Main Manuscript for

Sterol side chain alkylation by sponge and bacterial methyltransferases

Malory O. Brown[a] and Babatunde O. Olagunju[b], José-Luis Giner[b], and Paula V. Welander[a]

[a]Department of Earth System Science, Stanford University, Stanford, CA 94305

[b]Department of Chemistry, State University of New York-Environmental Science and Forestry, Syracuse, NY 13210

*Paula V. Welander

Email: welander@stanford.edu

Author Contributions: M.O.B, J.-L.G., and P.V.W. designed research; M.O.B., B.O.O., and J.-L.G. performed research; B.O.O. and J.-L.G. contributed reagents; M.O.B., B.O.O., J.-L.G., and P.V.W. analyzed data; and M.O.B, B.O.O., J.-L.G., and P.V.W. wrote the paper.

Competing Interest Statement: The authors declare no conflict of interest.

Classification: PHYSICAL SCIENCES: Earth, Atmospheric, and Planetary Sciences; BIOLOGICAL SCIENCES: Microbiology

Keywords: sterols, steranes, biomarkers, sterol methyltransferase, 24-isopropylcholestane

This PDF file includes:

Main Text
Figures 1 to 3
Table 1

Abstract

Fossilized sterols, or steranes, are broadly interpreted as biomarkers for past eukaryotic life. Steranes with side chain methylations can act as more specific eukaryotic biomarkers given the limited taxonomic distribution of their sterol precursors in extant eukaryotes. One side chain methylated sterane, 24-isopropylcholestane, has been interpreted as a biomarker for marine sponges of the Demospongiae class as they are the only organisms known to produce 24-isopropyl sterols as their major sterols. The presence of 24-isopropylcholestane in late Neoproterozoic rocks may therefore represent some of the earliest evidence for animals on Earth. However, the use of 24-isopropylcholestane as a sponge biomarker has been questioned for a variety of reasons including the potential for alternative sources and a lack of information on the biochemical mechanism behind the formation of its sterol precursors. Specifically, enzymes capable of methylating sterols to give the 24-isopropyl side chain have not been identified, and the function of putative sponge sterol 24-C-methyltransferases has not been demonstrated experimentally. In this study, we verified the functionality of sponge sterol methyltransferases and explored the potential for bacterial sterol side chain methylation through in vitro protein expression and lipid analyses. We identified three bacterial sterol methyltransferases each capable of sequential methylations resulting in the 24-isopropyl side chain structure of the 24-isopropylcholestane biomarker. Two of these side chain propylating enzymes were identified in a metagenome from a demosponge microbiome, suggesting bacterial symbionts may play a role in 24-isopropyl sterol biosynthesis in the demosponge holobiont.

Significance Statement

According to the sponge biomarker hypothesis, geologic 24-isopropylcholestane can act as an indicator for ancient demosponges and may represent some of the first evidence for animal life on Earth. However, enzymes capable of methylating sterols to give the 24-isopropyl sterol side chain have not been identified. We investigated the functionality of sponge sterol methyltransferases, enzymes known to alkylate the sterol side chain, through in vitro protein expression. We also analyzed the function of several putative bacterial sterol methyltransferases,

three of which produced 24-isopropyl sterols. Our results indicate uncultured bacteria, including demosponge symbionts, have the genomic capacity to synthesize side chain alkylated sterols and should therefore be considered when interpreting ergostane, stigmastane, and 24-isopropylcholestane biomarkers in the geologic record.

Main Text

Introduction

Sterols are polycyclic triterpenoid lipids required by most eukaryotes for many critical cellular functions, such as maintaining membrane integrity, phagocytosis, stress tolerance, embryogenesis, and cellular signaling (1–6). In addition, sterols are quite recalcitrant and are preserved as sterane hydrocarbons in sedimentary rocks up to 1.6 billion years old (7). Given their persistence in the geologic record and wide distribution amongst extant eukaryotes, steranes are broadly interpreted as biomarkers for ancient eukaryotic life (8, 9). But particular modifications to the sterol side chain, such as alkylation at the C-24 and C-26 positions, remain unique to certain eukaryotic lineages and can therefore function as more specific biomarkers (10). Two side chain alkylated steranes, 24-isopropylcholestane (24-ipc) and 26-methylstigmastane (26-mes), have been used as biomarkers for demosponges (11, 12) as they are the only extant organisms known to contain sterols with the 24-ipc and 26-mes side chain structures as their major sterols (12–18). The co-occurrence of 24-ipc and 26-mes in certain late Neoproterozoic rocks (660-540 Ma) predates the earliest sponge macrofossils by ~100 million years (19), potentially representing the first evidence of animals in the geologic record.

However, the use of 24-ipc and 26-mes as biomarkers for demosponges has been controversial and continues to be debated for several reasons (19–26). For example, 24-isopropyl sterols have been identified in pelagophyte algae which produce the structurally related compound 24-*n*-propylcholesterol as their major sterol (27–30). Trace amounts of 24-ipc and 26-mes have also been reported in hydrogenated lipid extracts from Rhizaria (31), although attempts to replicate these results have not been successful (23). Subsequent studies demonstrated 24-ipc and 26-mes

53 form abiotically from certain C_{29} sterols during laboratory pyrolysis under conditions that may mimic
54 diagenesis (26, 32). These studies suggest that the previously observed Rhizarian 24-isopropyl
55 and 26-methyl steranes may have been generated through laboratory pyrolysis rather than by in
56 vivo synthesis and that Neoproterozoic 24-ipc and 26-mes may have formed abiotically from algal
57 sterols. Further, microbial symbionts can constitute up to 60% of sponge biomass (33, 34) and
58 diverse bacteria are known to produce sterols (35). Bacterial symbionts may therefore contribute to
59 24-isopropyl and 26-methyl sterol biosynthesis in the sponge holobiont – defined as the sponge
60 host and the various microbes and viruses that reside within it (36). However, bacteria with the
61 genomic capacity to alkylate the sterol side chain have not previously been identified.

62 One approach to constrain these uncertainties in the sponge biomarker hypothesis is to
63 better understand the biochemical mechanism underpinning these distinct alkylations at C-24 and
64 C-26. Sterol 24-C-methyltransferases (SMTs), which alkylate the C-24 position of the sterol side
65 chain at an isolated double bond through an *S*-adenosylmethionine (SAM) dependent mechanism
66 (37), are the only enzymes known to methylate the sterol side chain. These proteins have been
67 characterized extensively from fungal and plant sources but only recently has an attempt been
68 made to address this biochemical reaction in sponges. Phylogenetic analyses performed to
69 understand how sponges could be synthesizing 24-propyl sterols suggested that the number of
70 SMT copies in a eukaryotic genome corresponds to the number of carbons added at C-24 in all
71 analyzed lineages except sponges, which contain one less SMT in their genome or transcriptome
72 than expected based on their sterol profile (38). This prompted the hypotheses that at least two
73 SMT copies in a genome are required to produce 24-propyl sterols and that sponge SMTs are
74 capable of multiple methylations at C-24. However, no experimental evidence demonstrating this
75 activity by any sponge SMT currently exists, and SMTs capable of alkylating the sterol side chain
76 to produce 24-propyl sterols have not been identified in any extant organism.

77 Here, we characterize a variety of putative SMTs from sponge and bacterial sources to
78 better understand the biochemical requirements for generating propylated sterols. We demonstrate
79 that sponge sterol 24-C-methyltransferase homologs are bona fide SMTs that can perform multiple
80 rounds of alkylation to produce sterols methylated and ethylated at the C-24 position. In addition,

81 we show that putative bacterial SMTs identified in metagenomic datasets, including those from
82 demosponge microbial symbiont communities, are capable of alkylating at the C-24 position.
83 Surprisingly, 3 SMTs of bacterial origin can each perform all 3 of the methylation steps required to
84 produce the 24-isopropyl side chain. This is the first in vitro demonstration of propylation of a sterol
85 side chain by an SMT and the first demonstration that bacteria have the potential to alkylate sterols
86 at C-24. Taken together, these data suggest alternative sources of 24-ipc are plausible and
87 demonstrate the utility of understanding the biochemistry of lipid biosynthesis to allow for more
88 robust biomarker interpretations.

89 **Results**

90 **Identification of functional sponge sterol methyltransferases.** Previous studies
91 identified SMT homologs in a variety of sponge genomes and transcriptomes (38). However, there
92 is currently no experimental evidence demonstrating that these putative SMTs alkylate the sterol
93 side chain much less methylate multiple times to generate 24-propyl sterols. Our first goal was to
94 heterologously express these putative SMTs from sponge sources and show methylation activity.
95 We tested 8 SMT homologs identified in the genomes and transcriptomes of 8 sponges of the
96 classes Demospongiae, Homoscleromorpha, and Calcarea which have each been shown to
97 contain 24-ethyl sterols at the species (39–42) or genus (43, 44) level. SMTs were codon-optimized
98 and artificially synthesized for expression in *Escherichia coli*. Cell extracts were generated from *E.*
99 *coli* overexpressing each sponge SMT and incubated with two potential SMT substrates,
100 desmosterol or 24-methylenecholesterol, and resulting sterol products were analyzed using GC-
101 MS. All 8 of the analyzed sponge SMTs were functional in our system and alkylated desmosterol
102 to produce 24-methyl sterols including 24-methylenecholesterol and codisterol (Table 1, Fig. 1*B*).
103 Six of the sponge SMTs also alkylated 24-methylenecholesterol to 24-ethyl sterols including
104 fucosterol, isofucosterol, and (epi)clerosterol demonstrating that some sponge SMTs can perform

105 two sequential side chain methylations. None of the analyzed sponge SMTs produced 24-propyl
106 sterols in our expression system.

107 **Bacterial SMTs are functional and produce 24-isopropyl sterols**. Although we
108 demonstrated sponge SMT homologs are true sterol methyltransferases, none of the proteins we
109 tested produced 24-isopropyl sterols–the diagenetic precursors to the 24-ipc biomarker. Our
110 analyses of sponge microbiomes, however, revealed several potential sterol biosynthesis proteins,
111 including SMT homologs, from bacterial symbionts. We hypothesized that these bacterial SMTs
112 may contribute to side chain alkylation in the sponge holobiont. However, no putative bacterial
113 SMTs, from either free-living or host-associated sources, have been experimentally shown to
114 alkylate the sterol side chain.

115 We chose 10 putative bacterial SMTs identified in the microbiome of 4 demosponge
116 species for in vitro expression and lipid analyses as previously performed with the sponge SMTs.
117 Of these 10 symbiont SMTs, 8 methylated desmosterol to 24-methyl sterols including codisterol,
118 24-methylenecholesterol, and 24-methyldesmosterol (Table 1, Fig. 1). Of the 8 SMTs that
119 methylated desmosterol, 3 also methylated 24-methylenecholesterol to 24-ethyl sterols including
120 fucosterol, isofucosterol, and (epi)clerosterol. Two SMTs, both identified in a metagenome from the
121 demosponge *Aplysina aerophoba*, also methylated 24-methylenecholesterol to two distinct C_{30}
122 sterols. Nuclear magnetic resonance (NMR) analysis (see Methods) confirmed the first symbiont
123 SMT produced 24-isopropenylcholesterol while the second produced 24-isopropylidenecholesterol
124 (Table S1). Both *A. aerophoba* metagenome SMTs were capable of all three methylation steps
125 required to synthesize the 24-ipc side chain. One propylating SMT was identified in a gene cluster
126 with several other homologs of sterol biosynthesis genes, including squalene monooxygenase,
127 oxidosqualene cyclase, sterol 14α-demethylase, delta(14)-sterol reductase, and the bacterial C-4
128 demethylation genes *sdmA* and *sdmB*, suggesting an unknown bacterial sponge symbiont has the
129 genomic capacity to synthesize 24-isopropyl sterols de novo (Fig. 2).

130 Identification of bacterial symbiont SMTs capable of producing side chain alkylated sterols,
131 including 24-isopropyl sterols, led us to question if bacteria from environments outside a sponge
132 host have the potential to alkylate the sterol side chain. Through additional searches, we identified

133 potential bacterial SMTs from a variety of publicly available environmental metagenomes and
134 metagenome-assembled genomes (MAGs). We tested the function of 14 of these SMT homologs
135 from marine, freshwater, thermal spring, and coral reef metagenomes and MAGs. Twelve of these
136 SMTs alkylated the sterol side chain to produce 24-methyl and/or 24-ethyl sterols (Table 1). One
137 SMT, identified in a drinking water filter MAG assigned to the phylum Chlamydiae, was capable of
138 alkylating desmosterol to both 24-isopropenylcholesterol and 24-isopropylidenecholesterol (Fig.
139 1*A*).

140 **Structural confirmation of C_{30} sterols produced by propylating bacterial SMTs.** It was
141 surprising to observe that one SMT was capable of propylating the sterol side chain on its own as
142 it has been hypothesized that at least two SMTs would be required for this reaction (38). To better
143 understand this unexpected biosynthetic pathway to 24-isopropyl sterols, we performed additional
144 in vitro reactions with one of the *A. aerophoba* metagenome SMTs and the Chlamydiae MAG SMT
145 with labeled substrates including [28-^{13}C] 24-methylenecholesterol, [29-^{13}C] isofucosterol, and [29-
146 ^{13}C] fucosterol. These ^{13}C-labeling experiments were used to assist the identification of the
147 biosynthetic products by 2D-NMR (heteronuclear single quantum coherence (HSQC) and
148 heteronuclear multiple bond correlation (HMBC)) by increasing the amount of ^{13}C at a specific
149 position by two orders of magnitude compared to natural abundance. The two terminal isopropyl
150 groups of 24-isopropyl sterols come from biosynthetically disparate origins, the isoprenoid pathway
151 and SAM methylation (Fig. S1A). Because of the possibilities of hydride-shifts and different
152 positions of deprotonation, each of the four terminal methyl groups can have multiple different
153 origins within the product (17). In addition to enhanced sensitivity, ^{13}C-labeling elucidated
154 mechanistic details of a pathway which, because of the intrinsic symmetry of 24-isopropyl sterols,
155 would be otherwise hidden.

156 In experiments using [28-^{13}C] 24-methylenecholesterol with the *A. aerophoba* symbiont
157 SMT, isofucosterol was found to be the major C_{29} sterol as shown by a ^{1}H-^{13}C cross peak
158 corresponding to C-28 at 5.108 and 116.48 ppm (Fig. S1B). The presence of fucosterol was shown
159 by an HSQC cross peak at 5.183 and 115.58 ppm. Comparison of the integral ratios of the cross
160 peaks (45) for these two isomers showed them to be present in a 7:1 ratio. An additional minor C_{29}

161 sterol was detected and shown to be epiclerosterol but not clerosterol by cross peaks for C-28 at
162 1.290/1.380 and 26.01 ppm (Table S1). Long-range HMBC correlations from the 29-methyl group
163 confirmed these identifications. Thus, isofucosterol showed an HMBC cross peak for H-29 and C-
164 28 at 1.59 and 115.58 ppm; fucosterol and epiclerosterol showed HMBC cross peaks at 1.57 and
165 116.48 ppm and 0.799 and 26.01 ppm, respectively. The major product of triple side chain
166 methylation was *24S*-24-isopropylcholesta-5,25-dienol (24-isopropenylcholesterol). The label was
167 not detected at C-28 by HSQC but was found exclusively at C-25 by HMBC from the 27-methyl
168 group (cross peak between 1.570 and 147.38 ppm). Moreover, its *24R* isomer was absent as
169 shown by the lack of HSQC and HMBC signals.

170 Further characterization of the *A. aerophoba* symbiont methylation pathway was carried
171 out with [29-^{13}C] fucosterol and [29-^{13}C] isofucosterol. Fucosterol was not a substrate for the SMT,
172 but isofucosterol yielded *24S*-24-isopropylcholesta-5,25-dienol (Fig. S3C) with the label distributed
173 in a 4:1 label ratio between C-27 and C-26, respectively, as shown by integration of the HSQC
174 signals for C-27 (1.570 and 18.95 ppm), and C-30 (4.602/4.740, and 111.85). This outcome shows
175 that interconversion of the C-28 and C-25 cations by hydride-shifts is not involved in the
176 biosynthesis of this compound and that deprotonation of the initially formed cation is not specific.

177 The same set of reactions with the propylating Chlamydiae MAG SMT gave similar results
178 - fucosterol was not a substrate and epiclerosterol, but not clerosterol, was detected. However,
179 minor differences were measured between the two bacterial SMTs tested. In the Chlamydiae MAG
180 SMT reaction, the ratio of isofucosterol and fucosterol was 30:1, and the ratio of label at C-27/C-
181 26 of *24S*-24-isopropylcholesta-5,25-dienol was 5:1. Also, the C_{30} Δ^{23} and Δ^{24} isomers were
182 detected in the Chlamydiae MAG SMT. In the experiment with [28-^{13}C] 24-methylenecholesterol,
183 HMBC showed two adjacent cross peaks correlating the C-26 and C-27 methyls with C-25 of 24-
184 isopropylcholesta-5,24-dienol (24-isopropylidenecholesterol, the Δ^{24} isomer) at 1.633/1.651 and
185 123.1 ppm. In this case, the label from [29-^{13}C] isofucosterol was detected only on the Z-vinylic
186 methyl by HSQC (1.651 and 19.73 ppm). 24-Isopropylcholesta-5,23-dienol (the Δ^{23} isomer) was
187 detected by label at 20.96 ppm exclusively, which is one of the two methyls of the Z-isopropyl
188 group. These results provide evidence for one hydride-shift to generate the C-24 cation. The ratio

189 of *24S*-24-isopropylcholesta-5,25-dienol and its Δ^{24} and Δ^{23} isomers was estimated by integration
190 of the HSQC to be 20:4:1.

191 **Phylogenetic analyses of bacterial SMTs.** The identification of functional SMTs in
192 metagenomes and MAGs assigned to the bacterial domain, including the three SMTs that propylate
193 the sterol side chain, is quite significant given that no bacterial proteins have ever been confirmed
194 to alkylate the sterol side chain. To provide further evidence that these SMTs are truly of bacterial
195 origin, we first performed phylogenetic analyses of 10 sponge SMTs and the 24 metagenomic
196 SMTs tested in our in vitro system. We added these SMTs and additional eukaryotic SMT homologs
197 identified in GenBank to a previously published SMT dataset (38) and repeated their analysis to
198 generate a maximum-likelihood phylogenetic tree from a concatenated alignment of conserved
199 methyltransferase and C-terminal domains as identified in Pfam. Through this analysis, we see that
200 23 out of the 24 putative bacterial SMTs, including those from sponge symbionts, form two clades
201 separate from the sponge SMT clade suggesting bacterial symbionts did not acquire their SMTs
202 from their sponge host via horizontal gene transfer. While the observed bootstrap values are not
203 robust enough to draw any meaningful evolutionary conclusions, the observed clustering suggests
204 that the majority of the metagenomic SMTs we tested are distinct and more closely related to each
205 other than to the SMTs of eukaryotes known to produce side chain alkylated sterols (Fig. 3).

206 The genomic context of the putative bacterial SMTs we tested can also provide evidence
207 of their bacterial origins. Of the 24 metagenomic SMTs tested here, 9 were in gene clusters or
208 MAGs with a homolog of oxidosqualene cyclase (OSC), a key enzyme in sterol biosynthesis
209 responsible for cyclizing oxidosqualene to lanosterol (35). Previous work has demonstrated that
210 OSCs cluster into robust phylogenetic clades with high bootstrap support (46). In addition,
211 phylogenetic analyses of environmental metagenomic OSCs from unknown sources can provide
212 insight into the types of uncultured organisms that may be producing sterols (35, 47). Thus,
213 phylogenetic analyses of OSCs associated with the SMTs we tested could provide more context of
214 the bacterial origins of these proteins. To generate the OSC tree, we identified 485 OSC sequences
215 from bacterial and eukaryotic genomes and MAGs using the OSC from the γ-proteobacterium
216 *Methylococcus capsulatus* BATH as a search query against the JGI IMG database. These OSC

217 protein sequences were aligned and used to generate a maximum-likelihood phylogenetic tree
218 using the squalene-hopene cyclase (SHC) protein as an outgroup. Seven of the OSCs, including 4
219 from demosponge metagenomes, fell into the Group 1 clade which contains only bacterial OSCs
220 (Fig. S2). The remaining 2 OSCs, from soil and hot spring sediment metagenomes, fell into
221 bacterial Group 2. The position of these metagenomic OSCs, which are associated with SMTs we
222 tested, provides additional evidence that these SMTs originate from a bacterial source.

223 To further demonstrate these SMTs originate from a bacterial source, we used the recently
224 developed classifier Whokaryote (48), which predicts whether a metagenomic contig is likely
225 eukaryotic or prokaryotic based on intergenic distance, gene density, gene length, and k-mer
226 frequencies. Whokaryote predicts all 19 of the contigs containing the metagenomic and MAG SMTs
227 tested here to be of prokaryotic origin (Table S4).

228 **Discussion**

229 The sponge biomarker hypothesis proposes that demosponges are the predominant
230 source of 24-isopropylcholestane biomarkers preserved in the sedimentary rock record. Support
231 for this hypothesis was put forth through phylogenetic and molecular clock analyses of putative
232 sponge sterol methyltransferases (SMTs) thought to catalyze the formation of propylated sterols
233 (38). These studies concluded that sponge SMTs are not only bona fide sterol methyltransferases
234 but that they are capable of multiple rounds of methylation. However, no experimental evidence
235 has been provided to confirm that sponge-derived SMTs methylate at the C-24 position much less
236 methylate multiple times.

237 The results presented here experimentally demonstrate that sponge genomes and
238 transcriptomes do indeed contain functional SMTs capable of sequentially methylating desmosterol
239 to 24-methylenecholesterol to 24-ethyl sterols, similarly to recently reported SMTs found in annelid
240 worms (49). However, as with annelid SMTs (50), we did not observe this activity across all sponge
241 SMTs analyzed in this study. Sterol profiles from *Oscarella carmela* (Homoscleromorpha) and
242 *Sycon ciliatum* (Calcarea) have not been reported, but 24-ethyl sterols have been identified in other
243 members of *Oscarella* and *Sycon* (43, 51). Yet, the *O. carmela* and *S. ciliatum* SMTs tested here

244 did not produce 24-ethyl sterols. SMTs capable of multiple rounds of methylation may therefore be
245 restricted to certain sponge species, although we recognize that these SMTs may be more selective
246 and require a different substrate than we have provided in our assays to produce 24-ethyl sterols.
247 Further, we did not observe the production of 24-propyl sterols by any sponge SMT, although this
248 was not surprising given that 24-propyl sterols have not been identified in any of the sponge species
249 analyzed here. One of the difficulties in trying to ascertain the potential for propylation by sponge-
250 derived SMTs is a lack of genomic data from demosponges that have been confirmed to produce
251 24-ipc. Future sequencing of genomes or transcriptomes from demosponges known to contain
252 sterols with the 24-ipc carbon skeleton may reveal sponge-derived SMTs sufficient for the
253 biosynthesis of 24-propyl sterols.

254 While we were unable to demonstrate the production of propylated sterols through the
255 expression of sponge SMTs, we did show that bacterial symbionts have the genomic capacity to
256 synthesize these lipids and may therefore contribute to the 24-isopropyl sterol pool in the
257 demosponge holobiont. Studies performed prior to the development of 16S community sequencing
258 attempted to demonstrate de novo sterol biosynthesis in demosponges. However, these
259 experiments were performed using cell-free sponge extracts or on whole sponge tissue from which
260 the presence of symbiont-derived enzymes cannot be excluded (52–54). Without confirmation that
261 the bacterial symbionts in these samples do not contain functional SMTs, the possibility remains
262 that bacterial symbionts contribute to sterol biosynthesis in demosponges. Further, holo-
263 transcriptomic sequencing of the demosponge *Vaceletia* sp. found that several key sterol
264 biosynthesis transcripts, including those for squalene synthase, squalene monooxygenase, and
265 oxidosqualene cyclase, were bacterially-derived (55). Both sponge and bacterial symbiont proteins
266 may therefore be required to produce 24-isopropyl sterols in vivo. However, the metagenomic gene
267 clusters containing functional SMTs and several other key genes for sterol biosynthesis identified
268 here, as well as steroid gene clusters containing SMT homologs previously identified in bacterial
269 MAGs associated with demosponges of the *Ircinia* genus (56), suggest bacterial demosponge
270 symbionts may produce sterols de novo without demosponge input. The 2 symbiont SMTs sufficient
271 for side chain propylation were both identified in metagenomes from the demosponge *A.*

272 *aerophoba*, which appears to contain only trace amounts of 24-isopropyl sterols (18). Our work
273 highlights the need for robust sequencing of both the sponge host and its microbiome coupled to
274 sterol analysis in demosponge species that produce 24-isopropyls as their major sterols.

275 From a biomarker perspective, our demonstration that a single bacterial SMT is sufficient
276 for propylating a sterol side chain is quite significant. Previous molecular clock analyses suggested
277 demosponges were the most likely source of Cryogenian 24-ipc biomarkers given that
278 demosponges acquired a second SMT copy prior to their deposition, while pelagophyte algae did
279 not acquire a third SMT copy until the Phanerozoic (38). Further, researchers have argued against
280 Rhizaria as a source of 24-isopropyl sterols given that *Bigelowiella natans* contains only 2 SMT
281 copies in its genome and is thus unlikely to synthesize 24-propyl sterols (23). These interpretations
282 rely on the hypothesis that multiple SMT copies are required to produce 24-isopropyl sterols.
283 However, we identified 3 SMTs each capable of producing the 24-isopropyl side chain via a single
284 protein. SMT copy number may therefore be an unreliable predictor of an organism's ability to
285 produce 24-propyl sterols. Now that propylating SMTs have been identified, site-directed
286 mutagenesis studies like those performed with yeast (57–59), algal (60), and plant (61) SMTs may
287 reveal specific amino acid changes that allow an SMT protein to perform a third methylation. If such
288 key residues are identified, they could act as a more reliable predictor of 24-isopropyl sterol
289 biosynthesis directly from sequencing data than SMT copy number alone.

290 In addition to its use as a demosponge biomarker, 24-ipc has also been used to indicate
291 marine input into terrestrial systems (62–64) given that 24-ipc precursor sterols have not been
292 identified in freshwater demosponges or algae. However, the production of 24-isopropyl sterols by
293 an SMT identified in a freshwater Chlamydiae MAG suggests the occurrence of these sterols may
294 not be limited to marine environments. Members of the phylum Chlamydiae have not been shown
295 to produce sterols, and homologs of other sterol biosynthesis genes were not present in the
296 Chlamydiae MAG. It is therefore unlikely that this organism synthesizes 24-isopropyl sterols de
297 novo. However, some endosymbiotic Chlamydiae are known to recruit sterols from their eukaryotic
298 host into their membranes (65). The possibility therefore remains that certain freshwater
299 Chlamydiae alkylate the side chain of host-acquired sterols to give the 24-isopropyl side chain via

300 their SMTs for reasons yet unknown. If this is the case, it may explain the presence of 22-dehydro-
301 24-isopropylcholestanol throughout a freshwater sediment core from Motte Lake, France (66). We
302 would therefore exercise caution when, for example, interpreting the presence of 24-ipc as
303 evidence for seawater incursion events.

304 Finally, although we identified SMTs capable of 24-isopropyl sterol biosynthesis, proteins
305 capable of methylating sterols to give other geologically important side chain structures remain
306 undiscovered. Identification of enzymes capable of methylating sterols at the C-26 position to give
307 the 26-mes side chain structure remains especially important when considering the biomarker
308 evidence for Precambrian sponges. Studies in which radiolabeled sterol precursors were fed to
309 whole sponges or sponge extracts have demonstrated codisterol or epicodisterol are the preferred
310 substrates for methylation at C-26 via a hypothetical 26-SMT protein (52, 67, 68). We identified
311 (epi)codisterol as a product of several SMTs, from both sponges and bacteria, suggesting both
312 groups have the genomic capacity to produce these potential 26-SMT substrates. However, it
313 remains to be seen if sponges and/or bacteria, sponge-associated and otherwise, encode SMTs
314 capable of methylation at C-26. Once enzymes with this functionality are identified, in-depth kinetics
315 studies with purified SMT proteins and multiple sterol substrates may reveal if SMTs from sponges
316 and/or bacteria are likely to produce 24-ipc and 26-mes sterol precursors in ratios that correspond
317 to the C_{30} sterane distributions observed in the Cryogenian. Similarly, confirmed side chain
318 alkylated sterol biosynthesis in a bacterial isolate along with identification of any unique isotopic
319 signatures resulting from the alkylation of sterol side chains by bacteria could help determine if
320 bacteria may have been a potential source of Cryogenian 24-ipc. Until then, the co-occurrence of
321 24-ipc and 26-mes in high ratios relative to other steranes continues to provide strong evidence for
322 derived sponges in the Precambrian, consistent with molecular clock analyses (69–71). However,
323 sterol side chain alkylation by symbiotic and free-living bacteria should now be considered both
324 when analyzing the sterols of extant eukaryotes and when interpreting the presence of geologic
325 steranes including ergostane, stigmastane, and 24-isopropylcholestane.

326

327 **Materials and Methods**

328 **Bioinformatics Analysis.** SMT homologs were identified in the Joint Genome Institute
329 Integrated Microbial Genomes & Microbiomes (JGI IMG), GenBank, and Compagen
330 (http://www.compagen.org) databases using the *Saccharomyces cerevisiae* S299C SMT amino
331 acid sequence as the BLASTP search query (locus tag: YML008C, maximum e-value: $1e^{-50}$,
332 minimum percent identity: 30%). Additional demosponge SMT sequences from *C. nucula*, *S.*
333 *fasciculatas*, and *P. ficiformis* transcriptomes were provided by Prof. David A. Gold (38). Identifying
334 sequence information is available in Table S2. Conserved sterol methyltransferase C-terminal and
335 methyltransferase 11 domains were identified using the HmmerWeb tool (version 2.41.1).

336 Oxidosqualene cyclase homologs were identified through BLASTP searches of the
337 metagenome scaffolds and MAGs containing the SMTs analyzed in this study and all bacterial and
338 eukaryotic genomes available in JGI IMG as of October 2019 using the *Methylococcus capsulatus*
339 Bath OSC protein sequence (locus tag: MCA2873, maximum e-value: $1e^{-50}$, minimum percent
340 identity: 20%, minimum sequence length: 350 aa). Redundancy was decreased through manual
341 removal of OSC homologs with 100% similarity at the amino acid level. The N-termini of fused
342 SMO/OSC homologs were trimmed out of the dataset to begin the sequences at the squalene-
343 hopene cyclase N-terminal domain as identified using HmmerWeb.

344 All sequence datasets were aligned via MUSCLE. Conserved SMT domains were aligned
345 separately, then concatenated using Geneious. Maximum-likelihood trees were generated using
346 IQ-Tree on XSEDE with 1,000 ultra-fast bootstrap replicates using the best amino acid substitution
347 models, gamma shape parameters, and invariable sites proportions under the Bayesian
348 Information Criterion and Akaike Information Criterion using ModelTest-NG on XSEDE (Table S3).
349 Resulting phylogenetic trees were visualized in iTOL (72).

350 All contigs containing metagenomic SMTs were retrieved from JGI IMG or GenBank and
351 analyzed with Whokaryote using the default inputs with the minimum contig size decreased to 2,000
352 base pairs. Calculated features and predictions are shown in Table S4.

353 **Gene Synthesis and Molecular Cloning.** SMT DNA sequences were codon-optimized
354 for expression in *E. coli* and artificially synthesized by GeneArt (ThermoFisher Scientific) or through

355 the Department of Energy Joint Genome Institute (DOE JGI) DNA Synthesis Science Program.
356 SMT sequences from DOE JGI were obtained in the IPTG-inducible plasmid pSRKGm-*lac*UV5-
357 rbs5 in *E. coli* TOP10. SMT sequences from GeneArt were subcloned into pSRKGm-*lac*UV5-rbs5
358 by sequence and ligase independent cloning (SLIC) (73) and transformed into *E. coli* DH10B as
359 described in *SI Appendix*.

360 **Bacterial Culture and Heterologous Expression.** *E. coli* expression strains were
361 cultured in 50 mL terrific broth (TB) supplemented with gentamycin (15 μg/mL) at 37 °C while
362 shaking at 225 rpm. Cultures were induced with 500 μM isopropyl *β*-D-1-thiogalactopyranoside
363 (IPTG) at an OD_{600} of ~0.6, then incubated an additional 4 hours at 30 °C while shaking at 225 rpm.
364 Cells were harvested by centrifugation at 4,500 × *g* for 10 minutes at 4 °C. Cell pellets were stored
365 at -80 °C until sonication.

366 **Sterol Methyltransferase Assay.** Cells pellets were resuspended in 5 mL buffer
367 containing the following: 50 mM Tris-HCl, 2 mM $MgCl_2$, 20% glycerol (v/v), and 0.1% β-
368 mercaptoethanol (v/v), pH 7.5. Cells were then lysed on ice via a Qsonica Sonicator Q500 equipped
369 with a 3.2 mm probe at 30% amplitude pulsing at 5 seconds on, 15 seconds off for 8 minutes of
370 total on time. Lysates were then partially clarified by centrifugation at 4,500 × *g* for 20 minutes at 4
371 °C. Total protein concentration in the resulting supernatant was quantified using a Coomassie
372 (Bradford) Protein Assay Kit (Thermo Scientific) according to the manufacturer's Standard
373 Microplate Protocol and a BioTek Synergy HT Microplate Reader. SMT assays were then
374 immediately prepared using enough supernatant to give 4500 μg total protein. Other reaction
375 components included 100 μM desmosterol (Sigma-Aldrich), 24-methylenecholesterol (Avanti Polar
376 Lipids, Inc.), [28-^{13}C] 24-methylenecholesterol (Avanti Polar Lipids, Inc.), or synthesized [29-^{13}C]
377 fucosterol or isofucosterol (see *SI Appendix*), 100 μM *S*-(5′-sdenosyl)-*L*-methionine chloride
378 dihydrochloride (Sigma-Aldrich), and 0.1% Tween-80 (v/v) to a total reaction volume of 400 μL.
379 Reactions were held at 30 °C for 20 hours and then stored at -20 °C until lipid extraction.

380 **Lipid Extraction.** Lipids were extracted using a modified Bligh-Dyer method (74, 75) with
381 the completed in vitro reactions as the water phase. Reactions were sonicated in 10:5:4 (vol:vol:vol)
382 methanol:dichloromethane:water for 1 hour. The organic phase was then separated with twice the

383 volume of 1:1 (vol:vol) dichloromethane:water followed by storage at -20 °C for >1 hr. Following
384 centrifugation at 2,800 x *g* for 10 min at 4 °C, the organic phase was transferred and evaporated
385 under N_2 to give total lipid extracts (TLEs). TLEs were derivatized to trimethylsilyl ethers in 1:1
386 (vol:vol) pyridine:Bis(trimethylsilyl)trifluoroacetamide for 1 hr at 70 °C prior to GC-MS analysis.

387 **GC-MS Analysis.** Lipids were separated with an Agilent 7890B Series GC equipped with
388 two Agilent DB-17HT columns (30 m x 0.25 mm i.d. x 0.15 µm film thickness) in tandem with helium
389 as the carrier gas at a constant flow of 1.1 ml/min and programmed as follows: 100°C for 2 min,
390 then 12°C/min to 250°C and held for 10 min, then 10°C/min to 330°C and held for 17.5 min. 2 uL
391 of each sample was injected in splitless mode at 250°C. The GC was coupled to an Agilent 5977A
392 Series MSD with the ion source at 230°C and operated at 70 eV in EI mode scanning from 50 to
393 850 Da in 0.5 s. Sterols were identified based on retention time and comparison to previously
394 published spectra and laboratory standards and NMR analysis.

395 **Structure confirmation of sterol products by NMR.** In addition to GC-MS analysis,
396 structures of sterol products formed by the metagenomic sterol methyltransferases referred to as
397 "Aplysina aerophoba JGIcombinedJ30088_10000835 1 of 2", "Petrosia ficiformis LXNJ01000389",
398 and "Chlamydiae sp., drinking water 1 of 2" were confirmed by NMR. Briefly, total lipid extracts from
399 pooled in vitro reactions were fractionated using small silica pipet columns (1 mL) eluting with 5 mL
400 each of Hexane/EtOAc 9:1, 2:1, and 1:1. Steroidal constituents were found in the 2:1 solvent
401 fraction and concentrated with a stream of N_2 prior to further purification by preparative TLC
402 (Hexane/EtOAc 4:1). Sterols were separated by HPLC and the fractions were analyzed by 800
403 MHz NMR (45). Sterols were identified by comparison of the HSQC-DEPT and HMBC spectra with
404 reference compounds.

405 **Acknowledgments**

406 We thank Prof. David A. Gold for providing demosponge SMT protein sequences, Jeremy
407 H. Wei for technical assistance, and Hanon McShea for helpful discussions. We also thank the
408 Department of Energy Joint Genome Institute DNA Synthesis Science Program (Project 503267)

for synthesis and cloning of the majority of the SMTs analyzed in this study. Funding for this study was provided by National Science Foundation Grant EAR-1752564 (to P.V.W).

References

1. T. M. Swan, K. Watson, Stress tolerance in a yeast sterol auxotroph: role of ergosterol, heat shock proteins and trehalose. *FEMS Microbiol Lett* **169**, 191–197 (1998).

2. L. Miao, M. Nielsen, J. Thewalt, J. H. Ipsen, M. Bloom, M. J. Zuckermann, O. G. Mouritsen, From lanosterol to cholesterol: structural evolution and differential effects on lipid bilayers. *Biophys J* **82**, 1429–1444 (2002).

3. H. Schaller, The role of sterols in plant growth and development. *Prog Lipid Res* **42**, 163–175 (2003).

4. A. B. Castoreno, Y. Wang, W. Stockinger, L. A. Jarzylo, H. Du, J. C. Pagnon, E. C. Shieh, A. Nohturfft, Transcriptional regulation of phagocytosis-induced membrane biogenesis by sterol regulatory element binding proteins. *Proc Natl Acad Sci U S A* **102**, 13129–13134 (2005).

5. P. Huang, D. Nedelcu, M. Watanabe, C. Jao, Y. Kim, J. Liu, A. Salic, Cellular cholesterol directly activates smoothened in hedgehog signaling. *Cell* **166**, 1176-1187.e14 (2016).

6. K. E. Bloch, Sterol, structure and membrane function. *Crit Rev Biochem Mol Biol* **14**, 47–92 (1983).

7. J. J. Brocks, G. D. Love, R. E. Summons, A. H. Knoll, G. A. Logan, S. A. Bowden, Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. *Nature* **437**, 866–870 (2005).

8. R. E. Summons, A. S. Bradley, L. L. Jahnke, J. R. Waldbauer, Steroids, triterpenoids and molecular oxygen. *Philos Trans R Soc Lond B Biol Sci* **361**, 951–968 (2006).

9. I. Bobrovskiy, J. M. Hope, A. Ivantsov, B. J. Nettersheim, C. Hallmann, J. J. Brocks, Ancient steroids establish the Ediacaran fossil *Dickinsonia* as one of the earliest animals. *Science* **361**, 1246–1249 (2018).

10. J. J. Brocks, A. J. M. Jarrett, E. Sirantoine, C. Hallmann, Y. Hoshino, T. Liyanage, The rise of algae in Cryogenian oceans and the emergence of animals. *Nature* **548**, 578–581 (2017).

11. G. D. Love, E. Grosjean, C. Stalvies, D. A. Fike, J. P. Grotzinger, A. S. Bradley, A. E. Kelly, M. Bhatia, W. Meredith, C. E. Snape, S. A. Bowring, D. J. Condon, R. E. Summons, Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* **457**, 718–721 (2009).

12. J. A. Zumberge, G. D. Love, P. Cárdenas, E. A. Sperling, S. Gunasekera, M. Rohrssen, E. Grosjean, J. P. Grotzinger, R. E. Summons, Demosponge steroid biomarker 26-methylstigmastane provides evidence for Neoproterozoic animals. *Nat Ecol Evol* **2**, 1709–1714 (2018).

13. W. Hofheinz, G. Oesterhelt, 24-isopropylcholesterol and 22-dehydro-24-isopropylcholesterol, novel sterols from a sponge. *Helv Chim Acta* **62**, 1307–1309 (1979).

14. H. Q. Do, L. Van Landeghem, G. Wielgosz-Collin, M. Takoudju, J.-M. Huvelin, J.-M. Kornprobst, J.-M. Bard, G. Barnathan, H. Nazih, Unusual sterolic mixture, and 24-isopropylcholesterol, from the sponge *Ciocalypta* sp. reduce cholesterol uptake and basolateral secretion in Caco-2 cells. *J Cell Biochem* **106**, 659–665 (2009).

15. M. Bortolotto, J. C. Braekman, D. Daloze, B. Tursch, Chemical studies of marine invertebrates. XXXVI. Strongylosterol, a novel C-30 sterol from the sponge *Strongylophora durissima* Dendy. *Bull Soc Chim Belges* **87**, 539–543 (1978).

16. G. Barnathan, N. Velosaotsy, S. Al-Lihaibi, J.-M. Njinkoue, J.-M. Kornprobst, J. Vacelet, N. Boury-Esnault, Unusual sterol composition and classification of three marine sponge families. *Boll Mus Ist Biol Univ Genova* **68**, 201–208 (2004).

17. J.-L. Giner, Biosynthesis of marine sterol side chains. *Chem Rev* **93**, 1735–1752 (1993).

18. J. A. Zumberge, "A lipid biomarker investigation tracking the evolution of the Neoproterozoic marine biosphere and the rise of eukaryotes," University of California, Riverside, Riverside, CA. (2019).

19. J. B. Antcliffe, R. H. T. Callow, M. D. Brasier, Giving the early fossil record of sponges a squeeze: the early fossil record of sponges. *Biol Rev* **89**, 972–1004 (2014).

20. J. B. Antcliffe, Questioning the evidence of organic compounds called sponge biomarkers. *Palaeontology* **56**, 917–925 (2013).

21. G. D. Love, R. E. Summons, The molecular record of Cryogenian sponges - a response to Antcliffe (2013). *Palaeontology* **58**, 1131–1136 (2015).

22. J. B. Antcliffe, The oldest compelling evidence for sponges is still early Cambrian in age - reply to Love and Summons (2015). *Palaeontology* **58**, 1137–1139 (2015).

23. G. D. Love, J. A. Zumberge, P. Cárdenas, E. A. Sperling, M. Rohrssen, E. Grosjean, J. P. Grotzinger, R. E. Summons, Sources of C30 steroid biomarkers in Neoproterozoic–Cambrian rocks and oils. *Nat Ecol Evol* **4**, 34–36 (2020).

24. C. Hallmann, B. J. Nettersheim, J. J. Brocks, A. Schwelm, J. M. Hope, F. Not, M. Lomas, C. Schmidt, R. Schiebel, E. C. M. Nowack, P. De Deckker, J. Pawlowski, S. S. Bowser, I. Bobrovskiy, K. Zonneveld, M. Kucera, M. Stuhr, Reply to: Sources of C30 steroid biomarkers in Neoproterozoic–Cambrian rocks and oils. *Nat Ecol Evol* **4**, 37–39 (2020).

25. I. Bobrovskiy, J. M. Hope, B. J. Nettersheim, J. K. Volkman, C. Hallmann, J. J. Brocks, Algal origin of sponge sterane biomarkers negates the oldest evidence for animals in the rock record. *Nat Ecol Evol* **5**, 165–168 (2021).

26. L. M. van Maldegem, B. J. Nettersheim, A. Leider, J. J. Brocks, P. Adam, P. Schaeffer, C. Hallmann, Geological alteration of Precambrian steroids mimics early animal signatures. *Nat Ecol Evol* **5**, 169–173 (2021).

27. W. C. M. C. Kokke, J. N. Shoolery, W. Fenical, C. Djerassi, Biosynthetic studies of marine lipids. 4. Mechanism of side chain alkylation in E-24-propylidenecholesterol by a Chrysophyte alga. *J Org Chem* **49**, 3742–3752 (1984).

28. J.-L. Giner, C. Djerassi, Biosynthetic studies of marine lipids. 31. Evidence for a protonated cyclopropyl intermediate in the biosynthesis of 24-propylidenecholesterol. *J Am Chem Soc* **113**, 1386–1393 (1991).

29. J.-L. Giner, X. Li, G. L. Boyer, Sterol composition of *Aureoumbra lagunensis*, the Texas brown tide alga. *Phytochemistry* **57**, 787–789 (2001).

30. J.-L. Giner, H. Zhao, G. L. Boyer, M. F. Satchwell, R. A. Andersen, Sterol chemotaxonomy of marine pelagophyte algae. *C&B* **6**, 1111–1130 (2009).

31. B. J. Nettersheim, J. J. Brocks, A. Schwelm, J. M. Hope, F. Not, M. Lomas, C. Schmidt, R. Schiebel, E. C. M. Nowack, P. De Deckker, J. Pawlowski, S. S. Bowser, I. Bobrovskiy, K. Zonneveld, M. Kucera, M. Stuhr, C. Hallmann, Putative sponge biomarkers in unicellular Rhizaria question an early rise of animals. *Nat Ecol Evol* **3**, 577–581 (2019).

32. I. Bobrovskiy, J. M. Hope, B. J. Nettersheim, J. K. Volkman, C. Hallmann, J. J. Brocks, Algal origin of sponge sterane biomarkers negates the oldest evidence for animals in the rock record. *Nat Ecol Evol* **5**, 165–168 (2021).

33. H. Gross, J. Reitner, GabrieleM. König, Isolation and structure elucidation of azoricasterol, a new sterol of the deepwater sponge *Macandrewia azorica*. *Naturwissenschaften* **91**, 441–446 (2004).

34. M. W. Taylor, R. Radax, D. Steger, M. Wagner, Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *MMBR* **71**, 295–347 (2007).

35. J. H. Wei, X. Yin, P. V. Welander, Sterol synthesis in diverse bacteria. *Front Microbiol* **7**, 990 (2016).

36. N. S. Webster, M. W. Taylor, Marine sponges and their microbial symbionts: love and other relationships. *Environ Microbiol* **14**, 335–346 (2012).

37. W. D. Nes, Enzyme mechanisms for sterol C -methylations. *Phytochemistry* **64**, 75–95 (2003).

38. D. A. Gold, J. Grabenstatter, A. de Mendoza, A. Riesgo, I. Ruiz-Trillo, R. E. Summons, Sterol and genomic analyses validate the sponge biomarker hypothesis. *Proc Natl Acad Sci U S A* **113**, 2684–2689 (2016).

39. D. A. Gold, S. S. O'Reilly, J. Watson, B. M. Degnan, S. M. Degnan, J. O. Krömer, R. E. Summons, Lipidomics of the sea sponge *Amphimedon queenslandica* and implication for biomarker geochemistry. *Geobiology* **15**, 836–843 (2017).

40. W. Bergmann, F. H. McTigue, Contributions to the study of marine products. XXI. Chondrillasterol. *J Org Chem* **13**, 738–741 (1948).

41. Y. Venkateswarlu, M. V. R. Reddy, M. R. Rao, A new epoxy sterol from the sponge *Ircinia fasciculata*. *J Nat Prod* **59**, 876–877 (1996).

42. J. Fromont, S. Kerr, R. Kerr, M. Riddle, P. Murphy, Chemotaxonomic relationships within, and comparisons between, the orders Haplosclerida and Petrosida (Porifera: Demospongiae) using sterol complements. *Biochem Syst Ecol* **22**, 735–752 (1994).

43. A. Aiello, E. Fattorusso, S. Magno, M. Menna, Isolation of five new 5α-hydroxy-6-keto-Δ7 sterols from the marine sponge *Oscarella lobularis*. *Steroids* **56**, 337–340 (1991).

44. P. R. Bergquist, W. Hofheinz, W. Hofheinz, G. Oesterhelt, Sterol composition and the classification of the demospongiae. *Biochem Syst Ecol* **8**, 423–435 (1980).

527 45. A. K. Lee, A. B. Banta, J. H. Wei, D. J. Kiemle, J. Feng, J.-L. Giner, P. V. Welander, C-4
528 sterol demethylation enzymes distinguish bacterial and eukaryotic sterol synthesis. *Proc*
529 *Natl Acad Sci USA* **115**, 5884–5889 (2018).

530 46. C. Santana-Molina, E. Rivas-Marin, A. M. Rojas, D. P. Devos, Origin and evolution of
531 polycyclic triterpene synthesis. *Mol Biol Evol* **37**, 1925–1941 (2020).

532 47. L. Villanueva, W. I. C. Rijpstra, S. Schouten, J. S. S. Damsté, Genetic biomarkers of the
533 sterol-biosynthetic pathway in microalgae. *Environ Microbiol Rep* **6**, 35–44 (2014).

534 48. L. J. U. Pronk, M. H. Medema, Whokaryote: distinguishing eukaryotic and prokaryotic contigs
535 in metagenomes based on gene structure. *Microbial Genomics* **8** (2022).

536 49. D. Michellod, T. Bien, D. Birgel, M. Jensen, M. Kleiner, S. Fearn, C. Zeidler, H. R. Gruber-
537 Vodicka, N. Dubilier, M. Liebeke, "*De novo* phytosterol synthesis in animals". *bioRxiv*,
538 https:/doi.org/10.1101/2022.04.22.489198 (2022).

539 50. D. Gold, T. Brunoir, C. Mulligan, A. Sistiaga, K. Vuu, P. Shih, S. OReilly, R. Summons,
540 "Sterol methyltransferases in annelid worms rewrite the molecular fossil record". *Research*
541 *Square*, https:/doi.org/10.21203/rs.3.rs-1686449/v1 (2022).

542 51. P. R. Bergquist, A. Lavis, R. C. Cambie, Sterol composition and classification of the porifera.
543 *Biochem Syst Ecol* **14**, 105–112 (1986).

544 52. J.-L. Giner, C. Djerassi, Biosynthetic studies of marine lipids-XXVIII. Use of sponge cell-free
545 extracts in the study of marine sterol biosynthesis. *Tetrahedron Lett* **31**, 5421–5424
546 (1990).

547 53. R. G. Kerr, I. L. Stoilov, J. E. Thompson, C. Djerassi, Biosynthetic studies of marine lipids 16.
548 De novo sterol biosynthesis in sponges. Incorporation and transformation of cycloartenol
549 and lanosterol into unconventional sterols of marine and freshwater sponges. *Tetrahedron*
550 **45**, 1893–1904 (1989).

551 54. C. J. Silva, L. Wünsche, C. Djerassi, Biosynthetic studies of marine lipids 35. The
552 demonstration of de novo sterol biosynthesis in sponges using radiolabeled isoprenoid
553 precursors. *Comp Biochem Physiol B* **99B**, 763–773 (1991).

554 55. J. Germer, N. Cerveau, D. J. Jackson, The holo-transcriptome of a calcified early branching
555 metazoan. *Front Mar Sci* **4**, 81 (2017).

556 56. J. B. Kelly, D. Carlson, J. S. Low, R. W. Thacker, "Comparative metagenomics evidence
557 distinct trends of genome evolution between sponge-dwelling bacteria and their pelagic
558 counterparts". *bioRvix*, https:/doi.org/10.1101/2020.08.31.276493 (2022).

559 57. K. Ganapathy, C. W. Jones, C. M. Stephens, R. Vatsyayan, J. A. Marshall, W. D. Nes,
560 Molecular probing of the *Saccharomyces cerevisiae* sterol 24-C methyltransferase reveals
561 multiple amino acid residues involved with C2-transfer activity. *Biochim Biophys Acta Mol*
562 *Cell Biol Lipids* **1781**, 344–351 (2008).

563 58. W. D. Nes, B. S. McCourt, J. A. Marshall, J. Ma, A. L. Dennis, M. Lopez, H. Li, L. He, Site-
564 directed mutagenesis of the sterol methyl transferase active site from *Saccharomyces*
565 *cerevisiae* results in formation of novel 24-ethyl sterols. *J Org Chem* **64**, 1535–1542
566 (1999).

59. W. D. Nes, J. A. Marshall, Z. Jia, T. T. Jaradat, Z. Song, P. Jayasimha, Active site mapping and substrate channeling in the sterol methyltransferase pathway. *J Biol Chem* **277**, 42549–42556 (2002).

60. B. A. Haubrich, E. K. Collins, A. L. Howard, Q. Wang, W. J. Snell, M. B. Miller, C. D. Thomas, S. K. Pleasant, W. D. Nes, Characterization, mutagenesis and mechanistic analysis of an ancient algal sterol C24-methyltransferase: implications for understanding sterol evolution in the green lineage. *Phytochemistry* **113**, 64–72 (2015).

61. W. D. Nes, A. Sinha, P. Jayasimha, W. Zhou, Z. Song, A. L. Dennis, Probing the sterol binding site of soybean sterol methyltransferase by site-directed mutagenesis: functional analysis of conserved aromatic amino acids in region 1. *Arch Biochem Biophys* **448**, 23–30 (2006).

62. J. F. Hu, P. A. Peng, M. Y. Liu, D. P. Xi, J. Z. Song, X. Q. Wan, C. S. Wang, Seawater incursion events in a Cretaceous paleo-lake revealed by specific marine biological markers. *Sci Rep* **5**, 9508 (2015).

63. M. McCaffrey, J. Michaelmoldowan, P. Lipton, R. Summons, K. E. PETERs, D. Watt', Paleoenvironmental implications of novel C30 steranes in Precambrian to Cenozoic age petroleum and bitumen. *Geochim Cosmochim Acta* **58**, 529–532 (1994).

64. J. M. Moldowan, C30-steranes, novel markers for marine petroleums and sedimentary rocks. *Geochim Cosmochim Acta* **48**, 2767–2768 (1984).

65. C. A. Elwell, J. N. Engel, Lipid acquisition by intracellular Chlamydiae. *Cell Microbiol* **14**, 1010–1018 (2012).

66. L. Wünsche, F. O. Gülaçar, A. Buchs, Several unexpected marine sterols in a freshwater sediment. *Org Geochem* **11**, 215–219 (1987).

67. I. L. Stoilov, J. E. Thompson, J. Ho. Cho, Carl. Djerassi, Biosynthetic studies of marine lipids. 9. Stereochemical aspects and hydrogen migrations in the biosynthesis of the triply alkylated side chain of the sponge sterol strongylosterol. *J Am Chem Soc* **108**, 8235–8241 (1986).

68. C. A. N. Catalan, J. E. Thompson, W. C. M. C. Kokke, C. Djerassi, Biosynthetic studies of marine lipids—3: experimental demonstration of the course of side chain extension in marine sterols. *Tetrahedron* **41**, 1073–1084 (1985).

69. K. J. Peterson, J. A. Cotton, J. G. Gehling, D. Pisani, The Ediacaran emergence of bilaterians: congruence between the genetic and the geological fossil records. *Philos Trans R Soc Lond B Biol Sci* **363**, 1435–1443 (2008).

70. D. H. Erwin, M. Laflamme, S. M. Tweedt, E. A. Sperling, D. Pisani, K. J. Peterson, The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* **334**, 1091–1097 (2011).

71. M. Dohrmann, G. Wörheide, Dating early animal evolution using phylogenomic data. *Sci Rep* **7**, 3599 (2017).

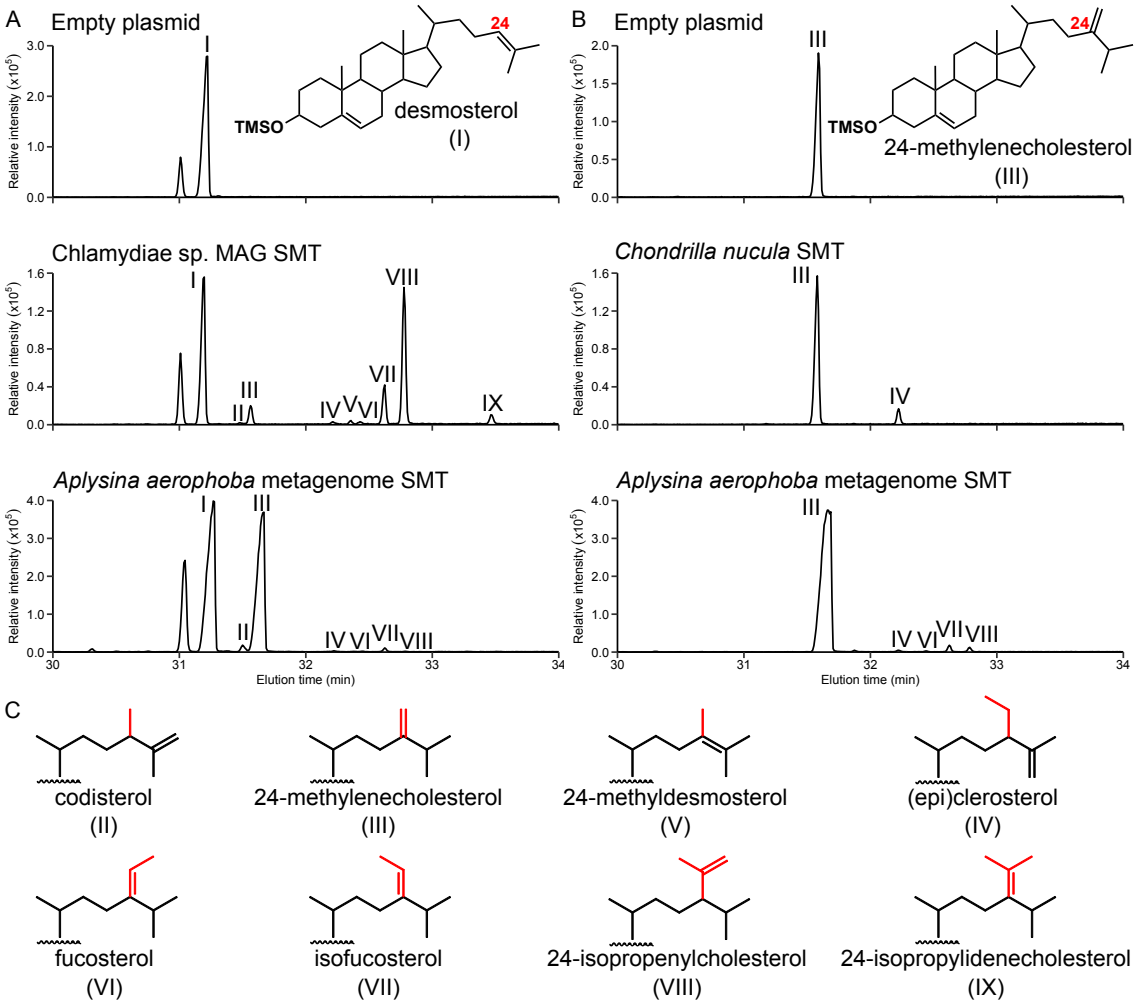
72. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**, W293–W296 (2021).

607 73. M. Z. Li, S. J. Elledge, SLIC: a method for sequence- and ligation-independent cloning.
608 *Methods Mol Biol* **852**, 51–59 (2012).

609 74. P. V. Welander, D. M. Doughty, C.-H. Wu, S. Mehay, R. E. Summons, D. K. Newman,
610 Identification and characterization of *Rhodopseudomonas palustris* TIE-1 hopanoid
611 biosynthesis mutants. *Geobiology* **10**, 163–177 (2012).

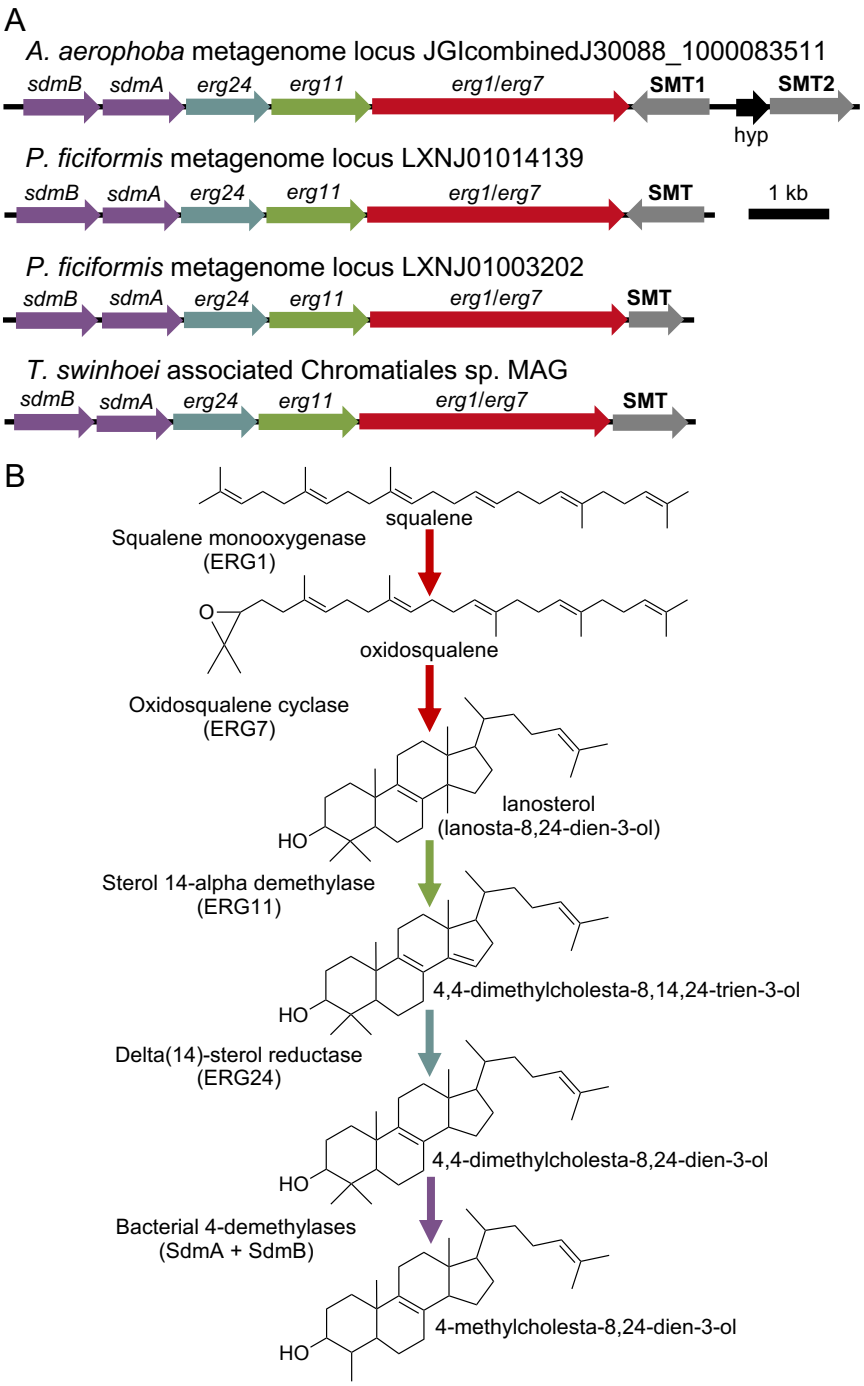
612 75. E. G. Bligh, W. J. Dyer, A rapid method of total lipid extraction and purification. *Can J*
613 *Biochem Physiol* **37**, 911–917 (1959).

614 **Figures and Tables**



615 **Figure 1.** Sterol methyltransferases from sponges and bacteria perform multiple side chain
616 methylations. (*A*) Representative extracted ion chromatograms (*m/z* 456, 470, 484, 498) of total
617 lipid extracts from in vitro reactions performed with *E. coli* lysates containing an empty plasmid
618 (*Upper*), a sterol methyltransferase (SMT) identified in a freshwater metagenome-assembled
619 genome assigned to the phylum Chlamydiae (*Middle*), and an SMT identified in a metagenome
620 from the microbiome of the demosponge *Aplysina aerophoba* (*Lower*) with desmosterol as the
621 substrate. The Chlamydiae MAG SMT produces codisterol, 24-methylenecholesterol, 24-
622 methyldesmosterol, epiclerosterol, fucosterol, isofucosterol, 24-isopropenylcholesterol, and 24-
623 isopropylidenecholesterol. The *A. aerophoba* metagenome SMT produces codisterol, 24-
624 methylenecholesterol, epiclerosterol, fucosterol, isofucosterol, and 24-isopropenylcholesterol from

625 desmosterol. (*B*) Representative extracted ion chromatograms (*m/z* 456, 470, 484, 498) of total
626 lipid extracts from in vitro reactions performed with *E. coli* lysates containing an empty plasmid
627 (*Upper*), an SMT identified in a transcriptome of the demosponge *Chondrilla nucula* (*Middle*), and
628 the same SMT identified in the *A. aerophoba* metagenome as in (A) (*Lower*) with 24-
629 methylenecholesterol as the substrate. The *C. nucula* SMT produces (epi)clerosterol, and the *A.*
630 *aerophoba* metagenome SMT produces the same sterols as in (A). (*C*) Side chain structures of
631 sterols identified in (A) and (B). All lipids were derivatized to trimethylsilyls. Mass spectra of
632 identified sterols are shown in *SI Appendix*, Fig. S1.



633 **Figure 2.** Sterol gene clusters in demosponge metagenomes suggest de novo biosynthesis by
634 bacterial symbionts. (*A*) Putative sterol biosynthesis gene clusters identified in metagenomes from
635 the microbiomes of the demosponge species *Aplysina aerophoba, Petrosia ficiformis*, and
636 *Theonella swinhoei*, each containing a sterol methyltransferase (gray arrows). Each gene cluster
637 contains homologs of a fused squalene monooxygenase/oxidosqualene cyclase (red arrows),

638 sterol 14-alpha demethylase (green arrows), delta(14)-sterol reductase (blue arrows), and the
639 bacterial sterol demethylases *sdmA* and *sdmB* (purple arrows). *(B)* Schematic of the bacterial sterol
640 biosynthesis pathway from squalene to 4-methylcholesta-8,24,-dien-3-ol. Arrow colors at each step
641 correspond to the putative homologs identifiied in (A). Detailed information on the source of each
642 metagenomic locus is available in *SI Appendix*, Table S2.

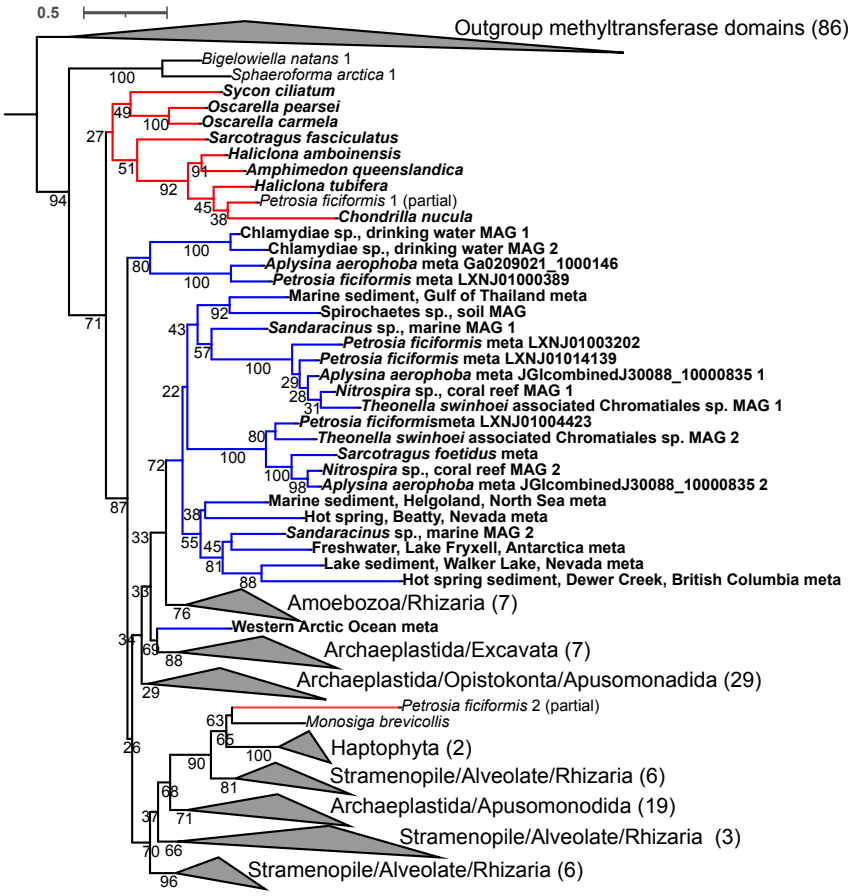


Figure 3. Maximum likelihood tree of sterol methyltransferases. The majority (23 out of 24) of the putative bacterial SMTs tested here (blue) form two clades distinct from sponge SMTs (red) and other eukaryotic SMTs (black), although bootstrap support is low. The numbers in parentheses indicate the number of proteins in collapsed clades. Bolded labels indicate SMTs tested in this study.

648 **Table 1.** Sterols identified as products of sponge and bacterial sterol methyltransferases.

	C_{28}			C_{29}			C_{30}	
	codisterol (II)	24-methylenecholesterol (III)	24-methyldesmosterol (V)	(epi)clerosterol (IV)*	fucosterol (VI)	isofucosterol (VIII)	24-isoprenylcholesterol (VIII)	24-isopropylidenecholesterol (IX)
DEMOSPONGIAE								
Amphimedon queenslandica		+			+			
Chondrilla nucula	+	+		+				
Sarcotragus fasciculatus	+	+			+			
Haliclona amboinensis	+	+		+	+	+		
Haliclona tubifera	+	+		+				
HOMOSCLEROMORPHA								
Oscarella pearsei	+	+		+	+	+		
Oscarella carmela	+	+						
CALCAREA								
Sycon ciliatum	+	+						
DEMOSPONGE METAGENOME								
Aplysina aerophoba JGIcombinedJ30088_10000835 1 of 2	+	+		+	+	+	+	
Aplysina aerophoba JGIcombinedJ30088_10000835 2 of 2								
Aplysina aerophoba Ga0209021_1000146	+	+	+		+	+		+
Petrosia ficiformis LXNJ01014139		+						
Petrosia ficiformis LXNJ01003202		+						
Petrosia ficiformis LXNJ01000389	+	+	+		+			
Petrosia ficiformis LXNJ01004423								
Sarcotragus foetidus		+						
DEMOSPONGE METAGENOME-ASSEMBLED GENOME								
Theonella swinhoei associated Chromatiales sp. 1 of 2		+						
Theonella swinhoei associated Chromatiales sp. 2 of 2								
OTHER METAGENOME								
Marine sediment, Gulf of Thailand		+						
Marine sediment, Helgoland, North Sea		+						
Western Arctic Ocean	+	+		+	+	+		
Freshwater, Lake Fryxell, Antarctica	+	+			+	+		
Lake sediment, Walker Lake, Nevada		+						
Hot spring, Beatty, Nevada								
Hot spring sediment, Dewer Creek, British Columbia		+						
OTHER METAGENOME-ASSEMBLED GENOME								
Chlamydiae sp., drinking water 1 of 2**	+	+	+	+	+	+	+	+
Chlamydiae sp., drinking water 2 of 2	+		+					
Nitrospira sp., coral reef 1 of 2	+	+		+				
Nitrospira sp., coral reef 2 of 2								
Sandaracinus sp., marine 1 of 2	+	+						
Sandaracinus sp., marine 2 of 2	+	+		+	+	+		
Spirochaetes sp., soil		+						

649 *NMR confirmed epiclerosterol and not clerosterol as products of the "*Aplysina aerophoba*
650 JGIcombinedJ30088_10000835 1 of 2" and "Chlamydiae sp., drinking water 1 of 2" sterol
651 methyltransferases.

652 **24-isopropylcholesta-5,23-dienol was also detected as a product of the "Chlamydiae sp., drinking
653 water 1 of 2" sterol methyltransferase by NMR, but was not detected by GC-MS.