1 # Characterization of sterol synthesis in bacteria

2

3 Jeremy H. Wei, Xinchi Yin and Paula V. Welander*

4 Department of Earth System Science, Stanford University, Stanford, CA, USA

5

6

7 *Correspondence:

8 Dr. Paula V. Welander

9 Stanford University

10 Department of Earth System Science

11 473 Via Ortega, Rm 140

12 Stanford, CA 94305

13 welander@stanford.edu

14

15 **Running Title:** Bacterial sterol synthesis

18

19   **Abstract**

20

21   Sterols are essential components of eukaryotic cells whose biosynthesis and function in

22   eukaryotes has been studied extensively. Sterols are also recognized as the diagenetic precursors

23   of steranes preserved in sedimentary rocks where they can function as geological proxies for

24   eukaryotic organisms and/or aerobic metabolisms and environments. However, production of

25   these lipids is not restricted to the eukaryotic domain as a few bacterial species also synthesize

26   sterols. Phylogenomic studies have identified genes encoding homologs of sterol biosynthesis

27   proteins in the genomes of several additional species, indicating that sterol production may be

28   more widespread in the bacterial domain than previously thought. Although the occurrence of

29   sterol synthesis genes in a genome indicates the potential for sterol production, it provides

30   neither conclusive evidence of sterol synthesis nor information about the composition and

31   abundance of basic and modified sterols that are actually being produced. Here, we coupled

32   bioinformatics with lipid analyses to investigate the scope of bacterial sterol production. We

33   identified oxidosqualene cyclase (Osc), which catalyzes the initial cyclization of oxidosqualene

34   to the basic sterol structure, in 34 bacterial genomes from 5 phyla (Bacteroidetes, Cyanobacteria,

35   Planctomycetes, Proteobacteria and Verrucomicrobia) and in 176 metagenomes. Our data indicate that

36   bacterial sterol synthesis likely occurs in diverse organisms and environments and also provides

37   evidence that there are as yet uncultured groups of bacterial sterol producers. Phylogenetic

38   analysis of bacterial and eukaryotic Osc sequences revealed two potential lineages of the sterol

39   pathway in bacteria indicating a complex evolutionary history of sterol synthesis in this domain. We

40   characterized the lipids produced by Osc-containing bacteria and found that we could generally

41   predict the ability to synthesize sterols. However, predicting the final modified sterol based on our current

42   knowledge of bacterial sterol synthesis was difficult. Some bacteria produced demethylated and

43   saturated sterol products even though they lacked homologs of the eukaryotic proteins required

44   for these modifications emphasizing that several aspects of bacterial sterol synthesis are still completely

45   unknown. It is possible that bacteria have evolved distinct proteins for catalyzing sterol

46   modifications and this could have significant implications for our understanding of the

47   evolutionary history of this ancient biosynthetic pathway.

48   **Introduction**

49   Sterols are tetracyclic triterpenoid lipids that are required by all eukaryotes for critical cellular

50   functions including maintaining membrane fluidity, phagocytosis, stress tolerance and cell

51   signaling (Bloch, 1991;Swan and Watson, 1998;Castoreno et al., 2005;Xu et al., 2005;Riobo,

52   2012). Studies on the biosynthesis of sterols in eukaryotes have revealed a variety of novel

53   biochemical reactions while molecular and cell biological studies have revealed unique

54   regulatory mechanisms and key insights into sterol transport (Dimster-Denk and Rine,

55   1996;Yang, 2006;Nes, 2011). Geochemists also have an interest in these molecules as they have

56   the potential to function as 'molecular fossils' (Summons et al., 2006;Love et al., 2009). Sterols,

57   like many polycyclic triterpenoid lipids, are quite recalcitrant and their degradation products, the

58   steranes, are readily preserved in ancient sediments. Sterane signatures in the rock record date as

59   far back as 1.6 billion years (Brocks et al., 2005) and, based on their distribution in modern

60   eukaryotes, are utilized as biomarkers for the existence of specific eukaryotic organisms at the

61   time of deposition (Peters et al., 2007a;b). Because eukaryotes are the predominant extant

62   producers of sterols and because they require sterols for growth, the use of steranes as

63   biomarkers for eukaryotes seems robust. However, sterol production has been observed in a few

64   bacterial species raising the question as to whether bacterial sterol production is significant for

65   the interpretation of sterane signatures (Volkman, 2003;2005).

66   Bacterial sterol production was first discovered in the aerobic methanotroph

67   *Methylococcus capsulatus* Bath (Bird et al., 1971;Bouvier et al., 1976). *M. capsulatus* was

68   shown to produce 4,4-dimethylcholesta-8,24-dien-3-ol, 4,4-dimethylcholesta-8-en-3-ol, 4-

69   methylcholesta-8,24-dien-3-ol, 4-methylcholesta-8-en-3-ol (Figure 1). Subsequent studies have

70   demonstrated the production of similar sterols in other aerobic methanotrophs of the

71   Methylococcales order within the γ-Proteobacteria (Schouten et al., 2000;Banta et al., 2015). In

72   addition, sterol biosynthesis has also been observed in a few myxobacteria of the δ-

73   Proteobacteria (Bode et al., 2003) and the plantomycete *Gemmata obscuriglobus* (Pearson et al.,

74   2003). *G. obscuriglobus* produces the least biosynthetically complex sterols, lanosterol and the

75   rare lanosterol isomer parkeol. The myxobacteria tend to produce more modified sterols. In

76   particular, *Nannocystis excedens* has been shown to produce primarily cholest-7-en-3-ol

77   (lathosterol) and cholest-8-en-3-ol (Bode et al., 2003).

78   Recent phylogenetic analyses of bacterial genomes have revealed that the potential for

79   sterol production in this domain might extend beyond what is currently known (Desmond and

80   Gribaldo, 2009;Villanueva et al., 2014). Desmond and Gribaldo identified sterol biosynthesis

81   genes in the myxobacterium *Plesiocystis pacifica* and proposed that this organism had the

82   potential to produce cholesta -7,24-dienol-3β-ol. In addition, Villanueva et al. observed

83   oxidosqualene cyclase (Osc) homologs, required for the initial cyclization of oxidosqualene to

84   lanosterol or cycloartenol (Figure 1) in a variety of bacterial genomes including three aerobic

85   methanotrophs, two Bacteriodetes species and one cyanobacterium symbiont. Further,

86   phylogenetic analyses of bacterial Osc homologs have begun to hint at the evolutionary history

87   of the sterol biosynthetic pathway in bacterial species. In particular, two studies have proposed

88   that bacteria acquired the sterol biosynthetic pathway through horizontal gene transfer from

89   eukaryotes (Desmond and Gribaldo, 2009;Frickey and Kannenberg, 2009).

90   While genomic and phylogenetic data may provide some clues to the diversity and

91   evolutionary history of the sterol biosynthetic pathway in bacteria, it is important to note that the

92   occurrence of an oxidosqualene cyclase in a bacterial genome demonstrates the potential to

93   produce sterols but it is not conclusive evidence that sterol production is actually occurring.

94   Also, the presence of oxidosqualene cyclase is only indicative of the initial cyclization required

95   to produce the most basic sterols and it does not provide any insight into how sterols may be

96   modified in these Osc containing bacteria. Sterol production is not uniform across all eukaryotes

97   both in terms of the final products produced and in the proteins and enzymatic mechanisms

98   involved in their biosynthesis (Hartmann, 1998;Volkman, 2003;Summons et al., 2006).

99   Vertebrates synthesize cholesterol as a final product while fungi generate ergosterol and plants

100   tend to make stigmasterol (Figure 1). Biosynthetically, plants have a distinct cyclase in which

101   they generate the cyclopropylsterol cycloartenol after cyclization of squalene while vertebrates

102   and fungi generate lanosterol (Figure 1) (Desmond and Gribaldo, 2009). Downstream

103   modifications, including methylations, unsaturation and isomerization, in eukaryotes also differ

104   and it is unclear whether sterol biosynthesis in bacteria is more similar to one (or none) of these

105   pathways.

106   To fully understand both the potential for sterol production in the bacterial domain and

107   the evolutionary history of the sterol biosynthetic pathway, further studies are needed to

108   characterize sterol production in bacterial species. In this study, we identify potential sterol

109   biosynthesis genes in a variety of bacterial genomes and metagenomes. We then characterized

110   the lipid profiles of a subset of these potential sterol-producers and demonstrate that all but one

111   of the organisms we tested were capable of sterol production under laboratory conditions.

112   Through these studies, it is evident that sterol production is more widespread in the bacterial

113   domain than previously thought and that the bacterial sterol biosynthetic pathway has a complex

114   evolutionary history.

115   **Materials and Methods**

116   **Bioinformatics Analyses.** Homologs of the *Methylococcus capsulatus* Osc protein (locus

117   tag: MCA2873) were identified through BLASTP (Altschul et al., 1997) searches of all bacterial

118   (31,237) and eukaryotic (220) genomes as well as all environmental metagenomes (2,707) on the

119   Joint Genome Institute Integrated Microbial Genomes database (http://img.jgi.doe.gov/).

120   MUSCLE (Edgar, 2004) alignments of bacterial and eukaryotic Osc sequences with an e-value

121   of 1e-50 or lower were generated in Geneious (Biomatters). For metagenomic sequences, only

122   Osc candidates larger than 400 amino acids were included in the alignment. Redundancy in

123   metagenomic alignments was decreased utilizing the Decrease Redundancy web tool

124   (http://web.expasy.org/decrease_redundancy/). Due to large gaps in the metagenomic alignments

125   it was also necessary to extract conserved regions with GBLOCKS (Talavera and Castresana,

126   2007) prior to building phylogenetic trees containing metagenomics sequences. PhyML

127   (Guindon and Gascuel, 2003) was utilized to generate maximum likelihood phylogenetic trees

128   using the LG+gamma model, four gamma rate categories, ten random starting trees, NNI branch

129   swapping, and substitution parameters estimated from the data. Resulting phylogenetic trees

130   were edited in the Interactive Tree of Life (iTOL) website (http://itol.embl.de/) (Letunic and

131   Bork, 2007;2011).

132   The following yeast and bacterial proteins were used to search bacterial genomes

133   (BLASTP) for homologs of other sterol biosynthesis proteins: squalene epoxidase (*M. capsulatus*

6

134   locus tag: MCA2872); lanosterol 14-α-demethylase ERG11 (*Saccharomyces cerevisiae* locus

135   tag: YHR007C); C-14 sterol reductase ERG24 (*S. cerevisiae* locus tag: YNL280C); C-4 methyl

136   sterol oxidase ERG25 (*S. cerevisiae* locus tag: YGR060W); C-3 sterol dehydrogenase ERG26 (*S.*

137   *cerevisiae* locus tag: YGL001C); 3-keto sterol reductase ERG27 (*S. cerevisiae* locus tag:

138   YLR100W); C-24(28) sterol reductase ERG4 (*S. cerevisiae* locus tag: YGL012W0; 24-

139   dehydrocholesterol reductase (*Homo sapiens* locus tag: HGNC:2859). The e-value cut-off for a

140   potential homolog of these proteins was set at 1e-10 or lower with a minimum 20% identity.

141   **Lipid Analysis.** Bacterial strains surveyed for sterol production and their growth

142   conditions are described in Table 1. All strains were grown in our laboratory except for

143   *Enhygromyxa salina* DSM15201, *Plesiocystis pacifica* SIR-1 DSM14875, and *Sandaracinus*

144   *amylolyticus* DSM53668. For lipid analysis of these three strains, cells were scraped directly

145   from the agar plates purchased from the German Collection of Microorganisms and Cell Cultures

146   (DSMZ; https://www.dsmz.de/), placed in 2 ml of deionized water and stored at -20°C. All

147   liquid cultures were centrifuged at 5,000 x g for 10 minutes at 4°C and the supernatant was

148   discarded. Cell pellets were frozen at -20°C prior to lipid extraction.

149   Frozen cell pellets were resuspended in 2 ml of deionized water and transferred to a

150   solvent washed Teflon centrifuge tube. 5 ml of methanol and 2.5 ml of dichloromethane were

151   added and the cell mixture was sonicated for 1 hour. 10 ml of deionized water and 10 ml of

152   dichloromethane were added to samples after sonication, mixed and incubated at -20°C

153   overnight. Samples were centrifuged for 10 minutes at 2800 x g and the organic layer was

154   transferred to a 40 ml baked glass vial. The total lipid extract was evaporated under N_2 and

155   derivatized to acetate or trimethylsilyl (TMS) esters. For derivatization, TLEs were treated with

156   50 ul of pyridine and 50 ul of acetic anhydride to create acetate derivatives, or with 25 ul of

157   pyridne and 25 ul of TMS + 1% N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA) for

158   trimethylsilyl derivative compounds. Samples were dried under N_2 after derivatization and

159   resuspended in 50-200 μl of dichloromethane prior to high temperature gas chromatography-

160   mass spectrometry (GC-MS) analysis (Sessions et al., 2013).

161   Lipid extracts were separated on an Agilent 7890B Series GC with helium as the carrier

162   gas at a constant flow of 1.0-1.2 ml/min and programmed as follows: 100°C for 2 min, ramp

163   15°C/min to 320°C and hold 28-30 min. Analyses were done on a DB5-HT column (30 m X 0.25

164   mm i.d. X 0.1 μm film thickness) or a DB17-HT column (30 m x 0.25 mm i.d. x 0.15 μm film

165   thickness). 2 μl of the sample were injected into a Gerstel-programmable temperature

166   vaporization (PTV) injector, operated in splitless mode at 320°C. The GC was coupled to a

167   5977A Series MSD with the source at 230°C and operated at 70 eV in EI mode scanning from

168   50-850 Da in 0.5 s. All lipids were identified based on their retention time and mass spectra

169   (Figure 2) as well as comparison to prepared internal standards and previously published spectra.

170   **Results**

171   **Identification of oxidosqualene cyclase homologs in bacterial genomes**

172   To identify potential bacterial sterol producers, we queried all bacterial genomes in the

173   Joint Genome Institute Integrated Microbial Genomes database (http://img.jgi.doe.gov/) for

174   homologs of the *M. capsulatus* Bath oxidosqualene cyclase (locus tag: MCA2873).

175   Oxidosqualene cyclases catalyze the conversion of oxidosqualene to lanosterol in vertebrates and

176   fungi and to cycloartenol in land plants (Desmond and Gribaldo, 2009). Deletion of this protein

177   in yeast completely blocks sterol production and therefore its occurrence in a genome is a good

178   indicator of sterol production by an organism (Lees et al., 1995). BLASTP analysis recovered 34

179   bacterial Osc homologs in five different bacterial phyla with an e-value equal to or lower than e[-]

180   ^{100} and greater than 30% similarity (Table 2). As expected, Osc homologs are found in the

181   genomes of five organisms that have been previously shown to produce sterols: *M. capsulatus*

182   Bath, *N. exedens*, *Cystobacter fuscus*, *Stigmatella aurantiaca* and *G. obscuriglobus* (Table 2).

183   The myxobacterium *Corallococcus coralloides* also contains an Osc homolog, however, a

184   previous study of myxobacterial species did not detect any sterols in this bacterium (Bode et al.,

185   2003). Prior phylogenetic studies have also identified Osc homologs in the genomes of

186   *Pleisocystis pacifica* (Desmond and Gribaldo, 2009), *Eudoraea adriatica*, *Fluviicola taffensis*,

187   *Methylobacter marinus*, *Methylomicrobium buryatense* and *Methylomicrobium alcaliphilum*

188   (Villanueva et al., 2014) which we also observed here. However, with the exception of *M.*

189   *alcaliphilum* (Banta et al., 2015), lipid analysis of these species have not been undertaken to

190   verify sterol production. *Prochloron didenmi* is a cyanobacterial obligate symbiont of the marine

191   ascidian *Lissoclinum patella*. *P. didenmi* has not been isolated in pure culture but partial genome

192   sequencing of this symbiont has previously revealed an Osc homolog and lanosterol has been

193   observed in whole ascidian extracts (Donia et al., 2011). Our bioinformatics analysis also

194   detected Osc homologs in strains not previously shown to produce sterols including three

195   myxobacteria, ten Methylococcales, two Cyanobacteria, one α-Proteobacterium

196   (*Methyloceanibacter caenitepidi*) and one Verrucomicrobia (*Verrucomicrobiaceae bacterium*).

197   A maximum likelihood tree of bacterial and eukaryotic Osc protein homologs was

198   created by aligning the 34 bacterial Osc homologs with 70 eukaryotic Osc sequences and 23

199   bacterial squalene-hopene cyclase sequences (Shc) as the outgroup. Squalene-hopene cyclases

200   catalyze the conversion of squalene to the polycyclic hopanoid diploptene and are structurally

201   and functionally similar to oxidosqualene cyclases (Siedenburg and Jendrossek, 2011). The

202   phylogenetic tree revealed two bacterial clades of Osc homologs (Figure 3) which is similar to

9

203   what was observed by Villanueva et al. (Villanueva et al., 2014). One bacterial Osc group

204   (Group 2) falls within the eukaryotic Osc clades while the Group 1 sequences form a sister clade

205   to the eukaryotic cyclases (Figure 3). Further, the clustering of these two bacterial clades of Osc

206   sequences was not congruent with taxonomic phylogeny based on 16S rDNA. For example, we

207   observe Osc homologs from strains of myxobacteria and aerobic methanotrophs clustering in

208   both Group 1 and 2 (Figure 3). This distribution of bacterial Osc homologs suggests a

209   complicated evolutionary history that potentially involves both gene transfer and gene loss.

210   The 34 bacterial species with Osc homologs in their genomes were isolated from a

211   variety of environments indicating that bacterial sterol producers are not restricted to a specific

212   ecological niche (Table 2). The majority of the myxobacterial sterol producers were acquired

213   from soil environments while two other myxobacterial strains originated from marine

214   ecosystems. The Methylococcales species were enriched from a diverse set of ecological settings

215   including sewage sludge, marine water columns, hot springs, freshwater lake sediments and soda

216   lake sediments. Several of the other organisms with Osc homologs in their genomes are also

217   from marine and freshwater environments with *M. caenitepidi* originating from a marine

218   hydrothermal vent, *E. adriatica* from coastal sediments of the Adriatic Sea and *F. taffensis* from

219   sediments of the River Taff.

220   **Distribution of oxidosqualene sequences in metagenomes**

221   To better understand the ecological distribution of sterol-producing bacteria, we

222   performed a BLASTP search of the environmental metagenomics database on the JGI/ IMG

223   website again utilizing the *M. capsulatus* Bath Osc protein as the query sequence. A total of 176

224   Osc metagenome sequences were identified with an e-value of e^{-50} or lower. The majority of the

225   metagenomic sequences were from soil, marine or freshwater environments similar to the

10

226   distribution of isolate environments described above (Figure 4). In addition, Osc sequences were

227   found in metagenomes from estuarine microbial mats, hydrothermal vent fluids and two

228   sequences from sponge symbionts.

229   To distinguish bacterial from eukaryotic Osc sequences in the metagenomes, we

230   generated a maximum likelihood phylogenetic tree of metagenomic and genomic Osc homologs.

231   Because many of the sequences retrieved from metagenomic samples were truncated, we

232   selected the 67 of the 176 metagenomic sequences that were at least 400 amino acids (Osc

233   proteins are generally about 600-650 amino acids) to generate a reliable alignment. After

234   reducing redundancy in the alignment we generated a phylogenetic tree that included a total of

235   55 metagenomic sequences, 65 eukaryotic genomic sequences and 25 bacterial genomic

236   sequences as well as 18 bacterial squalene-hopene cyclase sequences as the outgroup.

237   Thirty-seven of the Osc metagenomics sequences retrieved clustered within the two

238   bacterial Osc clades (Figure 5). Some of these sequences grouped with known sterol producers

239   like the Methylococcales and Myxococcales. However, some of these sequences formed their

240   own clades within the bacterial groups or clustered with organisms that have yet to be shown to

241   produce sterols. Thus, identification of these bacterial Osc homologs in metagenome datasets

242   indicates that there are novel sterol-producing bacteria yet to be discovered and that the sterol

243   producing bacteria inhabit diverse environments. Our bioinformatics analysis of metagenomic

244   databases did identify 18 eukaryotic Osc sequences (Figure 6) which were related to algal, plant

245   or fungal Osc homologs. Given how widespread sterol synthesis is in eukaryotes, we had

246   expected to detect more eukaryotic Osc sequences than bacterial sequences in metagenomic

247   databases. However, it has been documented that metagenomic sequencing tends to recover few

248   eukaryotic sequences in general (Lindahl and Kuske, 2013). Therefore, the low number of

11

249 eukaryotic metagenomic sequences more likely reflects a limited number of eukaryotic

250 sequences in metagenome databases rather than the true prevalence of eukaryotic sterol

251 producers in the environment.

Lipid analysis of potential sterol producers

253 Our identification of Osc homologs in bacterial genomes demonstrates that the potential

254 for sterol synthesis exist in a variety of bacteria. However, the majority of these potential sterol-

255 producing bacterial strains have not been tested for sterol production. In addition, the occurrence

256 of Osc in a genome only suggests the production of the most basic sterols, lanosterol or

257 cycloartenol. Thus, lipid analysis is needed not just to verify sterol production but also to

258 determine if and how sterols are modified in bacteria. We performed lipid analysis on 11 Osc-

259 containing bacteria that included five myxobacteria, four Methylococcales, one Bacteriodetes

260 and one α-Proteobacterium (Table 1). In addition, we searched the genomes of these 11

261 organisms for other sterol biosynthesis protein homologs. Our goal was to link the occurrence of

262 these downstream biosynthesis genes with any sterol modifications, such as saturations and

263 demethylations, these bacteria may be carrying out.

Sterol production in the myxobacteria

265 Four of the five myxobacterial strains tested were found to produce sterols (Table 3). *C.

266 fuscus* strains were previously reported to produce either lanosterol or cycloartenol (Bode et al.,

267 2003) and the *C. fuscus* strain we analyzed produced cycloartenol. We identified homologs for

268 C-14 demethylation and C-24 reduction in the *C. fuscus* genome but did not observe any sterols

269 with these modifications (Table 4). The other three myxobacteria, *E. salina*, *P. pacifica* and *S.

270 amylolyticus* all produced lanosterol rather than cycloartenol and all three strains modified

271 lanosterol to generate zymosterol (cholesta-8,24-dien-3-ol) (Figure 1 and Figure 7). The

12

272   conversion of lanosterol to zymosterol requires demethylation at C-4 and C-14 and a reduction at

273   C-14. A previous study had identified homologs for these biosynthetic steps in the *P. pacifica*

274   genome (Desmond and Gribaldo, 2009) and we also observe these protein homologs in *E. salina*

275   and *S. amylolyticus*. However, C-4 demethylation requires three proteins in yeast and vertebrates

276   (ERG25, ERG26 and ERG27) and we do not observe homologs to all of these proteins in these

277   three strains. Desmond et al. pointed out that *P. pacifica* did not have a homolog for one of the

278   three proteins required for C-4 demethylation (ERG27) and we demonstrate that *S. amylolyticus*

279   is also missing this protein. Further, we were only able to identify one homolog of these three

280   proteins in *E. salina* (Table 4). Thus, it is unclear how these myxobacterial strains are fully

281   demethylating at the C-4 position.

282   In agreement with a previous study, the myxobacterium *C. coralloides* produces

283   significant amounts of squalene but no sterol-like molecules despite having a copy of both

284   squalene epoxidase (SE), required for the conversion of squalene to oxidosqualene prior to

285   cyclization, and oxidosqualene cyclase in its genome (Table 4 and Figure 7) (Bode et al., 2003).

286   To determine if the *C. coralloides* SE and Osc proteins were missing any necessary functional

287   residues, we constructed an alignment of a subset of the bacterial SE and Osc homologs with

288   four eukaryotic SE and Osc sequences (Figure 8). Both of these alignments indicate that key

289   functional amino acid positions in the *C. coralloides* SE and Osc proteins are conserved and so

290   the proteins are likely to be functional. It is also possible that the lack of sterol production may

291   be due a lack of expression under the specific laboratory growth conditions we tested. Current

292   studies are focused on growing *C. coralloides* under various conditions to induce sterol synthesis

293   as well expressing the *C. coralloides* SE and Osc homologs in a heterologous system to verify

294   that these proteins are functional.

Sterol production in the methanotrophs

295

296   The lipid profiles of the four Methylococcales species tested were similar to what was

297   previously observed in *M. capsulatus* Bath (Volkman, 2005), with some exceptions (Figure 9).

298   *M. lacus* did not saturate the sterol side chain at C-24 as would be predicted because it lacks a

299   homolog of the C-24(28) sterol reductase (ERG4 in yeast or DHCR24 in humans) (Table 4). *M.*

300   *luteus*, on the other hand, only produced sterols that were saturated at the C-24 position (Figure

301   9). Interestingly, while all of the Methylococcales tested produced sterols that were partially

302   demethylated at the C-4 position, none had homologs of any of the eukaryotic C-4 demethylase

303   genes (Table 4). These methanotrophs also had sterols in which the unsaturation generated

304   during C-14 demethylation was subsequently removed even though they lack a homolog of the

305   C-14 reductase (ERG24) (Table 4). This is in contrast to the previously tested *M. alcaliphilum*

306   (Banta et al., 2015) which does have a homolog of the C-14 reductase (ERG24) indicating that

307   there may be more than one mechanism for this reaction even within the Methylococcales.

Sterol production in other bacterial species

308

309   We also observed production of cycloartenol in one Bacteriodetes species, *F. taffensis*,

310   and one α-Protebacterium, *M. caenitepidi* (Figure 10 and Table 3). Neither of these strains had

311   homologs of sterol biosynthesis genes downstream of *osc* in their genomes and this was in

312   agreement with our observations of only cycloartenol production (Table 4).

Cycloartenol versus lanosterol synthesis is likely correlated with a single residue

313

314   The production of cycloartenol by some strains in our survey and lanosterol by others

315   prompted us to investigate if specific residues were indicative of whether a cyclase was a

316   lanosterol or cycloartenol synthase. Site-directed mutagenesis studies have previously identified

317   three amino acids changes that seem to control the product profile of oxidosqualene cyclases

14

318 (Meyer et al., 2000;Meyer et al., 2002;Lodeiro et al., 2004). Specifically, the amino acid residues

319 T381/C,Q449/V453 (numbering based on human Osc) were indicative of a lanosterol synthase

320 while Y381/H449/I453 suggested a cycloartenol synthase (Summons et al., 2006). Comparative

321 genomics of three bacterial cyclases with eukaryotic cyclases revealed that only one of these

322 residues was conserved and suggested that a valine (V) or isoleucine (I) at residue 453 suggested

323 lanosterol or cycloartenol production, respectively (Summons et al., 2006). Our lipid analyses

324 and alignments (Figure 8) verify that the bacterial oxidosqualene cyclases in the organisms we

325 tested completely correlated with the observation that a V453 was indicative of lanosterol

326 production while I453 signified cycloartenol production.

327 **Discussion**

328   Sterol biosynthesis is primarily viewed as a eukaryotic feature that is rarely observed in

329 the bacterial domain. Here, we coupled bioinformatics with lipid analyses to demonstrate that

330 sterol production occurs in diverse bacteria and that this pathway may exist in yet to be

331 discovered bacterial species. Our phylogenetic analysis of one of the key proteins involved in

332 sterol biosynthesis, the oxidosqualene cyclase (Osc), demonstrates that the evolutionary history

333 of this pathway in the bacterial domain is complex. In two previous phylogenomic studies it was

334 concluded that sterol biosynthesis in the bacterial domain was most likely acquired through

335 horizontal gene transfer (Desmond and Gribaldo, 2009;Frickey and Kannenberg, 2009).

336 However, these phylogenetic analyses were limited as only three bacterial Osc sequences were

337 available at the time these studies were undertaken. Our phylogenetic reconstruction with a

338 larger data set demonstrates that bacterial Osc homologs fall into two clades – one monophyletic

339 with the eukaryotic Osc sequences (Group 2) and one forming a sister clade to the eukaryotic

340 sequences (Group 1). This topology implies that Group 2 *osc* genes were most likely acquired

341   via horizontal transfer from a eukaryote as previously proposed (Frickey and Kannenberg, 2009)

342   while Group 1 *osc* genes are more divergent from those found in eukaryotes. It seems plausible

343   that Group 1 Osc sequences may represent a primitive lineage of cyclases suggesting that these

344   bacteria may contain a more ancestral sterol biosynthetic pathway. However, more bacterial and

345   eukaryotic Osc sequences as well as more rigorous phylogenetic analyses are needed to better

346   interpret the evolutionary history of sterol biosynthesis in both the bacterial and eukaryotic

347   domains.

348   While our phylogenetic analyses suggest a complex evolutionary history of sterol

349   biosynthesis in bacteria, our lipid analyses demonstrate less modification of sterols in bacteria

350   compared to what is usually observed in eukaryotes. Many of the bacteria we tested produced

351   lanosterol or cycloartenol as the end product. Production of these basic sterols only requires two

352   biosynthetic steps of the canonical eukaryotic sterol pathway - the epoxidation of squalene to

353   oxidosqualene and the subsequent cyclization of oxidosqualene to lanosterol or cycloartenol. The

354   myxobacteria and methanotrophs, however, did make certain modifications such as C-4 and C-

355   14 demethylations and isomerization of double bonds in the main ring structure. Interestingly,

356   not all proteins required to make those modification in eukaryotes were found in the genomes of

357   these bacteria. In particular, the removal of the C-4 methyl groups requires the activity of three

358   eukaryotic proteins, a C-4 methyl oxidase (ERG25), a C-4 decarboxylase (ERG26) and a C-3

359   ketoreducatse (ERG27) (Bard et al., 1996;Gachotte et al., 1998;Gachotte et al., 1999). These

360   three proteins were first identified in yeast and homologs have been identified in most sterol

361   producing eukaryotic genomes, with the exception of plants which seem to be missing an ERG27

362   homolog (Desmond and Gribaldo, 2009). In the myxobacteria, we observed that two of the

363   organisms that removed the C-4 methyl groups had homologs of ERG25 and ERG26 but not

16

364   ERG27 and a third organism only had a homolog of ERG25. Desmond and Gribaldo attempted

365   to identify potential ERG27 homologs in the genome of *P. pacifica* through comparative

366   genomics. One potential gene candidate was identified (*P. pacifica* locus tag:

367   Ga0067453_11974) and the myxobacteria we tested do have a homolog of this protein in their

368   genomes. However, further studies are needed to determine if this protein is necessary for C-4

369   demethylation in the myxobacteria.

370   It is also possible that downstream sterol modifications in bacteria occur via distinct

371   biochemical pathways than what is observed in eukaryotes. This is a particularly compelling in

372   the aerobic methanotrophs. In these organisms, one methyl group is removed at the C-4 position

373   but we could not identify homologs of the eukaryotic C-4 demethylase genes (ERG25, ERG26 or

374   ERG27). In addition, we observed saturation of the C-14 double bond in the sterols of all

375   methanotrophs tested but did not identify a C-14 reductase (ERG24) in their genomes. The

376   discrepancies in the sterols produced by methanotrophs and the proteins identified in their

377   genomes points to the possibility that novel sterol biosynthesis proteins may exist in bacteria.

378   Identification and characterization of these bacterial sterol proteins could reveal unique

379   biochemical and regulatory mechanisms. In addition, a full understanding of the proteins

380   involved in bacterial sterol production will allow for studies to discern what functional role these

381   lipids play in the bacterial cell and would provide significant insight into the evolution of this

382   ancient biosynthetic pathway.

383   Finally, this work demonstrates the utility of combining bioinformatics with lipid

384   analyses to get a broader picture of not just sterol synthesis in bacteria but potentially other

385   geologically relevant lipids. The increasing amount of genome and metagenome sequence data

386   available along with advancements in developing culturing and genetic systems in nontraditional

387   microbes provides an excellent opportunity for exploring many aspects of biomarker lipids in

388   microbes – their biosynthesis, their function and their evolutionary history. Ultimately, a full

389   understanding of microbial biomarker lipids will provide valuable information for more nuanced

390   interpretation of microbial lipid biosignatures in both modern ecosystems and ancient

391   sedimentary rocks.

392   **Author contributions**

393   PW designed the study. JW, XY and PW acquired the data. JW and PW analyzed the data. PW

394   wrote the manuscript.

395   **Funding**

398   **Acknowledgments**

403

References

Alain, K., Intertaglia, L., Catala, P., and Lebaron, P. (2008). *Eudoraea adriatica* gen. nov., sp. nov., a novel marine bacterium of the family *Flavobacteriaceae*. *Int. J. Syst. Evol. Microbiol.* 58**,** 2275-2281.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25**,** 3389-3402.

Banta, A.B., Wei, J.H., and Welander, P.V. (2015). A distinct pathway for tetrahymanol synthesis in bacteria. *Proc. Natl. Acad. Sci. U.S.A* 112**,** 13478-13483.

Bard, M., Bruner, D.A., Pierson, C.A., Lees, N.D., Biermann, B., Frye, L., Koegel, C., and Barbuch, R. (1996). Cloning and characterization of ERG25, the *Saccharomyces cerevisiae* gene encoding C-4 sterol methyl oxidase. *Proc. Natl. Acad. Sci. U.S.A* 93**,** 186-190.

Bird, C.W., Lynch, J.M., Pirt, F.J., and Reid, W.W. (1971). Steroids and squalene in *Methylococcus capsulatus* grown on methane. *Nature* 230**,** 473-474.

Bloch, K. (1991). "Cholesterol: evolution of structure and function.," in *Biochemistry of lipids, lipoproteins, and membranes,* eds. D.E. Vance & J. Vance. (Amsterdam: Elsevier), 363-381.

Bode, H.B., Zeggel, B., Silakowski, B., Wenzel, S.C., Reichenbach, H., and Muller, R. (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.

Bouvier, P., Rohmer, M., Benveniste, P., and Ourisson, G. (1976). Δ8(14)-steroids in the bacterium *Methylococcus capsulatus*. *Biochem. J.* 159**,** 267-271.

Bowman, J.P. (2014). "The Family *Methylococcaceae*," in *The Prokaryotes: Gammaproteobacteria,* eds. E. Rosenberg, E.F. Delong, S. Lory, E. Stackenbrandt & F. Thompson. (Berlin, Germany: Springer-Verlag), 411-440.

Bowman, J.P., Sly, L.I., Nichols, P.D., and Hayward, A.C. (1993). Revised taxonomy of the methanotrophs: Description of *Methylobacter* gen. nov., emendation of *Methylococcus*, validation of *Methylosinus* and *Methylocystis* species, and a proposal that the family *Methylococcaceae* includes only the group I methanotrophs. *Int. J. Syst. Bacteriol.* 43**,** 735-753.

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

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

Desmond, E., and Gribaldo, S. (2009). Phylogenomics of sterol synthesis: Insights into the origin, evolution, and diversity of a key eukaryotic feature. *Genome. Biol. Evol.* 2009**,** 364-381.

Dimster-Denk, D., and Rine, J. (1996). Transcriptional regulation of a sterol-biosynthetic enzyme by sterol levels in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 16**,** 3981-3989.

Donia, M.S., Fricke, W.F., Partensky, F., Cox, J., Elshahawi, S.I., White, J.R., Phillippy, A.M., Schatz, M.C., Piel, J., Haygood, M.G., Ravel, J., and Schmidt, E.W. (2011). Complex microbiome underlying secondary and primary metabolism in the tunicate-Prochloron symbiosis. *Proc. Natl. Acad. Sci. U.S.A* 108, E1423-E1432.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792-1797.

Frickey, T., and 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.

Gachotte, D., Barbuch, R., Gaylor, J., Nickel, E., and 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. U.S.A* 95, 13794-13799.

Gachotte, D., Sen, S.E., Eckstein, J., Barbuch, R., Krieger, M., Ray, B.D., and Bard, M. (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.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696-704.

Hamilton, R., Kits, K.D., Ramonovskaya, V.A., Rozova, O.N., Yurimoto, H., Iguchi, H., Khmelenina, V.N., Sakai, Y., Dunfield, P.F., Klotz, M.G., Knief, C., Op Den Camp, H.J., Jetten, M.S., Bringel, F., Vuilleumier, S., Svenning, M.M., Shapiro, N., Woyke, T., Trotsenko, Y.A., Stein, L.Y., and Kalyuzhnaya, M.G. (2015). Draft genomes of gammaproteobacterial methanotrophs isolated from terrestrial ecosystems. *Genome Announc.* 3.

Hartmann, M.A. (1998). Plant sterols and the membrane environment. *Trends Plant Sci.* 3, 170-175.

Huntley, S., Hamann, N., Wegener-Feldbrugge, S., Treuner-Lange, A., Kube, M., Reinhardt, R., Klages, S., Muller, R., Ronning, C.M., Nierman, W.C., and Sogaard-Andersen, L. (2011). Comparative genomic analysis of fruiting body formation in Myxococcales. *Mol. Biol. Evol.* 28, 1083-1097.

Huntley, S., Zhang, Y., Treuner-Lange, A., Kneip, S., Sensen, C.W., and Sogaard-Andersen, L. (2012). Complete genome sequence of the fruiting myxobacterium *Corallococcus corralloides* DSM 2259. *J. Bacteriol.* 194, 3012-3013.

Iizuka, T., Jojima, Y., Fudou, R., Hiraishi, A., Ahn, J.W., and Yamanaka, S. (2003a). *Plesiocystis pacifica* gen. nov., sp. nov., a marine myxobacterium that contains dihydrogenated menaquinone, isolated from the Pacific coasts of Japan. *Int. J. Syst. Evol. Microbiol.* 53, 189-195.

Iizuka, T., Jojima, Y., Fudou, R., Tokura, M., Hiraishi, A., and Yamanaka, S. (2003b). *Enhygromyxa salina* gen. nov., sp. nov., a slightly halophilic myxobacterium isolated from the coastal areas of Japan. *Syst. Appl. Microbiol.* 26, 189-196.

Kaluzhnaya, M., Khmelenina, V., Eshinimaev, B., Suzina, N., Nikitin, D., Solonin, A., Lin, J.L., Mcdonald, I., Murrell, C., and Trotsenko, Y. (2001). Taxonomic characterization of new alkaliphilic and alkalitolerant methanotrophs from soda lakes of the Southeastern Transbaikal region and description of Methylomicrobium buryatense sp.nov. *Syst. Appl. Microbiol.* 24, 166-176.

494 Kalyuzhnaya, M.G., Khmelenina, V., Eshinimaev, B., Sorokin, D., Fuse, H., Lidstrom, M., and
495   Trotsenko, Y. (2008). Classification of halo(alkali)philic and halo(alkali)tolerant
496   methanotrophs provisionally assigned to the genera *Methylomicrobium* and
497   *Methylobacter* and emended description of the genus *Methylomicrobium*. *Int. J. Syst.*
498   *Evol. Microbiol.* 58**,** 591-596.
499 Kalyuzhnaya, M.G., Stolyar, S.M., Auman, A.J., Lara, J.C., Lidstrom, M.E., and Chistoserdova,
500   L. (2005). *Methylosarcina lacus* sp. nov., a methanotroph from Lake Washington,
501   Seattle, USA, and emended description of the genus *Methylosarcina*. *Int. J. Syst. Evol.*
502   *Microbiol.* 55**,** 2345-2350.
503 Khmelenina, V.N., Kalyuzhnaya, M.G., Starostina, N.G., Suzina, N.E., and Trotsenko, Y.A.
504   (1997). Isolation and characterization of halotolerant alkaliphilic methanotrophic bacteria
505   from Tuva Soda Lakes. *Curr. Microbiol.* 35**,** 257-261.
506 Kleiveland, C.R., Hult, L.T., Kuczkowska, K., Jacobsen, M., Lea, T., and Pope, P.B. (2012).
507   Draft genome sequence of the methane-oxidizing bacterium *Methylococcus capsulatus*
508   (Texas). *J. Bacteriol.* 194**,** 6626.
509 Lees, N.D., Skaggs, B., Kirsch, D.R., and Bard, M. (1995). Cloning of the late genes in the
510   ergosterol biosynthetic pathway of Saccharomyces cerevisiae--a review. *Lipids* 30**,** 221-
511   226.
512 Letunic, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic
513   tree display and annotation. *Bioinformatics* 23**,** 127-128.
514 Letunic, I., and Bork, P. (2011). Interactive Tree Of Life v2: online annotation and display of
515   phylogenetic trees made easy. *Nucleic Acids Res.* 39**,** 475-478.
516 Lindahl, B.D., and Kuske, C.R. (2013). "Metagenomics for Study of Fungal Ecology," in *The*
517   *Ecological Genomics of Fungi,* ed. F. Martin. (Hoboken, NJ: John Wiley & Sons, Inc).
518 Lodeiro, S., Segura, M.J., Stahl, M., Schulz-Gasch, T., and Matsuda, S.P. (2004). Oxidosqualene
519   cyclase second-sphere residues profoundly influence the product profile. *Chembiochem*
520   5**,** 1581-1585.
521 Love, G.D., Grosjean, E., Stalvies, C., Fike, D.A., Grotzinger, J.P., Bradley, A.S., Kelly, A.E.,
522   Bhatia, M., Meredith, W., Snape, C.E., Bowring, S.A., Condon, D.J., and Summons, R.E.
523   (2009). Fossil steroids record the appearance of Demospongiae during the Cryogenian
524   period. *Nature* 457**,** 718-721.
525 Mccurdy, H.D. (1969). Studies on the taxonomy of the Myxobacterales. I. Record of Canadian
526   isolates and survey of methods. *Can. J. Microbiol.* 15**,** 1453-1461.
527 Meyer, M.M., Segura, M.J., Wilson, W.K., and Matsuda, S.P. (2000). Oxidosqualene Cyclase
528   Residues that Promote Formation of Cycloartenol, Lanosterol, and Parkeol. *Angew.*
529   *Chem. Int. Ed. Engl.* 39**,** 4090-4092.
530 Meyer, M.M., Xu, R., and Matsuda, S.P. (2002). Directed evolution to generate cycloartenol
531   synthase mutants that produce lanosterol. *Org. Lett.* 4**,** 1395-1398.
532 Mohr, K.I., Garcia, R.O., Gerth, K., Irschik, H., and Muller, R. (2012). *Sandaracinus*
533   *amylolyticus* gen. nov., sp. nov., a starch-degrading soil myxobacterium, and description
534   of Sandaracinaceae fam. nov. *Int. J. Syst. Evol. Microbiol.* 62**,** 1191-1198.
535 Nes, W.D. (2011). Biosynthesis of cholesterol and other sterols. *Chem. Rev.* 111**,** 6423-6451.
536 O'sullivan, L.A., Rinna, J., Humphreys, G., Weightman, A.J., and Fry, J.C. (2005). *Fluviicola*
537   *taffensis* gen. nov., sp. nov., a novel freshwater bacterium of the family *Cryomorphaceae*
538   in the phylum *'Bacteroidetes'*. *Int. J. Syst. Evol. Microbiol.* 55**,** 2189-2194.

539   Pearson, A., Budin, M., and Brocks, J.J. (2003). Phylogenetic and biochemical evidence for
540   sterol synthesis in the bacterium *Gemmata obscuriglobus*. *Proc. Natl. Acad. Sci. U.S.A*
541   100**,** 15352-15357.
542   Peters, K.E., Waters, C.C., and Moldowan, J.M. (2007a). *The Biomarker Guide Volume 1:*
543   *Biomarkers and Isotopes in the Environment and Human History.* Cambridge University
544   Press.
545   Peters, K.E., Waters, C.C., and Moldowan, J.M. (2007b). *The Biomarker Guide Volume 2:*
546   *Biomarkers and Isotopes in Petroleum Exploration and Earth History.* Cambridge
547   University Press.
548   Reichenbach, H. (1970). *Nannocystis exedens* gen. nov., spec. nov., a new myxobacterium of the
549   family *Sorangiaceae*. *Arch Mikrobiol* 70**,** 119-138.
550   Riobo, N.A. (2012). Cholesterol and its derivatives in Sonic Hedgehog signaling and cancer.
551   *Curr. Opin. Pharmacol.* 12**,** 736-741.
552   Schouten, S., Bowman, J.P., Rijpstra, W.I., and Sinninghe Damste, J.S. (2000). Sterols in a
553   psychrophilic methanotroph, *Methylosphaera hansonii*. *FEMS Microbiol. Lett.* 186**,** 193-
554   195.
555   Sessions, A.L., Zhang, L.C., Welander, P.V., Doughty, D., Summons, R.E., and Newman, D.K.
556   (2013). Identification and quantification of polyfunctionalized hopanoids by high
557   temperature gas chromatography-mass spectrometry. *Org. Geochem.* 56**,** 120-130.
558   Siedenburg, G., and Jendrossek, D. (2011). Squalene-hopene cyclases. *Appl. Environ. Microbiol.*
559   77**,** 3905-3915.
560   Strand, S.E., and Lidstrom, M.E. (1984). Characterization of a new marine methylotroph. *FEMS*
561   *Microbiol. Lett.* 21**,** 247-251.
562   Stratmann, K., Moore, R.E., Bonjouklian, R., Deeter, J.B., Patterson, G.M.L., Shaffer, S., Smith,
563   C.D., and Smitka, T.A. (1994). Welwitindolinones, Unusual Alkaloids from the Blue-
564   Green Algae *Hapalosiphon welwitschii* and *Westiella intricata* - Relationship to
565   Fischerindoles and Hapalindoles. *J. Am. Chem. Soc.* 116**,** 9935-9942.
566   Summons, R.E., Bradley, A.S., Jahnke, L.L., and Waldbauer, J.R. (2006). Steroids, triterpenoids
567   and molecular oxygen. *Philos. Trans. R. Soc. Lond.* 361**,** 951-968.
568   Swan, T.M., and Watson, K. (1998). Stress tolerance in a yeast sterol auxotroph: role of
569   ergosterol, heat shock proteins and trehalose. *FEMS Microbiol. Lett.* 169**,** 191-197.
570   Takeuchi, M., Katayama, T., Yamagishi, T., Hanada, S., Tamaki, H., Kamagata, Y., Oshima, K.,
571   Hattori, M., Marumo, K., Nedachi, M., Maeda, H., Suwa, Y., and Sakata, S. (2014).
572   *Methyloceanibacter caenitepidi* gen. nov., sp. nov., a facultatively methylotrophic
573   bacterium isolated from marine sediments near a hydrothermal vent. *Int. J. Syst. Evol.*
574   *Microbiol.* 64**,** 462-468.
575   Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent
576   and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56**,** 564-
577   577.
578   Villanueva, L., Rijpstra, W.I., Schouten, S., and Damste, J.S. (2014). Genetic biomarkers of the
579   sterol-biosynthetic pathway in microalgae. *Environ. Microbiol. Rep.* 6**,** 35-44.
580   Volkman, J.K. (2003). Sterols in microorganisms. *Appl. Microbiol. Biotechnol.* 60**,** 495-506.
581   Volkman, J.K. (2005). Sterols and other triterpenoids: source specificity and evolution of
582   biosynthetic pathways. *Org. Geochem.* 36**,** 139-159.

22

583   Welander, P.V., and Summons, R.E. (2012). Discovery, taxonomic distribution, and phenotypic
584   characterization of a gene required for 3-methylhopanoid production. *Proc. Natl. Acad.*
585   *Sci. U.S.A.* 109**,** 12905-12910.
586   Whittenbury, R., Phillips, K.C., and Wilkinson, J.F. (1970). Enrichment, isolation and some
587   properties of methane-utilizing bacteria. *J. Gen. Microbiol.* 61**,** 205-218.
588   Xu, F., Rychnovsky, S.D., Belani, J.D., Hobbs, H.H., Cohen, J.C., and Rawson, R.B. (2005).
589   Dual roles for cholesterol in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 102**,** 14551-
590   14556.
591   Yamamoto, E., Muramatsu, H., and Nagai, K. (2014). *Vulgatibacter incomptus* gen. nov., sp.
592   nov. and *Labilithrix luteola* gen. nov., sp. nov., two myxobacteria isolated from soil in
593   Yakushima Island, and the description of Vulgatibacteraceae fam. nov., Labilitrichaceae
594   fam. nov. and Anaeromyxobacteraceae fam. nov. *Int. J. Syst. Evol. Microbiol.* 64**,** 3360-
595   3368.
596   Yang, H. (2006). Nonvesicular sterol transport: two protein families and a sterol sensor? *Trends*
597   *Cell Biol.* 16**,** 427-432.

598

599

600

Table 1. Bacterial strains tested for sterol biosynthesis

Bacteria surveyed in this study	Growth conditions	Source
Corallococcus coralloides DSM 2259	DSMZ Medium 222 at 30°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	DSMZ
Cystobacter fuscus DSM 2262	DSMZ Medium 222 at 30°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	DSMZ
Enhygromyxa salina DSM 15201	DSMZ Medium 958 on agar plates at 30°C ; lipid analysis from cells scraped off agar plates	DSMZ
Fluviicola taffensis RW262 DSM 16823	DSMZ Medium 1 for liquid at 30°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	DSMZ
Methylobacter luteus IMV-B-3098	NMS Medium (Welander and Summons, 2012) plus methane at 5 psi over ambient at 30°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	M.G. Kalyuzhnaya, San Diego State University
Methylobacter whittenburyi ACM 3310	NMS Medium (Welander and Summons, 2012) plus methane at 5 psi over ambient at 30°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	M.G. Kalyuzhnaya, San Diego State University
Methyloceanibacter caenitepidi DSM 27242	DSMZ Medium 1488 plus 1% methanol at 37°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	M.G. Kalyuzhnaya, San Diego State University
Methylococcus capsulatus Texas ATCC 19069	NMS Medium (Welander and Summons, 2012) plus methane at 5 psi over ambient at 30°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	ATCC
Methylosarcina lacus LW14	NMS Medium (Welander and Summons, 2012) plus methane at 5 psi over ambient at 30°C with shaking at 200 RPM; lipid analysis from 50 ml of a stationary phase culture	M.G. Kalyuzhnaya, San Diego State University
Plesiocystis pacifica SIR-1 DSM 14875	DSMZ Medium 958 on agar plates at 30°C; lipid analysis from cells scraped off agar plates	DSMZ
Sandaracinus amylolyticus DSM 53668	DSMZ Medium 67 on agar plates at 30°C; lipid analysis from cells scraped off agar plates	DSMZ

601

602

24

Table 2. Bacterial genomes that contain oxidosqualene cyclase homologs

Genome	Locus Tag	Isolation/Environment
δ-Proteobacteria (myxobacteria)		
Corallococcus coralloides DSM 2259	COCOR_01777	Soil, Canada (Huntley et al., 2012)
Cystobacter fuscus DSM 2262	D187_003104	Soil, Canada (McCurdy, 1969)
Enhygromyxa salina DSM 15201	Ga0055550_114516	Marine, intertidal zone (Iizuka et al., 2003b)
Labilithrix luteola DSM 27648	AKJ09_09404	Soil, Yakushima Island (Yamamoto et al., 2014)
Nannocystis exedens ATCC 25963	Ga0008035_02275	Soil, Desert (Reichenbach, 1970)
Plesiocystis pacifica SIR-1	PPSIR1_04883	Marine, intertidal zone (Iizuka et al., 2003a)
Sandaracinus amylolyticus DSM 53668	Ga0055546_18051	Soil, India (Mohr et al., 2012)
Stigmatella aurantiaca DW4/3-1	STAUR_5418	Soil (Huntley et al., 2011)
γ-Proteobacteria (Methylococcales)		
Methylobacter luteus IMV-B-3098	MetluDRAFT_1255	Sewage (Bowman et al., 1993)
Methylobacter sp. BBA5.1	EK22DRAFT_03359	Environmental, JGI
Methylobacter whittenburyi ACM 3310	GY38DRAFT_3867	Terrestrial (Hamilton et al., 2015)
Methylobacterium marinus A45	MetmaDRAFT_3943	Marine water column, Framvaren Fjord, Norway (Strand and Lidstrom, 1984)
Methylocaldum sp. 175	JC06DRAFT_3873	Environmental, JGI
Methylocaldum szegediense O-12	MetszDRAFT_3769	Hot spring (Bowman, 2014)
Methylococcaceae sp. 73a	EK23DRAFT_02566	Environmental, JGI
Methylococcus capsulatus Bath	MCA2873	Hot Spring (Whittenbury et al., 1970)
Methylococcus capsulatus Texas (ATCC 19069)	H156DRAFT_2530	Sewer sludge (Kleiveland et al., 2012)
Methylomicrobium alcaliphilum 20Z	MEALZ_0768	Sediment, Tuva soda lakes (Khmelenina et al., 1997)
Methylomicrobium buryatense 5G	METBUDRAFT_4052	Sediment, Transbaikal soda lakes (Kaluzhnaya et al., 2001)
Methylomicrobium kenyense AMO1	IQ34DRAFT_3157	Surface sediment Kenyan soda lake (Kalyuzhnaya et al., 2008)
Methylosarcina lacus LW14	MetlaDRAFT_0845	Sediment, Lake Washington (Kalyuzhnaya et al., 2005)
α-Proteobacteria (Rhizobiales)		
Methyloceanibacter caenitepidi str. Gela4	Ga0077927_11952	Marine sediment hydrothermal vent (Takeuchi et al., 2014)
Bacteriodetes		
Eudoraea adriatica DSM 19308	G504DRAFT_2316	Coastal waters, Adriatic Sea (Alain et al., 2008)
Eudoraea sp. SCGC 5358	K485DRAFT_00205	Marine, North Sea, JGI
Eudoraea sp. SCGC 5441	K506DRAFT_00376	Marine, North Sea, JGI
Eudoraea sp. SCGC 5444	K507DRAFT_00676	Marine, North Sea, JGI

25

Fluviicola taffensis RW262 DSM 16823	Fluta_3214	Freshwater, River Taff, (O'Sullivan et al., 2005)
Cyanobacteria		
Planktothrix sp. st147	st147_cleanDRAFT_00043800	Freshwater, Lake Langersee
Prochloron didemni P2-Fiji	Ga0040003_100917	Tunicate symbiont (*Lissoclinum patella*) (Donia et al., 2011)
Prochloron didemni P3-Solomon	Ga0040004_00276	Tunicate symbiont (*Lissoclinum patella*) (Donia et al., 2011)
Prochloron didemni P4-Papua New Guinea	Ga0040005_03818	Tunicate symbiont (*Lissoclinum patella*) (Donia et al., 2011)
Westiella intricata UH HT-29-1	HT291_04925	Soil, Moen Island (Stratmann et al., 1994)
Planctomycete		
Gemmata obscuriglobus UQM 2246	GobsU_010100006605	Freshwater dam, Queensland, Australia
Gemmata sp. IIL30	Ga0036985_07173	Environmental, JGI

603

604

Table 3. Sterols identified in bacterial strains

	squalene	cycloartenol	lanosterol	parkeol	4,4-dimethylcholesta-8,24-dien-3-ol	4,4-dimethylcholesta-8-en-3-ol	4-methylcholesta-8,24-dien-3-ol	4-methylcholesta-8-en-3-ol	zymosterol
Myxobacteria									
Corallococcus coralloides	+								
Cystobacter fuscus	+	+							
Enhygromyxa salina			+						+
Plesiocystis pacifica			+						+
Sandaracinus amylolyticus			+						+
Methylococcales									
Methylobacter luteus						+		+	
Methylobacter whittenburyi					+	+	+	+	
Methylococcus capsulatus Texas					+	+	+	+	
Methylosarcina lacus					+		+		
Bacteriodetes									
Fluviicola taffensis		+							
α-Proteobacteria									
Methyloceanibacter caenitepidi		+							
Strains tested in previous studies									
Nannocystis species (Bode et al., 2003)	+		+		+	+	+	+	+
Stigmatella aurantica (Bode et al., 2003)	+	+							
Cystobacter species (Bode et al., 2003)	+	+	+						
Corallococcus species (Bode et al., 2003)	+								
Gemmata obscuriglobus (Pearson et al., 2003)			+	+					
Methylococcus capsulatus Bath (Bird et al., 1971;Bouvier et al., 1976)					+	+	+	+	
Methylomicrobium alcaliphilum 20Z (Banta et al., 2015)						+			

605

606

Table 4. Identification of sterol biosynthesis genes in bacterial genomes

| | squalene oxygenation SE | oxidosqualene cyclization - Las (lanosterol) | oxidosqualene cyclization - Cas (cycloartenol) | C-14 demethylation ERG11 | C-14 reduction ERG24 | C-4 demethylation | | | C-24(28) reduction ERG4/DHCR24 |
						C-4 methyl oxidation ERG25	C-3 dehydrogenation ERG26	C-3 ketoreduction ERG27	
Strains tested in this study									
Corallococcus coralloides	COCOR_01775		COCOR_01777	COCOR_02429					COCOR_02182
Cystobacter fuscus	D187_003106		D187_003104	D187_005870		D187_005391			D187_001522
Enhygromyxa salina	Ga0055550_102654	Ga0055550_114516		Ga0055550_103211	Ga0055550_105926	Ga0055550_101078			Ga0055550_101115
Eudoraea adriatica	G504DRAF_2315	G504DRAFT_2316							
Fluviicola taffensis	Fluta_3221		Fluta_3214						
Methylobacter luteus	MetluDRAFT_1256	MetluDRAFT_1255		MetluDRAFT_1253					MetluDRAFT_1263
Methylobacter whittenburyi	GY38DRAFT_3868	GY38DRAFT_3867		GY38DRAFT_3865					GY38DRAFT_3875
Methyloceanibacter caenitepidi	GL4_3111		GL4_3110						
Methylococcus capsulatus Texas	H156DRAF_2531	H156DRAFT_2530		H156DRAFT_1746					H156DRAF_0889
Methylosarcina lacus	MetlaDRAFT_0846	MetlaDRAFT_0845		MetlaDRAFT_0843					
Plesiocystis pacifica	PPSIR1_02838	PPSIR1_02843		PPSIR1_36894	PPSIR1_23244	PPSIR1_14275	PPSIR1_17435		
Sandaracinus amylolyticus	Ga0055546_12562	Ga0055546_18051		Ga0055546_101121	Ga0055546_17127	Ga0055546_103226	Ga0055546_17162		Ga0055546_16346
Strains tested in previous studies									
Nannocystis excedens	Ga0008035_04851	Ga0008035_04852		Ga0008035_02418	Ga0008035_01127	Ga0008035_06338	Ga0008035_00875		Ga0008035_06674
Stigmatella aurantica	STAUR_5420		STAUR_5418	STAUR_2030		STAUR_2074			
Gemmata obscuriglobus	GobsU_010100006610	GobsU_010100006605							

Methylococcus capsulatus Bath	MCA 2872	MCA 2873		MCA 2711					MCA 1404
Methylomicrobium alcaliphilum 20Z	MEAL Z_ 0767	MEAL Z_ 0768		MEAL Z_ 0770					MEALZ _ 3890

607

608   **Figure Legends**

609

610   **Figure 1.** Sterol biosynthesis in eukaryotes. All sterol biosynthetic pathways begin with the

611   oxidation of squalene to oxidosqualene and subsequent cyclization to lanosterol (vertebrates and

612   fungi) or cycloartenol (plants). Shown are the initial enzymatic steps in the conversion of

613   lanosterol to zymosterol which occurs similarly in vertebrates and fungi. Proteins involved in

614   these steps have been characterized from a variety of eukaryotes and the locus to tags shown are

615   those from *Saccharomyces cerevisiae* (Erg).

616

617   **Figure 2.** Mass spectra of sterols identified in this study. Spectra of the following acetylated

618   sterols: cycloartenol (9,19-cyclolanost-24-en-3-ol), 4,4-dimethylcholesta-8,24-dien-3-ol, 4,4-

619   dimethylcholesta-8-en-3-ol, 4-methylcholesta-8,24-dien-3-ol, 4-methylcholesta-8-en-3-ol.

620   Spectra of the following trimethylsilylated sterols: lanosterol (lanosta-8,24-dien-3-ol) and

621   zymosterol (cholesta-8,24-dien-3-ol).

622

623   **Figure 3.** Maximum likelihood phylogenetic tree of oxidosqualene cyclase protein sequences

624   from bacterial and eukaryotic genomes. Bacterial squalene hopene cyclase (Shc) sequences were

625   used as the outgroup. Eukaryotic Osc and bacterial Shc branches are collapsed for better

626   visualization of the tree. Colored branches represent different bacterial phyla: δ-Proteobacteria

627   (blue), Cyanobacteria (green), Planctomycetes (cyan), γ-Proteobacteria (red), Bacteriodetes

628   (pink), α-Proteobacteria (brown) and Verrucomicrobia (orange).

629

630   **Figure 4.** Distribution of Osc protein sequences in metagenomes. Each bar represents the

631   number of Osc homologs identified in the metagenomes from that ecosystem. The majority of

632   homologs were found in freshwater, soil and marine metagenomes.

633

634   **Figure 5.** Maximum likelihood phylogenetic tree of bacterial genomic and metagenomic Osc

635   protein sequences. Red branches represent bacterial sequences and black collapsed branches are

636   eukaryotic sequences. Blue labels indicate metagenomic sequences and black labels indicate

637   sequences from genomes. Bacterial Shc sequences were used as the outgroup.

638

639   **Figure 6.** Maximum likelihood phylogenetic tree of eukaryotic genomic and metagenomic Osc

640   protein sequences. Red collapsed branches are bacterial sequences and black branches represent

641   eukaryotic sequences. Blue labels indicate metagenomic sequences and black labels indicate

642   sequences from genomes. Eukaryotic branches with no metagenomic Osc sequences were

643   collapsed to reduce the size of the tree. Bacterial Shc sequences were used as the outgroup.

644

645   **Figure 7.** Sterols production in the myxobacteria. Extracted ion chromatograms (m/z 69, 440,

646   442, 454, 456, 468, and 498) of total lipid extract (TLE) from five myxobacteria. *C. coralloides*

647   and *C. fuscus* TLEs were extracted from liquid cultures and were acetylated prior to running on

648   the GC-MS. *E. salina*, *P. pacifica* and *S. amylolyticus* TLEs were extracted from cultures on a

649   plate as growing them in liquid cultures was difficult. These TLEs were trimethylsilylated prior

650   to running on the GC-MS. Sterol peaks were identified based on their mass spectra as shown in
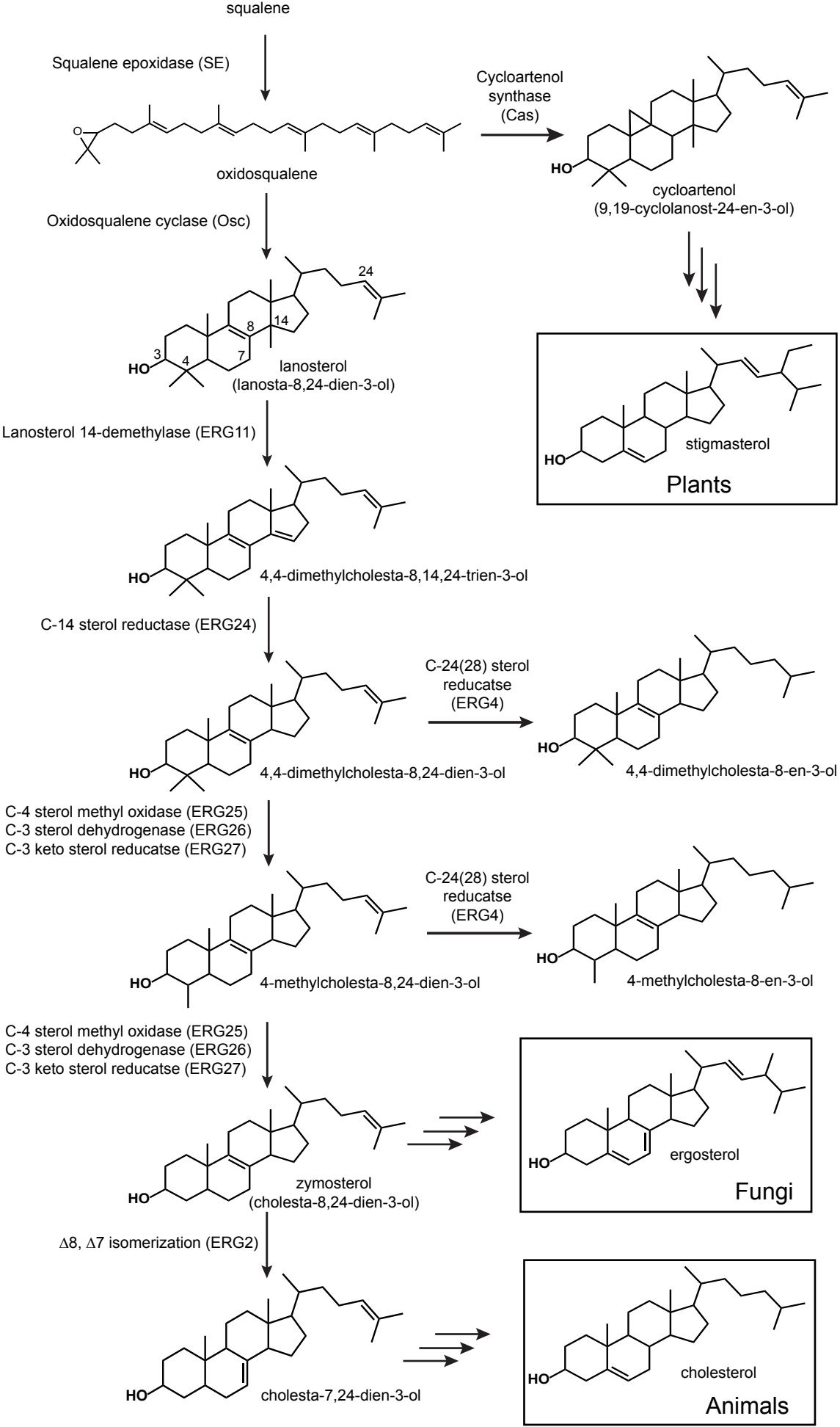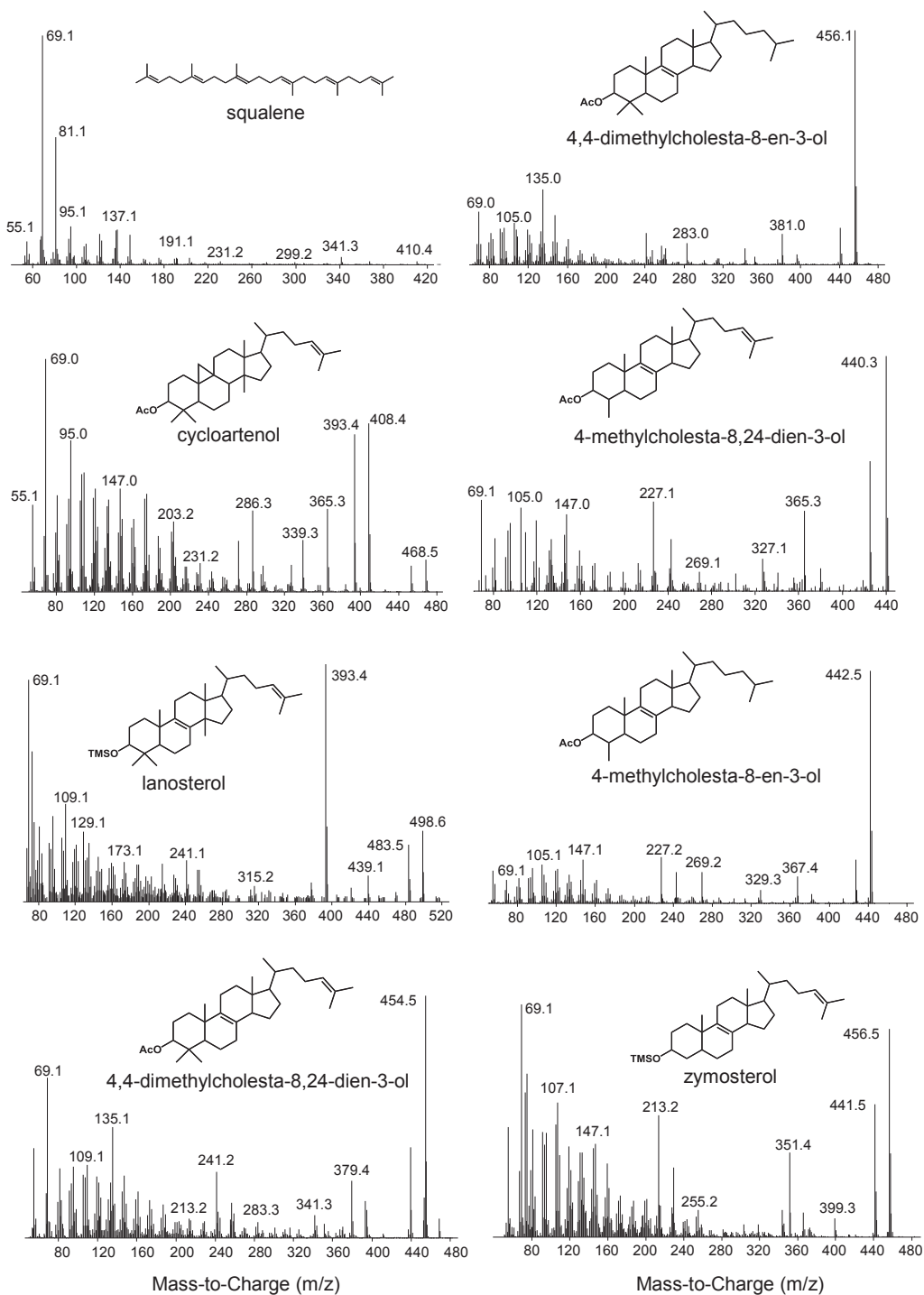
651   Figure 2.

652

653 **Figure 8**. Amino acid alignments of the critical functional domains of oxidosqualene cyclase **(A)**

654 and squalene epoxidase **(B)** homologs. Residues in black indicate residues that have been

655 demonstrated to have a role in the biosynthesis of sterols in eukaryotes. Grey residues are those

656 that differ from the conserved residue. In the Osc alignment, an isoleucine (I) at 453 (yellow)

657 indicates a cycloartenol synthase and a valine (V) at 453 (blue) indicates a lanosterol synthase.

658 Numbers correspond to residues in human Osc and SE. Bold labels indicate bacterial strains

659 tested in this study. #: eukaryotic sequences, *: bacteria that have been shown to produce sterols.

660

661 **Figure 9.** Sterols production in the aerobic methanotrophs. Extracted ion chromatograms (m/z

662 69, 440, 442, 454, 456, 468, and 498) of total lipid extract (TLE) from four aerobic

663 methanotrophs. All TLEs were extracted from liquid cultures and were acetylated prior to

664 running on the GC-MS. Sterol peaks were identified based on their mass spectra as shown in

665 Figure 2.

666

667 **Figure 10.** Sterols production in one Bacteriodetes and one α-Proteobacterium. Extracted ion

668 chromatograms (m/z 69, 440, 442, 454, 456, 468, and 498) of total lipid extract (TLE) from the

669 Bacteriodetes strains *F. taffensis* and the α-Proteobacterium *M. caenitepidi*. All TLEs were

670 extracted from liquid cultures. The F. taffensis TLE was acetylated prior to running on the GC-

671 MS. *M. caenitepidi* TLEs were trimethylsilylated prior to running on the GC-MS. Sterol peaks

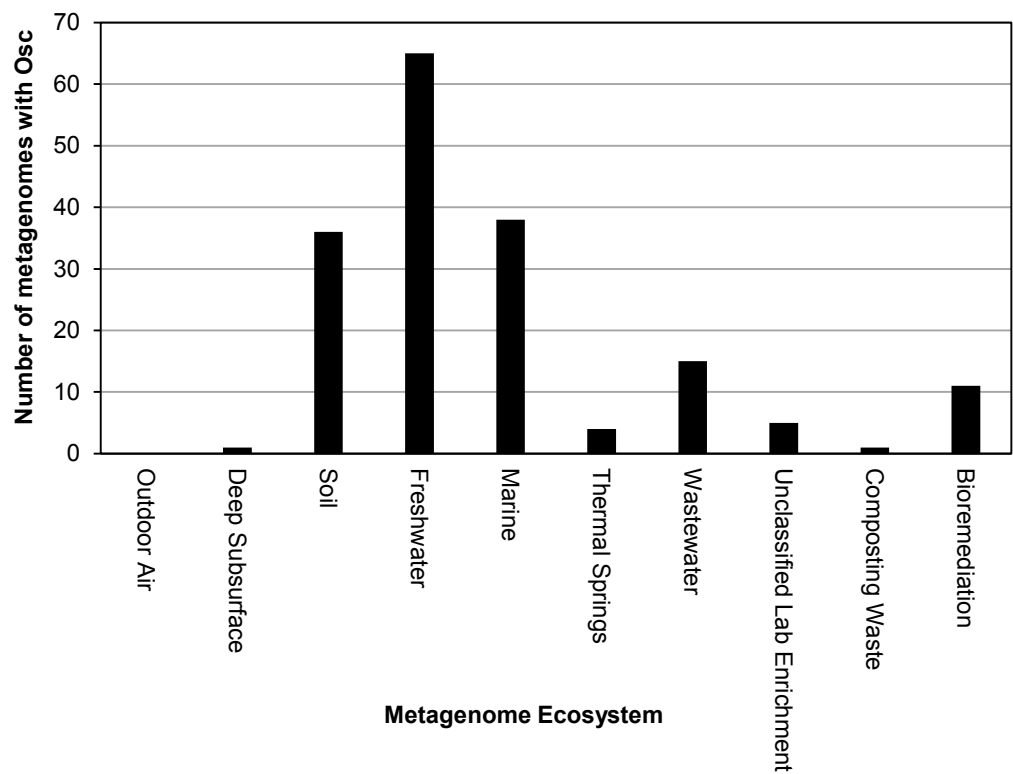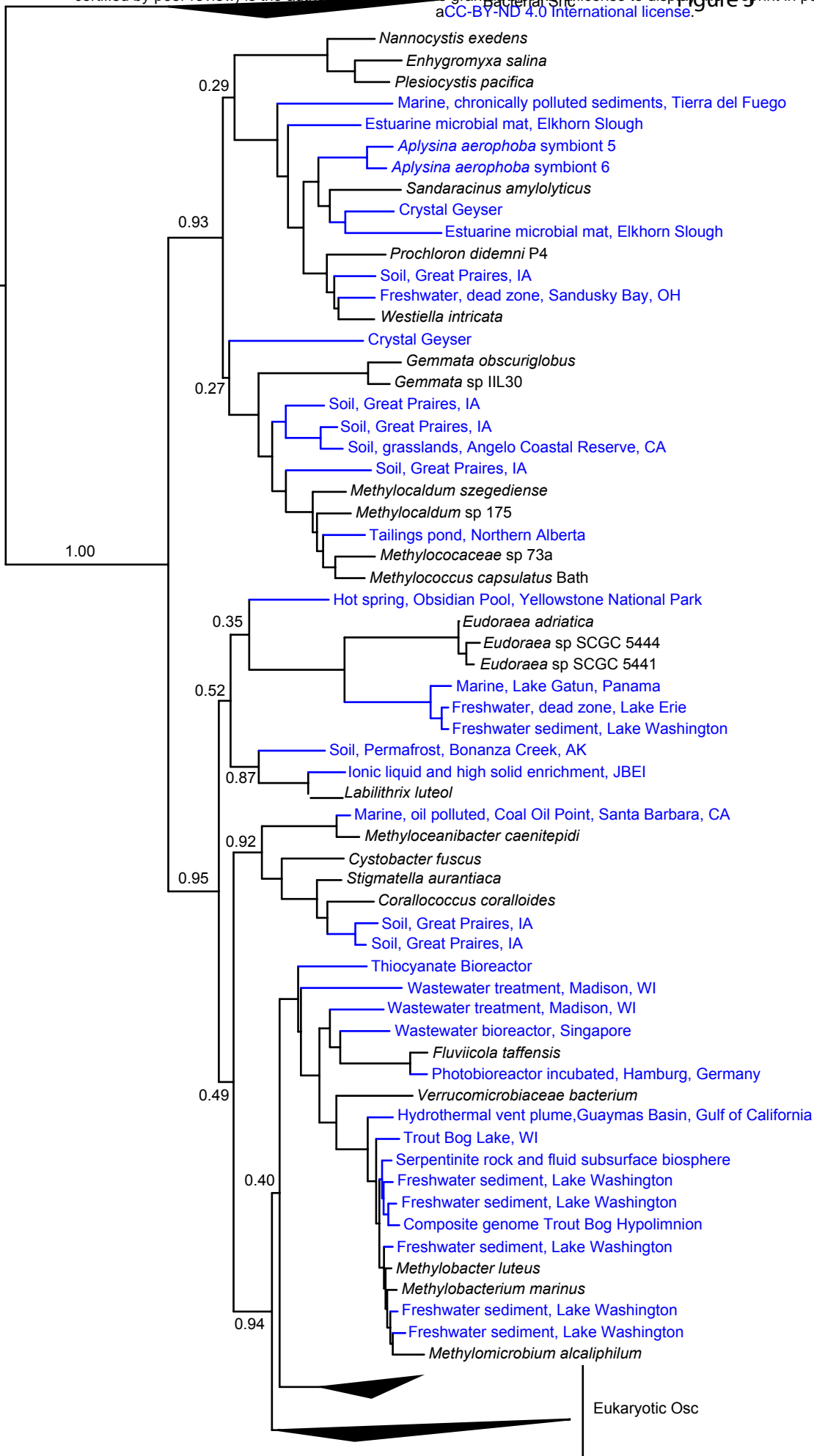672 were identified based on their mass spectra as shown in Figure 2.
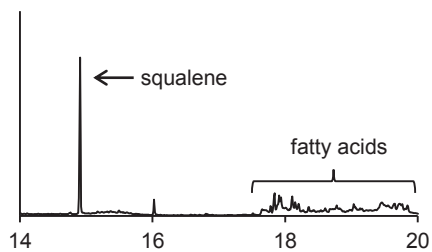
673

Figure 1

Figure 2

Figure 3

δ - Proteobacteria (Myxobacteria)

Cyanobacteria

Planctomycetes

γ - Proteobacteria (Methylococcales)

Bacteriodetes

α - Proteobacteria (Rhizobiales)

Verrucomicrobia

Figure 4

Number of metagenomes with Osc (y-axis)

Metagenome Ecosystem (x-axis)

Categories: Outdoor Air, Deep Subsurface, Soil, Freshwater, Marine, Thermal Springs, Wastewater, Unclassified Lab Enrichment, Composting Waste, Bioremediation
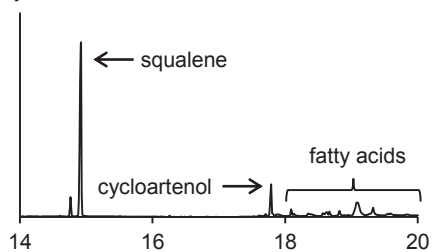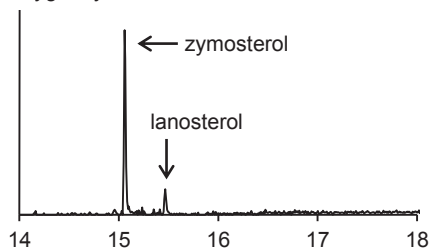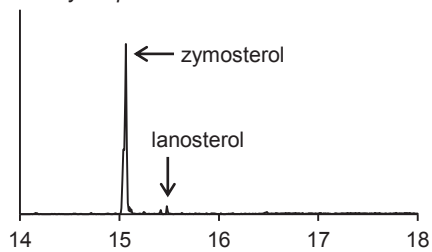
Figure 3

Figure 6

Figure 7
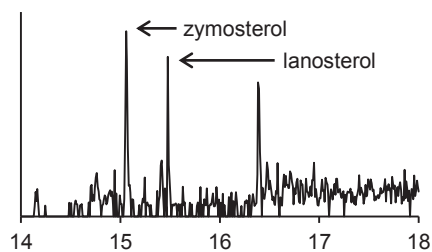
Figure 8

A. Oxidosqualene cyclase (Osc)

```
                    98        232       387       444    453       503       533       581       696       704 707
                    |         |         |         |      455         |         |         |         |         |  |
                                                        456
#G.max (Cas)        GDYGGPM WCHCR QLWDT WPFST WPISDCT ATYEL VECTS GSWGV GVFNK ITYAAYR
#D.discoideum (Cas) GDYGGPM WCHCR QLWDT WPFST WPISDCT ASYEN VECSA GSWGI GVFNF ISYSNYK
 C.coralloides (Cas) GDYSGPL WCHCR QLWDT WPFST WPISDCT ATYEQ VECTS GAWGV GVFSR IHYDAYL
*C.fuscus (Cas)     GDYGGPL WCHCR ELWDT WPFST WPISDCT ATYEL VECTS GSWGI GMFNK IHYDHYL
*F.taffensis (Cas)  GDYGGPL WCHNR QLWDT WPFST WAITDCT ASYEK TECSS GCWGV GIFNG ITYTSYR
*M.caenitepidi (Cas) GDYGGPV WCHSR RLWDT WPFSN WPISDCT PTYEK VECTS GAWGI GVFNR ITYDNYR
*S.aurantiaca (Cas) GDYGGPQ WCHCR ELWDT WPFST WPISDCT ATYEL VECTS GSWGV GIFNR IHYDAYL
*M.alcaliphilum    GDYGGPL WCHCR QLWDT WPFST WPVSDCT STYEL TECSA GGWAV GVFNY ITYANYR
*M.lacus           GDYGGPL WCHTR QLWDT WPFST WPVADCT ATYEL TECTA GGWAV GVFNY ITYANYR
*M.luteus          GDYGGPL WCHTR QLWDT WPFST WPVADCT ATYEL TECTA GGWAV GVFNY ITYANYR
*M.whittenburyi    GDYGGPL WCHTR QLWDT WPFST WPVADCT ATYEL TECTA GGWAV GVFNY ITYANYR
*S.amylolyticus    GEVVWCP YCHTR ASWDT YCFAG WPVSDCT GSYEA VECTA GMWGV GIFFH LDYVLYR
*E.salina          GEVVWNP YCHMR DIWDT WGFAN WPVSDCT GSYEP TECTA GFWGV GVFFN LDYRLYK
*P.pacifica        GEVVWNP YCHMR DIWDT WGFAD WPVSDCT GSYEE AECSA GFWGV GVFFN LDYRLYR
*M.capsulatus_Texas GEMVWCT YCHTR NAWDT WCFSD WPVSDCA GTYER VECTA GFWGI GVFFG LDYRLYP
*M.capsulatus_Bath GEMVWCT YCHTR NAWDT WCFSD WPVSDCA GTYER VECTA GFWGI GVFFG LDYRLYP
*N.exedens         GEVIWNP YCHMR DIWDT WGFAD WPVSDCT GSYEP TECTA GFWGI GVFFN LDYDLYR
*G.obscuriglobus   GEMIWCP YCHTR RTWDT WCLGD WPVSDCT GSYER IECTG GAWGV GVFFG LDYDLYR
#H.sapiens         GDYGGPL WCHCR QIWDT FSFST WIVSDCT ATYET VECTS GSWGV GVFNK ISYTSYR
#R.norvegicus      GDYGGPL WCHCR QTWDT FPFST WIVADCT ATYET VECTS GSWGV GVFNK ISYTNYR
```
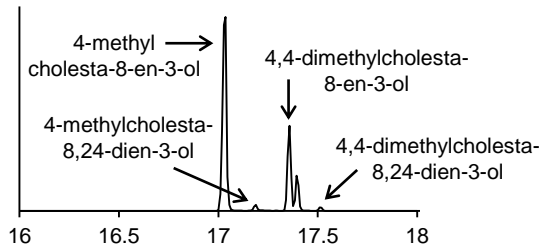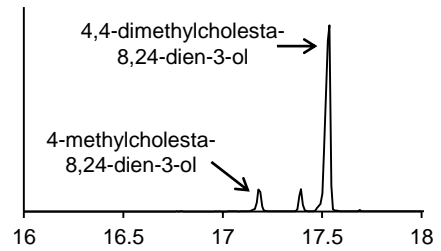
B. Squalene epoxidase (SE)

```
                    129       164       194       209       228       284     342       406   426
                    131                                               285           407
                    134
                    |||       |         |         |         |         ||        |         ||    |
#G.max             VVGAGVAGSA VGELL LGYVL KLPYP GRFIQ VCDGCF EVRCL LMGDAF LSDIV
#D.discoideum      VVGAGVAGSA VGELM FGYGI KLSYP GRFIQ VCDGCF EIRVL MIGDSW LSDCV
 C.coralloides     IAGGGPAGCA AGELL RGFAV VLPYG AALTL AADGRA RVRGC LVGDAA FHDAR
*C.fuscus          VIGGGPAGLV AGEFI FGFKI EMAYP DTLVE GADGRA AYRIL LVGDAL FLDAR
*F.taffensis       VICAGVACGV IGELM NGYNL TINYD GKFLT VCDGPM QVRLC LLGDSL FSDIL
*M.caenitepidi     VICAGPIGCA AGEWI RGFVA VLSYP HKLVD AADGRR VARVM LVGDAV LKDVE
*S.aurantiaca      IIGAGPAGCA AGELL QGFAV QLTYN ARFAQ AADGRS SWRVL LVGDAA ISDAM
*M.alcaliphilum    IIGAGMAGAT VGELL EGYAL TIPYP GRFLQ TSDGFF EVRIL MLGDAL FSDIE
*M.lacus           IIGAGMAGAT VGELL EGYSL TVPYP GRLLQ TSDGFF EVRIL MLGDAL FSDIQ
*M.luteus          IIGAGMAGAT VGELL GGYAL TIHYP GRFLQ TSDGFF EVRLL MLGDAL FSDIQ
*M.whittenburyi    IIGAGMAGAT VGELL EGYAL TIPYP GRFLQ TSDGFF EIRLL MLGDAL FSDIQ
*S.amylolyticus    IVGAGLAGCA SGEFI DGFVV VLDYA KTLVR AADGKA SFRLT LVGDAA LLDAK
*E.salina          IVGAGLAGCA SGEFI DGFVV VLDYA KTLVR AADGKA SFRLT LVGDAA LLDAK
*P.pacifica        IVGAGVAGCA SGEFI DGFVV RLDYA KRLVR AADGKA VSRLT LVGDAA LLDAH
*M.capsulatus_Texas IAGGSVACSA AGELI NGFAI LLPYG TLLKS GADGPM RARVM LVGDAR VKDAF
*M.capsulatus_Bath IAGGSVACSA AGELI NGFAI LLPYG TLLKS GADGPM RARVM LVGDAR VKDAF
*N.exedens         IIGAGMAGCA AGEFI DGFTV ELPYA KTLVK GADGKA TYRVT LVGDAA LRDAE
*G.obscuriglobus   VVGGGPVGCA GGEWL HGFAI VLPYV HLIIN GADGRQ QARMC LVGDAV FLDAV
#H.sapiens         IVGAGVLGSA VGEFL NGYMI QIPYP GRFIM VADGLF ETRVL LLGDAY FKDIK
#R.norvegicus      IIGSGVLGSA LGECL HGYVI QIPYP GKFIM VADGLF ETRVL LLGDAY LKDIK
```
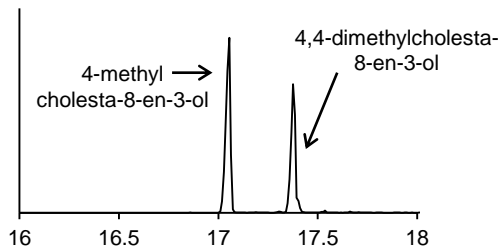
Figure 9

Methylococcus capsulatus Texas

4-methyl →
cholesta-8-en-3-ol

4,4-dimethylcholesta-
8-en-3-ol

4-methylcholesta-
8,24-dien-3-ol

4,4-dimethylcholesta-
8,24-dien-3-ol

16   16.5   17   17.5   18

Methylosarcina lacus

4,4-dimethylcholesta-
8,24-dien-3-ol →

4-methylcholesta-
8,24-dien-3-ol

16   16.5   17   17.5   18

Methylobacter luteus

4-methyl →
cholesta-8-en-3-ol

4,4-dimethylcholesta-
8-en-3-ol

16   16.5   17   17.5   18

Methylobacter whittenburyi

4-methylcholesta-
8,24-dien-3-ol

4,4-dimethylcholesta-
8-en-3-ol

4-methyl
cholesta-8-
en-3-ol

4,4-dimethylcholesta-
8,24-dien-3-ol

15   16   17   18
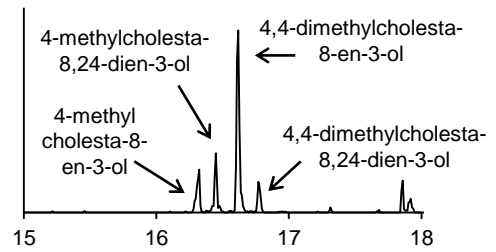
Figure 10

Fluviicola taffensis

cycloartenol ⟶

16   16.5   17   17.5   18

Methyloceanibacter caenitepidi

cycloartenol ⟶

14   14.5   15   15.5   16