–Dyer extraction method (43, 44) as described in detail in ref. 10. Briefly, cells were sonicated in 10:5:4 (vol:vol:vol) methanol:dichloromethane (DCM):water; the organic phase was separated with twice the volume of 1:1 (vol:vol) DCM:water, and then, it was transferred and evaporated under N_2 to yield the total lipid extract (TLE). The alcohol soluble fractions of some TLEs were further purified by Si column chromatography (45). Lipids were derivatized to trimethylsilyl ethers with 1:1 (vol:vol) Bis(trimethylsilyl)trifluoroacetamide:pyridine before analysis by GC-MS. Lipid extracts were separated on an Agilent 7890B Series GC using one of two methods as described in *SI Appendix, Methods*.

NMR Analysis. The lipids were fractionated by preparative TLC on glass-backed plates (10 cm in length) coated with a 0.25-mm layer of silica gel 60 F254 using 4:1 hexane:ethyl acetate as the developing solvent. The TLC fractions were characterized by NMR using a Bruker Avance III HD with an Ascend 800-MHz magnet and a 5-mm TCI cryoprobe at 30 °C using deuterated chloroform ($CDCl_3$) as the solvent. Calibration was by the residual solvent signal (7.26 ppm). All NMR spectral data are shown in *SI Appendix*, Figs. S7–S12. Description of structure determination for each sterol product is provided in *SI Appendix*.

^{13}C-Labeling Study. Expression strain PVW7128 was cultured in 75 mL of TB supplemented with 25 mg of DOX and appropriate antibiotics with shaking at 37 °C. At midexponential phase, expression was induced with 500 μM IPTG and 0.2% arabinose and further incubated for 40 h with shaking at 30 °C. After harvesting, the cells were extracted as above, and the extracted lipids were separated by preparative TLC. The purified lanosterol and 4-desmethyllanosterol were analyzed by 800-MHz NMR using a quantitative heteronuclear single-quantum correlation method (46). The ^{13}C sweep width was narrowed to provide enhanced resolution. Care was taken to avoid

overlap from signal folding. NMR analysis was carried out in triplicate. The NMR integrals of the methyl signals were compared with those obtained from the corresponding unlabeled sterols, which were also measured in triplicate.

Bioinformatics Analysis. The JGI IMG (https://img.jgi.doe.gov/) Phylogenetic Profiler was used to identify putative bacterial sterol demethylating proteins. Homologs of SdmA and SdmB were identified through BLASTP searches (e value < $1e^{-50}$, 30% identity) of genomic databases on the JGI IMG Portal (47–49). Metagenomic homologs of SdmA were also identified through BLASTP searches (e value < $1e^{-50}$; 20% identity; >300 amino acids) of the JGI environmental and host-associated metagenomic databases. Protein sequences were aligned via MUSCLE (50) using Geneious (Biomatters Limited). Redundant metagenomic sequences were removed from alignments using the Decrease Redundancy web tool (https://web.expasy.org/decrease_redundancy/), and large gaps were removed from metagenomic sequence alignments using GBLOCKS (51). Maximum likelihood trees were constructed with PhyML (52) using the LG+gamma model, four gamma rate categories, 10 random starting trees, nearest-neighbor interchange branch swapping, and substitution parameters estimated from the data. All tree files were imported and edited in the Interactive Tree of Life online tool (itol.embl.de/) (53).

ACKNOWLEDGMENTS. We thank Prof. John E. Cronan for sharing the plasmid pBAD1031K, J. Keasling and T. S. Lee for plasmid pJBEI2997 (through Addgene), and Clare Gill and Yirang Cho for technical assistance. We also thank two anonymous reviewers for their insightful comments. Funding for this study was provided by National Science Foundation Grants OCE-1061957 (to J.-L.G.) and EAR-1451767 (to P.V.W.) and NASA Exobiology Program NNX16AJ33G (to P.V.W.). The acquisition of the 800-MHz NMR spectrometer was made possible by NIH Grant S10 OD012254.

1. Huang P, et al. (2016) Cellular cholesterol directly activates smoothened in hedgehog signaling. *Cell* 166:1176–1187.e14.
2. Miao L, et al. (2002) From lanosterol to cholesterol: Structural evolution and differential effects on lipid bilayers. *Biophys J* 82:1429–1444.
3. Bi X, Liao G (2010) Cholesterol in Niemann-Pick type C disease. *Subcell Biochem* 51: 319–335.
4. Summons RE, Bradley AS, Jahnke LL, Waldbauer JR (2006) Steroids, triterpenoids and molecular oxygen. *Philos Trans R Soc Lond B Biol Sci* 361:951–968.
5. Brocks JJ, et al. (2017) The rise of algae in Cryogenian oceans and the emergence of animals. *Nature* 548:578–581.
6. Love GD, et al. (2009) Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* 457:718–721.
7. Bird CW, Lynch JM, Pirt FJ, Reid WW (1971) Steroids and squalene in *Methylococcus capsulatus* grown on methane. *Nature* 230:473–474.
8. Bode HB, et al. (2003) Steroid biosynthesis in prokaryotes: Identification of myxobacterial steroids and cloning of the first bacterial 2,3(S)-oxidosqualene cyclase from the myxobacterium *Stigmatella aurantiaca*. *Mol Microbiol* 47:471–481.
9. Wei JH, Yin X, Welander PV (2016) Sterol synthesis in diverse bacteria. *Front Microbiol* 7:990.
10. Banta AB, Wei JH, Gill CC, Giner JL, Welander PV (2017) Synthesis of arborane triterpenols by a bacterial oxidosqualene cyclase. *Proc Natl Acad Sci USA* 114:245–250.
11. Kohl W, Gloe A, Reichenbach H (1983) Steroids from the myxobacterium *Nannocystis exdens*. *J Gen Microbiol* 129:1629–1635.
12. Garcia R, Gemperlein K, Müller R (2014) *Minicystis rosea* gen. nov., sp. nov., a polyunsaturated fatty acid-rich and steroid-producing soil myxobacterium. *Int J Syst Evol Microbiol* 64:3733–3742.
13. Pearson A, Budin M, Brocks JJ (2003) Phylogenetic and biochemical evidence for sterol synthesis in the bacterium Gemmata obscuriglobus. *Proc Natl Acad Sci USA* 100: 15352–15357.
14. Patt TE, Hanson RS (1978) Intracytoplasmic membrane, phospholipid, and sterol content of *Methylobacterium organophilum* cells grown under different conditions. *J Bacteriol* 134:636–644.
15. Schouten S, Bowman JP, Rijpstra WI, Sinninghe Damsté JS (2000) Sterols in a psychrophilic methanotroph, *Methylosphaera hansonii*. *FEMS Microbiol Lett* 186: 193–195.
16. Villanueva L, Rijpstra WI, Schouten S, Damsté JS (2014) Genetic biomarkers of the sterol-biosynthetic pathway in microalgae. *Environ Microbiol Rep* 6:35–44.
17. Gold DA, Caron A, Fournier GP, Summons RE (2017) Paleoproterozoic sterol biosynthesis and the rise of oxygen. *Nature* 543:420–423.
18. Desmond E, Gribaldo S (2009) Phylogenomics of sterol synthesis: Insights into the origin, evolution, and diversity of a key eukaryotic feature. *Genome Biol Evol* 1: 364–381.
19. Frickey T, Kannenberg E (2009) Phylogenetic analysis of the triterpene cyclase protein family in prokaryotes and eukaryotes suggests bidirectional lateral gene transfer. *Environ Microbiol* 11:1224–1241.
20. Bard M, et al. (1996) Cloning and characterization of ERG25, the *Saccharomyces cerevisiae* gene encoding C-4 sterol methyl oxidase. *Proc Natl Acad Sci USA* 93: 186–190.
21. Gachotte D, Barbuch R, Gaylor J, Nickel E, Bard M (1998) Characterization of the *Saccharomyces cerevisiae* ERG26 gene encoding the C-3 sterol dehydrogenase (C-4 decarboxylase) involved in sterol biosynthesis. *Proc Natl Acad Sci USA* 95: 13794–13799.
22. Gachotte D, et al. (1999) Characterization of the Saccharomyces cerevisiae ERG27 gene encoding the 3-keto reductase involved in C-4 sterol demethylation. *Proc Natl Acad Sci USA* 96:12655–12660.
23. Rahier A (2011) Dissecting the sterol C-4 demethylation process in higher plants. From structures and genes to catalytic mechanism. *Steroids* 76:340–352.
24. Jones P, et al. (2014) InterProScan 5: Genome-scale protein function classification. *Bioinformatics* 30:1236–1240.
25. Mitropoulos KA, Gibbons GF, Reeves BE (1976) Lanosterol 14alpha-demethylase. Similarity of the enzyme system from yeast and rat liver. *Steroids* 27:821–829.
26. Penning TM, Chen M, Jin Y (2015) Promiscuity and diversity in 3-ketosteroid reductases. *J Steroid Biochem Mol Biol* 151:93–101.
27. Sharpless KB, Snyder TE, Spencer TA, Maheshwari KK, Nelson JA (1969) Biological demethylation of 4,4-dimethyl sterols, evidence for enzymic epimerization of the 4β-methyl group prior to its oxidative removal. *J Am Chem Soc* 91:3394–3396.
28. Giner JL, Jaun B (1998) Biosynthesis of isoprenoids in *Escherichia coli*: Retention of the methyl H-atoms of 1-deoxy-D-xylulose. *Tetrahedron Lett* 39:8021–8022.
29. Eisenreich W, Bacher A, Arigoni D, Rohdich F (2004) Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell Mol Life Sci* 61:1401–1426.
30. Rohmer M (2008) From molecular fossils of bacterial hopanoids to the formation of isoprene units: Discovery and elucidation of the methylerythritol phosphate pathway. *Lipids* 43:1095–1107.
31. Kallberg Y, Persson B (2006) Prediction of coenzyme specificity in dehydrogenases/reductases. A hidden Markov model-based method and its application on complete genomes. *FEBS J* 273:1177–1184.
32. Bouvier P, Rohmer M, Benveniste P, Ourisson G (1976) Delta8(14)-steroids in the bacterium *Methylococcus capsulatus*. *Biochem J* 159:267–271.
33. Xu F, et al. (2005) Dual roles for cholesterol in mammalian cells. *Proc Natl Acad Sci USA* 102:14551–14556.
34. Zhang X, et al. (2016) Sterol methyl oxidases affect embryo development via auxin-associated mechanisms. *Plant Physiol* 171:468–482.
35. Bloch K (1991) Cholesterol: Evolution of structure and function. *Biochemistry of Lipids, Lipoproteins, and Membranes*, eds Vance DE, Vance J (Elsevier, Amsterdam), pp 363–381.
36. Jahnke LL, Stan-Lotter H, Kato K, Hochstein LI (1992) Presence of methyl sterol and bacteriohopanepolyol in an outer-membrane preparation from *Methylococcus capsulatus* (Bath). *J Gen Microbiol* 138:1759–1766.
37. Hughes AL, Lee CY, Bien CM, Espenshade PJ (2007) 4-Methyl sterols regulate fission yeast SREBP-Scap under low oxygen and cell stress. *J Biol Chem* 282:24388–24396.
38. Hannich JT, et al. (2009) Methylation of the sterol nucleus by STRM-1 regulates dauer larva formation in *Caenorhabditis elegans*. *Dev Cell* 16:833–843.
39. Jahnke LL, Nichols PD (1986) Methyl sterol and cyclopropane fatty acid composition of *Methylococcus capsulatus* grown at low oxygen tensions. *J Bacteriol* 167:238–242.
40. Brocks JJ, et al. (2005) Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. *Nature* 437:866–870.
41. Whittenbury R, Phillips KC, Wilkinson JF (1970) Enrichment, isolation and some properties of methane-utilizing bacteria. *J Gen Microbiol* 61:205–218.
42. Peralta-Yahya PP, et al. (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nat Commun* 2:483.
43. Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 37:911–917.
44. Welander PV, et al. (2012) Identification and characterization of *Rhodopseudomonas palustris* TIE-1 hopanoid biosynthesis mutants. *Geobiology* 10:163–177.
45. Summons RE, et al. (2013) Lipid biomarkers in ooids from different locations and ages: Evidence for a common bacterial flora. *Geobiology* 11:420–436.
46. Yan J, Kiemle D, Liu S (2015) Quantification of xylooligomers in hot water wood extract by ^{1}H-^{13}C heteronuclear single quantum coherence NMR. *Carbohydr Polym* 117:903–909.
47. Chen IM, et al. (2016) Supporting community annotation and user collaboration in the integrated microbial genomes (IMG) system. *BMC Genomics* 17:307.
48. Markowitz VM, et al. (2014) IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res* 42:D560–D567.
49. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
50. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
51. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552.
52. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321.
53. Letunic I, Bork P (2016) 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.