1 *De novo* phytosterol synthesis in

2 animals

3 Dolma Michellod[1], Tanja Bien[2], Daniel Birgel[3], Marlene Jensen[4], Manuel Kleiner[4], Sarah Fearn[5],

4 Caroline Zeidler[1], Harald R Gruber-Vodicka[1], Nicole Dubilier[1,6*], Manuel Liebeke[1*]

5

6 [1] Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

7 [2] Institute of Hygiene, University of Münster, Robert-Koch-Str. 41, 48149, Münster, Germany

8 [3] Institute for Geology, Center for Earth System Research and Sustainability, University of Hamburg,

9 Bundesstraße 55, 20146 Hamburg, Germany

10 [4] Department of Plant and Microbial Biology NC State University, Raleigh, NC 27695, USA

11 [5] Department of Materials, Imperial College London, London SW7 2AZ, United Kingdom

12 [6] MARUM, Center for Marine Environmental Sciences, University of Bremen, 28359 Bremen, Germany

13

14 *corresponding authors: mliebeke@mpi-bremen.de, ndubilier@mpi-bremen.de

15 Abstract

16 Sterols are lipids that regulate multiple processes in eukaryotic cells, and are essential

17 components of cellular membranes. Sterols are currently assumed to be kingdom

18 specific, with phytosterol synthesis restricted to plants while animals are only able to

19 synthesize cholesterol. Here, we challenge this assumption by demonstrating that the

20 marine annelids *Olavius* and *Inanidrilus* synthesize the phytosterol sitosterol *de novo*.

21 Using multi-omics, high-resolution metabolite imaging, heterologous gene expression

22 and enzyme assays, we show that sitosterol is the most abundant (60%) sterol in these

23 animals and characterize its biosynthetic pathway. We show that phytosterol synthesis

24 partially overlaps with cholesterol synthesis and involves a non-canonical C-24 sterol

25 methyltransferase (C_{24}-SMT). C_{24}-SMT is an essential enzyme for sitosterol synthesis in

26 plants, but not known from animals with bilateral symmetry (bilaterians). Our

27 comparative phylogenetic analyses of C_{24}-SMT homologs revealed that these are

28 widely distributed across annelids and other animal phyla, including sponges and

29 rotifers. Our findings show that phytosterol synthesis and use is not restricted to the

30 plant kingdom, and indicate that the evolution of sterols in animals is more complex than

31 previously assumed.

Introduction

33 Sterols are essential lipids present in all eukaryotes. They regulate the physical

34 properties of biological membranes and are involved in the formation of specialized

35 lipid-protein microdomains critical for signal transduction (Yeagle, 1985; Simons &

36 Ikonen, 1997; Simons & Toomre, 2000). In addition, as precursors and cofactors of

37 signaling molecules, they participate in many signaling and regulatory pathways

38 (Chiang & Ferrell, 2020; Radhakrishnan et al., 2020; Sarkar & Chattopadhyay, 2022;

39 Thummel & Chory, 2002). While sterols are ubiquitous in eukaryotes, their distribution is

40 assumed to be kingdom specific: fungi synthesize ergosterol (C_{28}); plants harbor a

41 mixture of phytosterols (C_{28} to C_{29}) dominated by sitosterol, stigmasterol, and

42 campesterol (Lagarda et al., 2006) and animals use cholesterol (C_{27}). These inter-

43 kingdom differences reflect a complex evolutionary history of sterol synthesis. Previous

44 phylogenetic analyses suggest that most enzymes for the biosynthesis of plant, fungal,

45 and animal sterols were already present in the last eukaryotic common ancestor (LECA)

46 (Desmond & Gribaldo, 2009; Summons et al., 2006), with the kingdom-specific

47 distribution observed in extant species then evolving from LECA through multiple events

48 of enzyme losses and specializations.

49 Kingdom-specific sterols differ from each other in only small structural details. For

50 example, phytosterols differ from cholesterol by the presence of an extra methyl- or

51 ethyl-group at position C_{24}. This methylation is catalyzed by C_{24} sterol methyltransferase

52 (C_{24}-SMT), an enzyme widely distributed in plants, protists and fungi but absent in

53 nearly all animals (Haubrich et al., 2015; Volkman, 2005). The only exception are some

54 marine sponges, in which C_{24}-SMT homologues are not related to phytosterol synthesis

55 but assumed to participate in the synthesis of 24-isopropylcholesterol, a cholesterol

56 derivative that serves as a sponge biomarker (Germer et al., 2017; Gold et al., 2016).

57 As with most other sterol enzymes, an ancestral C_{24}-SMT was likely present in the

58 LECA but lost early in the evolution of animals, probably after the divergence of

59 sponges, explaining why this enzyme is absent from bilaterians and why animals lack

60 the ability to produce phytosterols (Desmond & Gribaldo, 2009; Gold et al., 2016;

61 Haubrich et al., 2015).

62 Here, we show that phytosterol biosynthesis and use is not restricted to plants. By

63 screening multiple members of a globally distributed group of marine gutless annelids

64 from the genera *Olavius* and *Inanidrilus*, we show that all eight species analyzed have

65 sitosterol as their main sterol and express the genes needed for its synthesis. We

66 characterized the sitosterol biosynthetic pathway used by these annelids and

67 demonstrate the importance of a previously uncharacterized group of C_{24}-SMT

68 homologs encoded in these animals' genomes. Using heterologous gene expression

69 and enzymatic assays, we show that two animal C_{24}-SMT homologs are functional and

70 catalyzes both C-24 and C-28 methylations. Finally, we discovered that C_{24}-SMT

71 homologs are widely distributed across annelids and other animal phyla, including

72 sponges and rotifers, suggesting that phytosterols may also be synthesized by other

73 animals Our findings demonstrate that animals are capable of synthesizing and using

74 phytosterol and that these molecules are not restricted to the plant kingdom, and

75 suggest that the use of phytosterols as geological biomarkers for plants should be

76 reconsidered

Results and discussion

77

Sitosterol is the main sterol in the marine gutless annelid *Olavius algarvensis*

78

79 *O. algarvensis* belongs to a group of gutless marine annelids that are found worldwide,

80 mainly in coral-reef and seagrass sediments. These annelids lack a digestive system

81 and are obligately associated with bacterial endosymbionts that provide them with

82 nutrition (Giere, 1981, 1985; Kleiner et al., 2012; Woyke et al., 2006). As part of our

83 ongoing efforts to characterize the *O. algarvensis* symbiosis, we analyzed the

84 metabolome of single worm individuals using both gas chromatography-mass

85   spectrometry (GC-MS) and high-performance liquid chromatography-mass

86   spectrometry (HPLC-MS). These analyses revealed an unusual sterol composition, with

87   sitosterol accounting for the majority of the sterols detected (60%), and the remainder

88   consisting of cholesterol (**Figure 1A**). This was unexpected, as cholesterol usually

89   dominates the sterol pool in bilaterians, often making up more than 90% of the total

90   sterol content (Goad, 1981; Sissener et al., 2018). Sitosterol is a phytosterol and,

91   among bilaterians, has only been reported as the most abundant sterol in a few

92   phytoparasitic nematodes that are incapable of *de novo* sterol synthesis (Chitwood et

93   al., 1985, 1987; Cole & Krusberg, 1967). In these plant parasites, it is unclear if the

94   detected phytosterol is only present in the nematode gut content, or incorporated into

95   their cells and tissue**.** The absence of a gut in *O. algarvensis* excludes sterol

96   contamination from plant matter in the digestive tract. However, sitosterol could

97   originate from the bacterial symbionts of *O. algarvensis*, which form a thick layer

98   between the cuticle and the epidermis of the animal (**Figure 1B**).

99   To investigate this possibility, we used two high spatial-resolution metabolite imaging

100 techniques to localize the two major sterols in *O. algarvensis.* Time-of-Flight Secondary

101 Ion Mass Spectrometry (TOF-SIMS) data revealed that, at a spatial resolution of 0.4

102 µm, both sitosterol and cholesterol were uniformly distributed throughout the animals'

103 tissues (**Figure 1C and 1D**). We found no evidence for a tissue-specific distribution of

104 these two sterols, that is, there was no correlation between symbiont location and

105 sitosterol distribution. These findings are supported by a second mass spectrometry

106 imaging method, matrix-assisted laser desorption ionization mass spectrometry imaging

107 (MALDI-2-MSI), of cross and longitudinal sections at a spatial resolution of 5 µm. The

108 MALDI imaging data of longitudinal worm sections confirmed a uniform distribution of

109 sitosterol and cholesterol throughout the animal (**Supplementary Figure 1 and 2**) and

110 the identity of these sterols (**Supplementary Table 1).** This homogeneous sterol

111 distribution sterols suggests that the bacterial symbionts are not the source for

112 phytosterol in *O. algarvensis*.
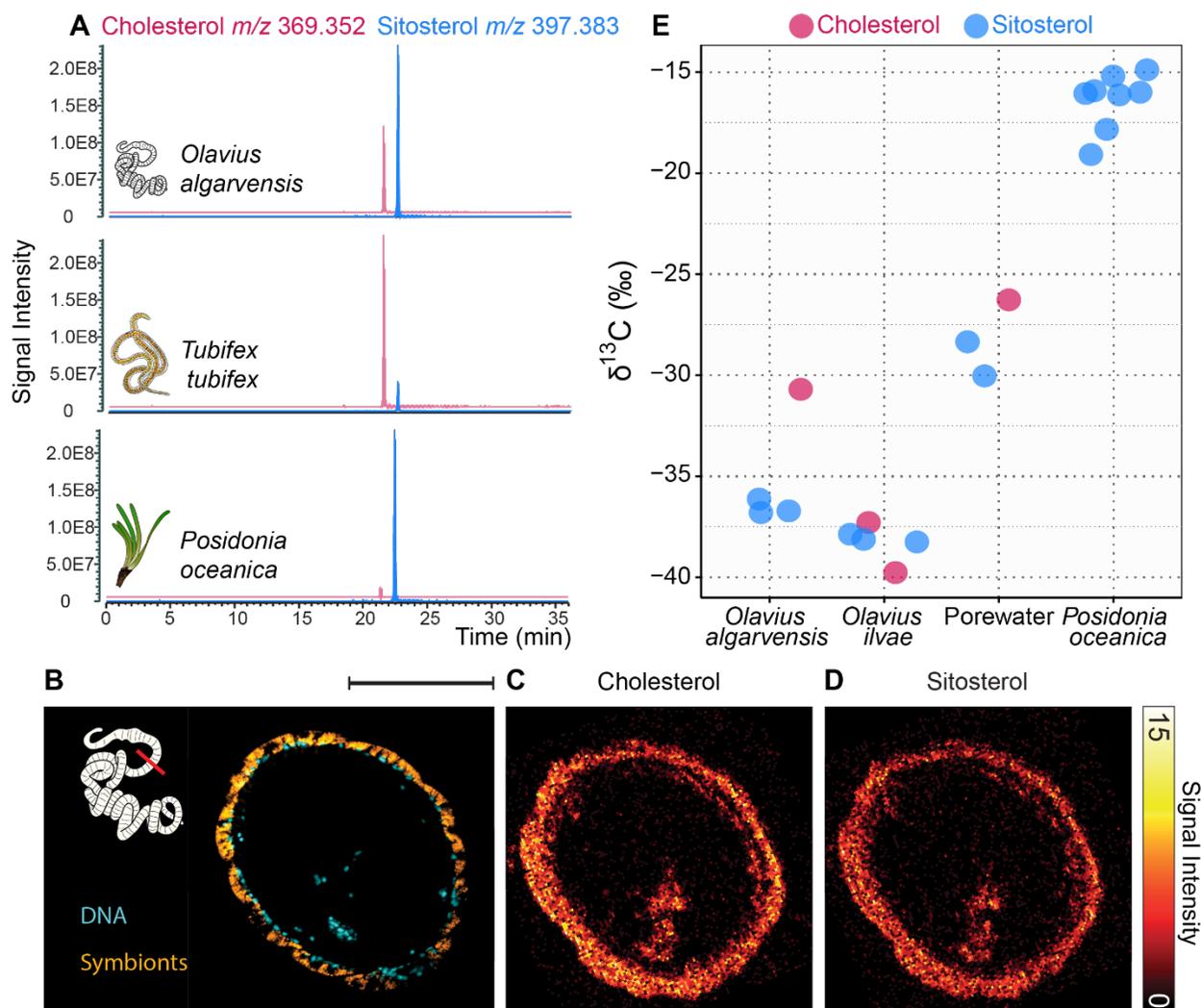
Figure 1 | *Olavius algarvensis* has an unusual sterol profile dominated by sitosterol, a plant sterol.
A, Extracted-ion chromatograms (XIC) of cholesterol ([M-H_2O+H]$^+$ $C_{27}H_{45}$ at *m/z* 369.352 (red)) and sitosterol ([M-H_2O+H]$^+$ $C_{29}H_{49}$ at *m/z* 397.383 (blue)). The XICs were generated from lipid extracts of (from top to bottom): the gutless marine annelid *O. algarvensis*, the freshwater annelid *Tubifex tubifex,* and the seagrass *Posidonia oceanica*. **B**, Chemosynthetic symbiotic bacteria are located just below the cuticle of *O. algarvensis*.16S rRNA fluorescence *in situ* hybridization (FISH) image of a cross section through a worm showing the symbionts in yellow (general eubacterial probe) and host nuclei in blue (DAPI). **C and D**, Distribution of sitosterol and cholesterol in *O. algarvensis* measured by TOF-SIMS **C**, Summed intensity of cholesterol ions (*m/z* 369.38, 385.34, 401.35) measured with TOF-SIMS. **D**, Summed intensity of sitosterol ions (*m/z*: 397.47, 383.37, 413.45) as measured by TOF-SIMS. **E**, The ^{13}C isotopic composition of sterols in gutless annelids differed from that of the neighboring seagrass (*P. oceanica*) and sediment porewater. Scale bar in **B-D** 100 μm.

O. algarvensis sterols have an isotopic composition that is distinct from their environment

Having found no evidence based on sterol distribution for a bacterial origin of sitosterol in *O. algarvensis*, we investigated if these animals acquire their sterols from the

5

129   environment. Chemical analyses of porewater collected in the vicinity of seagrass

130   meadows, the habitat of many gutless annelids (including *O. algarvensis*), showed that

131   sterols were present in the environment in concentrations sufficient to sustain the

132   growth of small sterol-auxotrophic invertebrates (**Supplementary information and**

133   **Supplementary Figure 3**). Therefore, we further investigated the origin of sterols in *O.*

134   *algarvensis* by analyzing the carbon isotopic signature ($\delta^{13}C$) of sterols in the worms,

135   their environment (which includes the seagrass *Posidonia oceanica*) and the porewater

136   of the sediments these worms live in. Carbon isotopic signatures are used to reveal

137   carbon sources and their paths through the food web. As a rule, the bulk $\delta^{13}C$ values of

138   animals reflect their dietary sources (0.5 ‰ to 2 ‰ difference) (McCutchan et al., 2003;

139   Tiunov, 2007). The $\delta^{13}C$ values of sterols are depleted in ^{13}C relative to bulk biomass by

140   5 ‰ to 8 ‰ (Canuel et al., 1997; Hayes, 2018). Results from gas chromatography

141   isotope ratio mass spectrometry (GC-IRMS) with single metabolite resolution showed

142   that sitosterol in the seagrass and porewater had $\delta^{13}C$ values ranging from −30 ‰ to

143   −15 ‰ (**Figure 1E and Supplementary information**). The sterols in *O. algarvensis*, as

144   well as those in another co-occurring gutless annelid species, *Olavius ilvae*, had much

145   lower $\delta^{13}C$ values: −38 ‰ to −36 ‰ for sitosterol and −40 ‰ to −31 ‰ for cholesterol

146   (**Figure 1E**). The difference in the isotopic signature of sterols in both *Olavius* species

147   and their environment excludes that these worms acquired sterols from their

148   environment, and instead indicates an endogenous origin. *O algarvensis*, as all other

149   *Olavius* and *Inanidrilus* species, derives all of its nutrition from its chemosynthetic

150   bacterial symbionts, and this is reflected in its bulk isotopic composition with $\delta^{13}C$ values

151   of −30.6 ‰ (Kleiner et al., 2015). The ^{13}C-depleted signatures of both cholesterol and

152   sitosterol by 1 to 10 ‰ compared to bulk biomass in *O. algarvensis* and *O. ilvae* led us

153   to hypothesize that these animals synthesize both sterols *de novo*, using organic carbon

154   derived from their chemosynthetic symbionts.

155   ***O. algarvensis* encodes and expresses enzymes involved in sitosterol synthesis**

156   **that overlap with those of cholesterol synthesis**

157   Having ruled out an external, environmental source of sitosterol in *O. algarvensis*, we

158   next investigated if the animals themselves can synthesize this phytosterol. To identify

6

159   and characterize the biosynthetic pathways involved in sterol production, we sequenced

160   and assembled the genome of *O. algarvensis* and analyzed metatranscriptomic and

161   metaproteomic data to search for enzymes involved in *de novo* sterol synthesis,

162   screening both *O. algarvensis* and its symbionts. These analyses revealed that the

163   symbionts, as most bacteria, do not encode enzymes involved in sterol synthesis. The

164   host, on the other hand, possessed the full enzymatic toolbox required for cholesterol

165   synthesis, with homologs of the 11 enzymes present in the genome of *O. algarvensis*

166   (**Supplementary Figure 4 and Supplementary Table 2**). The cholesterol biosynthesis

167   pathway, starting with squalene, is a series of ten connected enzymatic reactions

168   encoded by 11 genes (**Supplementary Figure 4 and Supplementary Table 3**).

169   Homologs of all enzymes were transcribed (11 out of 11 enzymes) and many of the

170   proteins were detected in the proteome of *O. algarvensis* (5 out of 11 proteins)

171   (**Supplementary Figure 4 and Supplementary Tables 4 and 5**), indicating active

172   expression of the genes involved in cholesterol synthesis. Phylogenetic analysis

173   allowed us to assign each homolog to an ortholog group and thus to a potential function

174   (**Supplementary Figures 5 to 14**). Collectively, these data show that *O. algarvensis*

175   has all the enzymes required for *de novo* cholesterol synthesis, which in combination

176   with the isotopic signature of their cholesterol, suggests that these worms are able to

177   synthesize cholesterol.

178   More importantly, our analyses also identified a homolog of C_{24}-SMT, an enzyme

179   essential to sitosterol synthesis, in the genome of *O. algarvensis* (**Figure 2**). As

180   described above, bilaterians are thought to lack C_{24}-SMT. C_{24}-SMT catalyzes the

181   transfer of a methyl group from S-adenosyl-L-methionine (SAM) to the sterol side chain

182   and is essential for the biosynthesis of plant sterols. The gene structure of the putative

183   C_{24}-SMT confirmed its eukaryotic origin and excluded bacterial contamination (**Figure**

184   **2**). The putative C_{24}-SMT gene is a 1071 bp open reading frame (ORF) encoding a 356

185   amino-acid polypeptide (**Supplementary Table 3**), and contains all the conserved

186   residues characteristic of C_{24}-SMT as well as the four conserved signature motifs

187   responsible for substrate binding (**Supplementary Figure 15**) (Bouvier-Navé et al.,

188   1998; Jayasimha & Nes, 2008; Nes et al., 2004, 2008; Nes & Heupel, 1986; Schaller,

189   2004; Veeramachaneni, P. P., 2005). We identified the C_{24}-SMT gene in *O. algarvensis*

190   transcriptomes and proteomes, confirming that these animals express this enzyme

191   **(Figure 2, Supplementary Table 4 and 5**). Our findings suggest that the *O. algarvensis*

192   C_{24}-SMT gene encodes a functional enzyme that may be involved in sitosterol

193   metabolism in these annelids, and represents the first discovery of a C_{24}-SMT enzyme

194   in bilaterians.

195   **The *O. algarvensis* C_{24}-SMT homolog is bifunctional and consecutively transfers**

196   **methyl groups to sterol intermediates**

197   Two methylation reactions are required for the final steps of sitosterol synthesis, one that

198   adds a methyl group at C-24 and one at C-28. These reactions can be catalyzed by the

199   same or different enzymes. For example, in most plants, C-24 and C-28 methylations are

200   mediated by different enzyme isoforms, SMT1 and SMT2 (Bouvier-Navé et al., 1998;

201   Hartmann, 2004; Neelakandan et al., 2009). Alternatively, in basal plants such as green

202   algae, a single C_{24}-SMT catalyzes both methylation steps, (Desmond & Gribaldo, 2009;

203   Haubrich et al., 2015). Sterol profiles and genome analyses of sponges, where C_{24}-SMTs

204   participate in the synthesis of the cholesterol derivative 24-isopropylcholesterol, suggest

205   that these enzymes are similarly bifunctional, although their biochemical activity remains

206   to be characterized (Gold et al., 2016). Since *O. algarvensis* encodes a single C_{24}-SMT

207   homolog, we hypothesized that this enzyme mediates both C-24 and C-28 methylation.

208
209   To test this hypothesis, we overexpressed *O. algarvensis* C_{24}-SMT in *Escherichia coli*

210   and examined its enzymatic activity, substrate preferences, and products by assaying

211   crude protein extracts with a methyl-donor (S-adenosylmethionine, SAM) and different

212   sterol substrates. The C_{24}-SMT from *O. algarvensis* expressed in *E. coli* was not able to

213   methylate classical plant sterol substrates (**Supplementary Table 14)**. However, the

214   enzyme was able to methylate zymosterol and desmosterol, two intermediates of the

215   cholesterol biosynthetic pathway. When incubated with either of these sterol substrates

216   and SAM, the *O. algarvensis* C_{24}-SMT produced a methylated C_{28} product (**Figure 2**

217   **and Supplementary Figure 16-19**). Zymosterol was methylated to fecosterol and

218   desmosterol to 24-methylene-cholesterol. The shift in retention times and changes in

219   mass spectra of the products indicated that a methyl group was added to their sterol

220 side chain, likely at the C_{24}-position (**Supplementary Figure 18-19**). These results

221 suggest that the cholesterol and sitosterol synthesis pathways overlap in *O. algarvensis*,

222 as the two C_{24}-SMT substrates, zymosterol and desmosterol, are intermediates

223 produced in the second half of the animal cholesterol synthesis pathway (**Figure 2)**.

224

225 **Figure 2 | *Olavius algarvensis* encodes and expresses a C_{24}-SMT which catalyzes two consecutive**

226 **methylations, using desmosterol as the first and 24-methylene-cholesterol as the second substrate. A,** *O.*

227 *algarvensis* encodes and expresses a homolog of C$_{24}$-SMT, an enzyme not previously found in bilaterian animals and

9

228   essential to sitosterol synthesis. Dots indicate detection in the genome, transcriptome and proteome of *O.*
229   *algarvensis*. **B,** The *O. algarvensis* C_{24}-SMT gene consists of 4 exons forming a 1071 bp open reading frame
230   encoding a 356 amino-acid polypeptide. The four conserved regions of the enzyme are highlighted by red arrows. **C,**
231   Chromatograms of enzymatic assays with desmosterol (top) and 24-methylene-cholesterol (bottom) as substrates**.** *O.*
232   *algarvensis* C_{24}-SMT, after overexpression in *E. coli*, added a methyl group to the side chains of desmosterol and 24-
233   methylene-cholesterol. In the first methylation desmosterol, an intermediate of cholesterol synthesis, was methylated
234   to produce 24-methylene-cholesterol (C_{28} sterol). In the second methylation, 24-methylene-cholesterol was
235   methylated to produce fucosterol (C_{29} sterol). **D,** Mass spectra of the different substrates and methylated products
236   from the enzymatic assays. Sterol intermediates differ by the number of methyl groups (CH_2 at *m/z* 14) attached to
237   their side chain. The side chain of desmosterol is not methylated, 24-methylene-cholesterol has a methyl group at C-
238   24, and fucosterol has two methyl groups at C-24 and C-28**.** The substrates and methylated products were identified
239   by MS, retention time and comparison with standards. The fragmentation pattern suggests that the methyl groups
240   were added to the side chain of the sterols. **E,** Structural representation of the two methylation steps in *O.*
241   *algarvensis*. **F,** Comparison of the enzyme used in the proposed sterol synthesis pathways in *Olavius* to the canonical
242   cholesterol and sitosterol synthesis pathways (similar enzymatic reactions are colored similarly)**.** The first six steps
243   are common to both cholesterol and sitosterol synthesis pathways. This trunk pathway branches off after the
244   synthesis of desmosterol. For sitosterol synthesis, desmosterol is first methylated by C_{24}-SMT to 24-methylene-
245   cholesterol, which is then methylated in a second, consecutive step by C_{24}-SMT to fucosterol. Fucosterol is reduced
246   to sitosterol by a sterol C24-reductase (DHCR24, DIM). Squalene monooxygenase (SQE), oxydosqualene cyclase
247   (LAS, CAS), sterol 14 demethylase (CYP51), sterol 14-reductase (LBR, FK), C-4 demethylation (C-4 dem.), Sterol
248   Δ7-Δ8 isomerase (EBP, HYD1), sterol 5-desaturase (SC5DL, DWF7), sterol Δ7 reductase (DHCR7, DWF5), and C-
249   24 sterol methyltransferase (C_{24}-SMT, SMT1, SMT2).

250   After confirming the first methylation step at C-24, we next searched for potential

251   substrates for the second methylation step at C-28. This second methylation is essential

252   as sitosterol is a C_{29} compound, characterized by the presence of two methyl groups on

253   its side chain. To test our hypothesis that both of these methylations are catalyzed by

254   the *O. algarvensis* C_{24}-SMT, we selected the product of the first methylation, 24-

255   methylene-cholestrol, as well as campesterol, as potential substrates for the second

256   methylation (**Supplementary Figure 20**). 24-methylene-cholesterol was the only

257   substrate to which the *O. algarvensis* C_{24}-SMT added a methyl group, producing the C_{29}

258   compound fucosterol (**Figure 2 and Supplementary Figure 20 and 21**). 24-methylene-

259   cholesterol is the product of the methylation of desmosterol, providing evidence to

260   support our hypothesis that in *O. algarvensis*, the C-24 and C-28 methylations are

261   catalyzed by the same enzyme and occur consecutively. That is, the *O. algarvensis* C_{24}-

262   SMT first methylates desmosterol at C-24 to produce the C_{28} sterol 24-methylene-

263   cholesterol, and then adds a second methyl group to 24-methylene-cholesterol at C-28,

264   to produce the C_{29} sterol fucosterol. Fucosterol differs from sitosterol by the presence of

265   a double bond at position C-24(28). This double bond is most likely removed by the

266   delta(24)-sterol reductase (DHCR24), which is expressed based on its presence in *O.*

267   *algarvensis* transcriptomes (**Supplementary Figure 4 and 14**). These results provide

268   the first evidence for an animal C_{24}-SMT able to catalyze the two methylation steps

269   needed to synthesize sitosterol from a cholesterol intermediate, revealing a previously

270   unknown pathway for phytosterol synthesis in animals (**Figure 2**).

271   **C_{24}-SMT homologs are widespread in annelids**

272   Having demonstrated the activity of an animal C_{24}-SMT homolog that enables *O.*

273   *algarvensis* to synthesize sitosterol *de novo*, we asked if other gutless annelids also

274   encode functional C_{24}-SMTs. To answer this question, we analyzed the sterol contents

275   of six additional gutless annelid species, collected at locations in the Mediterranean Sea

276   (Elba, Mallorca, Monaco) and the Caribbean Sea (Bahamas, Belize). All six species had

277   similar lipid profiles as *O. algarvensis*, with sitosterol as their major sterol

278   (**Supplementary Table 6 and Supplementary information**).

279   In addition to lipid profiling, we screened the transcriptomes of nine *Olavius* and

280   *Inanidrilus* species and found all nine species expressed a C_{24}-SMT homolog

281   (**Supplementary Table 7)**, including *O. ilvae*, which had similarly negative sitosterol

282   $\delta^{13}C$ values as *O. algarvensis* (**Figure 1E**) and also encodes a C_{24}-SMT in its genome

283   (**Supplementary Table 3**). We confirmed the C-24 methylation ability of a second C_{24}-

284   SMT homolog, from *O. clavatus*, by heterologously expressing this gene in *E. coli*. Our

285   biochemical assays demonstrated that the C_{24}-SMT from *O. clavatus* is also a

286   bifunctional sterol methyltransferase, capable of methylating zymosterol, desmosterol

287   and 24-methylene-cholesterol (**Supplementary Figure 16, 17 and 21**).

288   We next asked if C_{24}-SMT homologs are present in other annelids. We screened

289   published transcriptomes and identified C_{24}-SMTs homologs in three deep-sea gutless

290   tubeworm species and in 17 gut-bearing annelid species from marine, limnic and

291   terrestrial environments (**Supplementary Table 8**). Despite the presence of C_{24}-SMT

292   homologs in these annelids, the published sterol profiles of annelids, including four gut-

293   bearing species analyzed in this study, are dominated by cholesterol (**Supplementary**

294   **Table 8, Figure 1A)**. However, sitosterol and other methylated/ethylated sterols (C_{28}

295   and C_{29} sterols) account for a considerable proportion of total sterols at 15-30% in some

296   of these species, including the hydrothermal vent and cold seep tubeworms *Riftia*

11

297 *pachyptila* and *Paraescarpia echinospica* (**Supplementary Table 8**). These deep-sea

298 siboglinid annelids are only distantly related to *Olavius* and *Inanidrilus*, but also lack a

299 gut and gain all of their nutrition from their chemosynthetic symbionts (Bright & Giere,

300 2005). The sterol contents of these tubeworms are dominated by cholesterol and

301 desmosterol, but the C_{28} sterol campesterol and other phytosterols make up as much as

302 nearly one third of their sterol contents (Guan et al., 2021; Phleger et al., 2005; Rieley et

303 al., 1995). Our discovery of C_{24}-SMTs homologs in these deep-sea annelids indicate

304 that these tubeworms are able to synthesize phytosterols as well.

305 Bona fide C_{24}-SMT homologs are present in at least three animal phyla: sponges,

306 rotifers and annelids

307 To assess the broader distribution of C_{24}-SMT homologs among animals, we performed

308 protein searches against public databases (see Materials and Methods for details). Hits

309 were found in six other animal phyla: sponges, cnidarians, mollusks, nematodes,

310 chordata and rotifers. Of these, the C_{24}-SMT homologs detected in sponges, rotifers

311 and annelids were identified as bona fide animal C_{24}-SMTs (**Figure 3, Supplementary**

312 **information**) and are described in the next paragraph. The C_{24}-SMTs recovered from

313 cnidarians, mollusks and chordata are unlikely to originate from the animals themselves.

314 Their phylogenetic placement and high similarities to plant, algal and protist sequences

315 suggest that these sequences originated from these animals' diets or are contaminants

316 (**Supplementary Figure 22 and Supplementary information**). We also found hits in

317 some nematodes, but these belonged to a group of C-4 sterol methyltransferases (C_4-

318 SMTs), that are specific to nematodes (Chitwood, 1991; Darnet et al., 2020) and

319 phylogenetically distinct from C_{24}-SMTs (**Figure 3B, Supplementary Figure 22**).

320

321
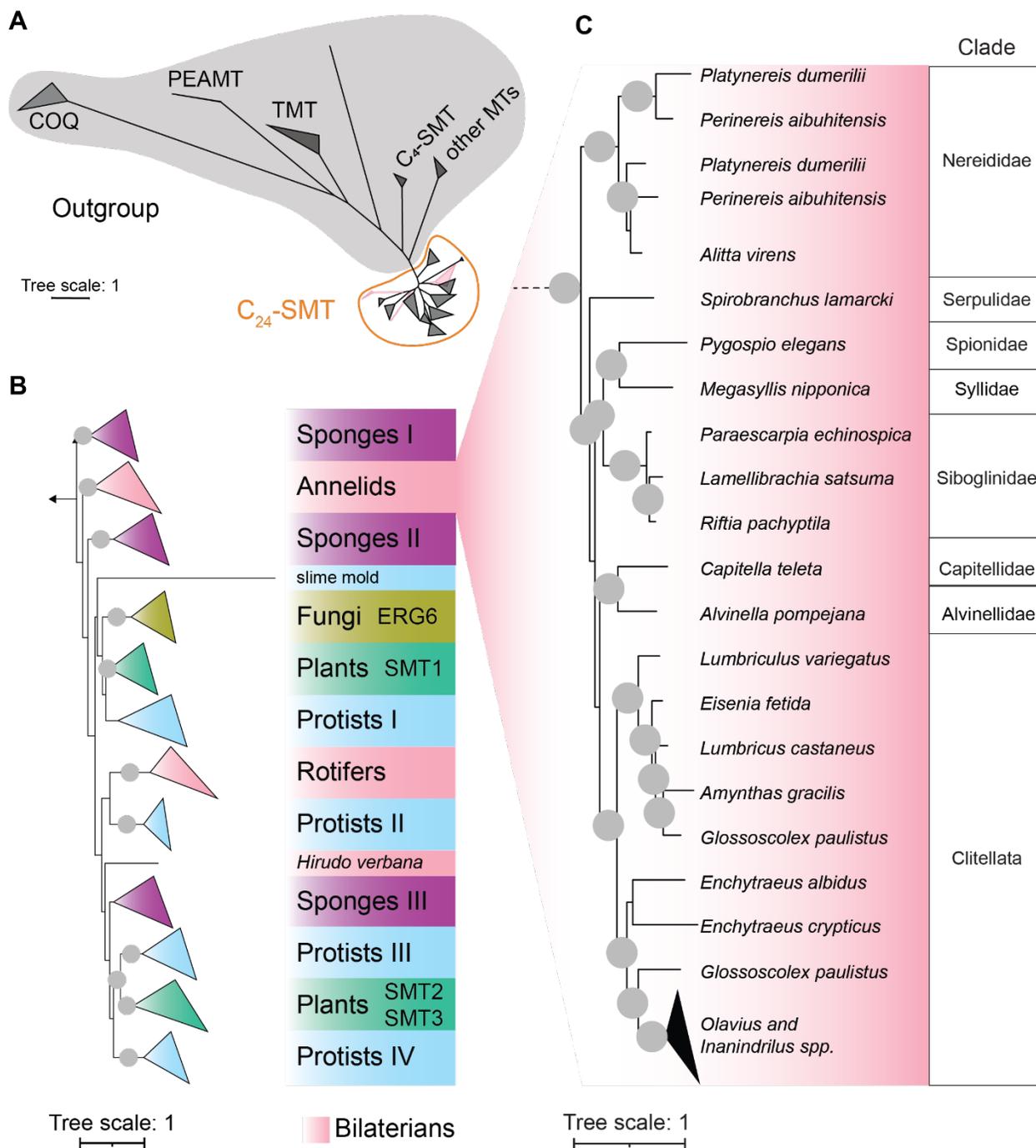
Figure 3 | C$_{24}$-SMTs are widely distributed across animals, particularly sponges, rotifers and annelids.

A, The animal C$_{24}$-SMT sequences cluster with bona fide plant and fungal C$_{24}$-SMTs and are phylogenetically distant from other SAM-dependent methyltransferases (outgroup: Ubiquinone biosynthesis O-methyltransferase (COQ), phosphoethanolamine N-methyltransferase (PEAMT), tocopherol O-methyltransferase **(**TMT), C$_4$ sterol methyltransferase (C$_4$-SMT)). Unrooted maximum likelihood amino acid tree for eukaryotic SAM-dependent methyltransferases. **B,** Maximum likelihood amino acid tree of eukaryotic C$_{24}$-SMTs. Sequences were clustered at 95% identity to make the tree more readable. Nodes with bootstrap value > 95 % are marked with grey circles. The

13

329   tree is rooted at midpoint. **C,** C_{24}-SMT homologs were detected in annelids from eight clades and were found in
330   marine, limnic and terrestrial species.

331   Our phylogenetic analyses of bona fide C_{24}-SMTs revealed that these fall into 12 clades

332   (**Figure 3**). These 12 clades include canonical C_{24}-SMTs from plants (SMT1 and

333   SMT2/SMT3) and fungi (ERG6), which have been extensively characterized in previous

334   studies (e.g. (Bouvier-Navé et al., 1998, p.; Haubrich et al., 2015; Nes et al., 2004)).

335   Animal C_{24}-SMTs do not cluster with these plant and fungi clades, suggesting that these

336   canonical C_{24}-SMTs were lost in the lineage leading to animals. Animal C_{24}-SMTs fell in

337   five subgroups. Three of these animal subgroups are from sponge C_{24}-SMTs and are

338   widely distributed across the tree. The remaining two animal clades contain sequences

339   from annelids and rotifers (**Supplementary Information**) The annelid clade includes

340   the C_{24}-SMTs we discovered in gutless *Olavius* and *Inanidrilus*, and have shown are

341   functional in producing sitosterol *de novo* in these animals. This clade also includes the

342   C_{24}-SMTs from all other annelids (26 species). These C_{24}-SMTs are widely distributed

343   across the annelid tree in eight orders, and in annelids from limnic, terrestrial and

344   marine environments (**Figure 3**). Given that the phylogeny of C_{24}-SMT corresponds well

345   to the phylogenetic evolution of annelids (**Figure 3**), it is likely that C_{24}-SMT was present

346   in the last common ancestor of annelids.

Conclusions

348   Here we describe and characterize a non-canonical animal C_{24}-SMT that enables

349   animals to synthesize sitosterol in a previously undescribed pathway that is distinct from

350   that of plants.

351   Our phylogenetic analyses of C_{24}-SMTs provide new insights into the evolutionary

352   history of sterols in eukaryotes. While ubiquitous in plants and fungi, C_{24}-SMTs were

353   thought to be largely absent from animals (Haubrich et al., 2015; Volkman, 2005). Prior

354   to this study, animal C_{24}-SMT sequences had only been reported in a few sponges (one

355   of the earliest diverging animal lineages) (Germer et al., 2017; Gold et al., 2016) and the

356   annelid *Capitella teleta* (Najle et al., 2016), but the activity of these enzymes remains

357   uncharacterized and their role in phytosterol synthesis was not explored. It has been

14

358   assumed that C_{24}-SMTs were present in the last eukaryotic common ancestor (LECA)

359   and then lost in the animal branch (Desmond & Gribaldo, 2009; Gold et al., 2016;

360   Haubrich et al., 2015). Our results, however, indicate that while the ancestral homologs

361   that led to extant plant and fungal C_{24}-SMTs were lost in the lineage leading to animals,

362   they retained non-canonical C_{24}-SMTs. These non-canonical C_{24}-SMTs are present in

363   at least three animal phyla, and sitosterol is the dominant sterol in at least some of

364   these animals, particularly *Olavius* and *Inanidrilus*. Our analyses also revealed that the

365   C_{24}-SMTs sequences isolated from sponges and protists are widely distributed across

366   the animal tree, suggesting that ancestral eukaryotes may have had several copies of

367   the C_{24}-SMT enzyme and flexible sterol synthesis pathways. While previous

368   phylogenetic analyses indicated that at least one copy of the C_{24}-SMT gene was

369   present in LECA (Desmond & Gribaldo, 2009; Gold et al., 2016), our data suggest that

370   more copies were present in the LECA and that the evolutionary history of sterol

371   synthesis pathways is more complex than previously assumed. Our results furthermore

372   challenge the widespread assumption that sterols can be used as geological biomarkers

373   for teasing apart the evolutionary history of early plants and animals (Summons et al.,

374   2022).

375   Given that C_{24}-SMTs are widespread in animals, how can we explain the unusually high

376   abundance of phytosterols in gutless *Olavius* and *Inanidrilus*? As described above,

377   other gutless animals that gain their nutrition from chemosynthetic symbionts (such as

378   vent and seep tubeworms) also have C_{28} and C_{29} sterol, but only *Olavius* and *Inanidrilus*

379   have higher amount of sitosterol than cholesterol. Sterols play many essential roles in

380   eukaryotes and their homeostasis is tightly regulated by complex mechanisms (Luo et

381   al., 2020; Wollam & Antebi, 2011). Animal membranes are usually dominated by

382   cholesterol, but studies have shown that phytosterols can be incorporated into animal

383   membranes (Mouritsen & Zuckermann, 2004) and can have beneficial effects on

384   animals. For example, they act as cholesterol-lowering agents, have anti-tumor, anti-

385   inflammatory, antibacterial and antifungal properties (Bin Sayeed & Ameen, 2015;

386   Saeidnia et al., 2014), and modulate interactions between bacterial pathogens and

387   eukaryotic hosts (van der Meer-Janssen et al., 2010). Therefore, the anti-inflammatory

388   and antibacterial properties of sitosterol, as well as its ability to protect animal cells

389 against toxins that target cholesterol (Li et al., 2015), might play a role in the symbiosis

390 between *Olavius* and *Inanidrilus* and their chemoautotrophic symbionts, by preventing

391 the symbionts from entering the host cytoplasm. In addition, changes in sterol

392 composition affect the fluidity and permeability of membranes, and these physical

393 changes in turn affect many cellular processes. For example, the high levels of

394 sitosterol in *Olavius* and *Inanidrilus* might increase the permeability of their membranes

395 for dissolved gases from their environment. These hosts compete with their aerobic

396 symbionts for the little oxygen available in their environment. Under oxic conditions,

397 these worms' symbionts consume 90% of the oxygen taken up by the worms (Häusler,

398 Lott and Dubilier, unpublished results), indicating that the worms may often experience

399 oxygen limitation. Furthermore, sitosterol has been shown to enhance mitochondrial

400 energy metabolism in a mouse cell line (Shi et al., 2013), and might enable *Olavius and*

401 *Inanidrilus* to gain more energy under low oxygen concentrations in their environment.

402 While additional studies are needed to elucidate the physiological and/or ecological

403 roles of sitosterol in animals, *Olavius and Inanidrilus* are valuable model systems for

404 studying the impact of phytosterols on animal membrane properties *in vivo* and

405 furthering our understanding of the roles sterols play in eukaryotic cells.

Materials and Methods

406

407 **Reagents.** All organic solvents were LC–MS grade: acetonitrile (ACN; Honeywell,

408 Honeywell Specialty Chemicals), chloroform (Merck), isopropanol (IPA; BioSolve),

409 methanol (MeOH; BioSolve), hexane (Sigma-Aldrich), acetone (Sigma-Aldrich), ethanol

410 (EtOH; Sigma-Aldrich) and formic acid (FA; Sigma-Aldrich). Water was deionized using

411 the Astacus MembraPure system (MembraPure). Pyridine (dried (max. 0.0075% H_2O)

412 SeccoSolv®) was obtained from Sigma-Aldrich. The reagents used for GC-MS

413 derivatization were obtain from Chromatographie Service and Sigma-Aldrich. The

414 internal standards (5α-cholestane and Ribitol) used for GC-MS analysis were obtained

415 from Sigma-Aldrich.

Sampling

Gutless annelids. Sediments in which gutless annelids occur were collected by scuba diving. The worms were extracted manually from the sediment and either directly fixed in MeOH or kept in aquaria with seagrass and sediment from the collection site for up to one year before use in experiments. Six gutless annelid species, belonging to two different genera were collected in five locations: the bay of Sant'Andrea (Island of Elba, Italy) (42° 48'29.4588" N; 10° 8' 34.4436" E), the bay of Magaluf (Mallorca, Spain) (39º 30' 14.814'' N; 2º 32' 35.868'' E), at Carrie Bow Cay (Belize) (16° 04' 59" N; 88° 04' 55" W), Twin Cayes (Belize) (16° 50' 3" N; 88° 6' 23" W), and Okinawa (Japan) (26° 29' 33.4" N; 127° 50' 31.6" E).

Gut-bearing annelids. *Cirratulidae sp.*, *Heronidrilus sp.* and *Rhyacodrilus sp.* were collected in Belize in the same environments as the gutless annelids. They were extracted manually from the sediment and directly fixed in MeOH. *Tubifex tubifex* specimens were purchased in an aquarium shop and kept in aquaria for two weeks before fixation in MeOH. *Capitella teleta* worms were provided by the Meyer Lab (https://wordpress.clarku.edu/nmeyer/) and kept in aquarium before fixation in MeOH. The MeOH fixed samples were stored at - 20°C until extraction.

Seagrass. Seagrass plants (*Posidonia oceanica*) were collected by scuba diving in the bay of Sant'Andrea (Elba, Italy) (42° 48' 29.4588" N; 10° 8' 34.4436" E). The leaves, roots and rhizomes were dissected using a razor blade, placed into individual bags and stored at -20°C.

Porewater. Porewater was collected from sediments for metabolomic analyses. We sampled in and near seagrass meadows in the Mediterranean Sea in the bay of Sant'Andrea (Elba, Italy) (42° 48' 29.4588" N; 10° 8' 34.4436" E) and in the Caribbean at Carrie Bow Cay (Belize) (16° 04' 59" N; 88° 04' 55" W) and Twin Cayes (Belize) (16° 50' 3" N; 88° 6' 23" W). Using a steel lance (1 m long, 2 mm inner diameter) fitted with a wire mesh (63 µm) to prevent the intake of sediment and seagrass, porewater was slowly extracted from sediments into polypropylene syringes. A porewater profile

17

444 consisted of top to bottom sampling of the sediments every 5 or 10 cm down to 30 cm.

445 For metabolomic analysis, 10 mL samples were stored at -20°C until further processing.

Metabolite extraction

447 **Gutless and gut-bearing annelids.** Metabolites were extracted from the worms using

448 the following method for metabolite profiling: Tissues from MeOH fixed worms were

449 transferred to 2 mL screwcap tubes containing a mix of silica beads (Sigmund Linder).

450 The residual methanol was added to the screw cap tube. The tubes were spiked with

451 100 µL 5α-cholestane (1 mM) and 40 µL ribitol (0.2 mg mL^{-1}). 0.5 mL pre-cooled

452 extraction solution (ACN:MeOH:water (v:v:v) 2:2:1) was added to each tube.

453 Tissues were disrupted by bead beating for 2 cycles of 40 sec (4 m sec^{-1}). The tissues

454 were pelleted by centrifugation (9,600 x g, 2 min), and the supernatants transferred to

455 new tubes. The pellets were extracted one more time with 1.5 mL of extraction solution.

456 The supernatants were combined and evaporated to dryness in a vacuum concentrator

457 without heating (approximately 1.5 h). The obtained aliquots were stored at -20°C until

458 metabolite derivatization.

459 **Seagrass.** The frozen plant tissues were ground to a fine powder in liquid nitrogen

460 using a pestle and mortar. 70 mg of the powder was transferred to 2 mL screw cap

461 tubes containing 1.2 mL MeOH (pre-cooled at -20°C). The tubes were vortexed for 10 s.

462 The internal standards were added to the tubes: 40 µL ribitol (0.2 mg mL^{-1}) and 100 µL

463 5α-cholestane (1 mM), and the tubes were vortexed for another 10 sec. The tubes were

464 placed on a thermomixer and shaken for 10 min at 70°C at 950 rpm. The plant powder

465 was pelleted by centrifugation (10 min, 11,000 × g, 4°C). The supernatant was

466 transferred into a new 2 mL Eppendorf tube and evaporated to dryness using a

467 Concentrator Plus (Eppendorf) (V-AL, 1.5 h, 45°C). The dried extracts were stored at -

468 20°C until metabolite derivatization.

469 **Porewater.** Sterols were extracted using Superclean LC-18 SPE tubes (6 mL, 0.5 g,

470 Supelco). The silica cartridge was equilibrated using an ultra-pure water (UPW):MeOH

471 dilution series (0:1, 1:4, 1:1, 4:1, 1:0 [v/v]). The porewater samples were spiked with

472   internal standard (100 μL 5α-cholestane (1mM)), before being loaded on the column.

473   Impurities were removed by three successive UPW washes, and the sterols eluted from

474   the cartridge with 3 × 5 mL of MeOH. The MeOH fractions were collected and

475   evaporated to dryness using a Concentrator Plus on V-AL mode with centrifugation at

476   30°C. Positive and negative controls were run in parallel. For negative controls sterols

477   were extracted from 10 mL of artificial sea water (ASW). As with positive controls, 10

478   mL ASW was spiked with 20, 40, or 80 nmol cholesterol and β-sitosterol. The dried

479   extracts were stored at -20°C until metabolite derivatization.

GC-MS analysis

480

481   **Derivatization.** To remove condensation formed during storage, we further dried the

482   extracts in a vacuum concentrator for 30 min prior to sample preparation for GC-MS

483   analysis.

484   **Gutless and gut-bearing annelids.** Metabolite derivatization was performed by adding

485   80 μL methoxyamine hydrochloride (MeOX) dissolved in pyridine (20 mg mL^{-1}) to the

486   dried pellet and incubating for 90 min at 37°C using a thermomixer (BioShake iQ,

487   Analytik Jena) under constant rotation at 1350 rpm. Following the addition of 100 μL

488   N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA), each extract was vortexed, and

489   incubated for another 30 min at 37°C on a thermomixer under constant rotation at 1350

490   rpm. After a short centrifugation, 100 μL of the supernatant was transferred to a GC-MS

491   vial (Insert G27, sping S27 and Mikor-KH-Vial G1; Chromatographic Service) for GC-

492   MS data acquisition.

493   **Seagrass.** 80 μL of MeOX (20 mg mL^{-1} dissolved in pyridine) was added to the dried

494   extracts. The resuspended dried extracts were vortexed for a few seconds and placed

495   on a thermomixer (BioShake iQ, Analytik Jena) for 90 min (37°C, 1200 rpm). 80 μL

496   BSTFA was added to the tubes. The tubes were vortexed for a few seconds and placed

497   on a thermomixer for 15 min (60°C, 1200 rpm). After a short centrifugation, 100 μL of

498   the supernatant was transferred into GC-MS vials (Insert G27, sping S27 and Mikor-KH-

499   Vial G1; Chromatographic Service) and analyzed by GC-MS.

500   **Porewater.** After complete evaporation, 80 µL BSTFA was added to the tubes. The

501   tubes were gently vortexed and placed on a thermomixer for 15 min (60°C, 950 rpm).

502   After a short centrifugation (1 min, 7,800 x g), the supernatant was transferred into GC-

503   MS vials and analyzed.

504   **Data acquisition.** The analysis of all metabolomic samples was conducted on a 7890B

505   GC system (Agilent Technologies) coupled to a 5977A single quadrupole mass

506   selective detector (Agilent Technologies). The gas chromatograph was equipped with a

507   DB-5 ms column (30 m × 0.25 mm, film thickness 0.25 µm; including 10 m DuraGuard

508   column, Agilent Technologies) and a GC inlet liner (ultra inert, splitless, single taper,

509   glass wool, Agilent). Helium was used as gas carrier at a constant flow (0.8 mL min^{-1}).

510   An Agilent 7693 autosampler injected 1 µL of derivatized sample in splitless mode. The

511   injector temperature was set at 290°C. The temperature program started at 60°C for 2

512   min, then increased to 300°C at 10°C min^{-1}, and held at 325°C for 7 min. Mass spectra

513   were acquired in electron ionization mode at 70 eV across the mass range of 50–600

514   *m/z* and a scan rate of 2 scans sec^{-1}. The retention time was locked using a standard

515   mixture of fatty acid methyl esters (Sigma-Aldrich).

516   **Data analysis.** Sterols were identified through comparison with standards using the

517   Mass Hunter Suite (Agilent) and through comparison to the NIST database. Sterols

518   were further quantified using the Mass Hunter Quantification Suite (Agilent).

519   # HPLC-MS

520   **High-resolution LC–MS/MS.** The analysis was performed using a QExactive Plus

521   Orbitrap (Thermo Fisher Scientific) equipped with a HESI probe and a Vanquish

522   Horizon UHPLC system (Thermo Fisher Scientific). The lipids were separated on an

523   Accucore C30 column (150 × 2.1 mm, 2.6 µm, Thermo Fisher Scientific), at 40°C, using

524   a solvent gradient. Buffer A (60:40 ACN:H2O, 10 mM ammonium formate, 0.1% FA)

525   and buffer B (90:10 IPA:ACN, 10 mM ammonium formate, 0.1% FA) (Breitkopf et al.,

526   2017) were used at a flow rate of 350 µl min^{-1}. The lipids were eluted from the column

527   with a gradient starting at 0% buffer B (**Supplementary Table 10**). The injection volume

528   was 10 µl. In the same run, MS measurements were acquired in positive-ion and

529   negative-ion mode for a mass detection range of *m/z* = 150–1,500. Resolution of the

530   mass analyzer was set to 70,000 for MS scans and 35,000 for MS/MS scans at *m/z* =

531   200. MS/MS scans of the eight most abundant precursor ions were acquired in positive-

532   ion and negative-ion modes. Dynamic exclusion was enabled for 30 sec and collision

533   energy was set to 30 eV (for more details see **Supplementary Table 11**). The data

534   were analyzed with Thermo FreeStyle™ (version 1.6) and Xcalibur Quan Browser

535   Software v. 2.0.3 (Thermo).

GC-IRMS

536

537   **Sample preparation**. Sterols were extracted from the gutless annelids *O. algarvensis*

538   and *O. ilvae*, as well as from the rhizome and leaves of *P. oceanica*. All samples were

539   extracted with dichloromethane:MeOH (2:1) three times. The resulting total lipid extracts

540   were then separated by solid phase extraction with a Machery & Nagel aminopropyl

541   modified silica gel column (500mg) into four fractions with increasing polarity (see

542   (Birgel et al., 2008)). For the sterols, the third fraction was used

543   (dichloromethane:acetone 9:1). Prior to measurement on the GC-MS and GC-IRMS, the

544   samples were silylated with BSTFA. The porewater samples were extracted as

545   described above.

546   **Data acquisition.** The resulting sterols were identified on a Thermo Electron Trace

547   DSQ II coupled gas chromatograph mass spectrometer (GC-MS). The GC-MS was

548   equipped with a 30 m HP-5 MS UI fused silica capillary column (0.25 mm diameter, 0.25

549   µm film thickness). The carrier gas was helium. The GC temperature program was as

550   follows: 60°C (1 min), from 60°C to 150°C at 10°C min⁻¹, from 150°C to 325°C at 4°C

551   min⁻¹, 25 min isothermal. Identification of compounds was based on retention times and

552   published mass spectral data. Compound-specific carbon stable isotope compositions

553   of sterols were measured on a gas chromatograph (Agilent 6890) coupled with a

554   Thermo Finnigan Combustion III interface to a Finnigan Delta Plus XL isotope ratio

555   mass spectrometer (GC-IRMS). The GC conditions were identical to those described

556   above for GC-MS analyses. All sterols were corrected for the additional carbons

557   introduced by derivatization with BSTFA. The standard deviation of the isotope

558   measurements was < 0.8‰.

MALDI-2-MSI

560   **Sample preparation.** The worms were prepared following Kadesch et al. (2019) with a

561   few modifications. Briefly, 20 µL 6.7% glutaraldehyde solution in marine phosphate

562   buffered saline were deposited on a glass slide. Gutless annelids were transferred to

563   the fixative using bent acupuncture needles. A coverslip was applied, and samples were

564   frozen in liquid nitrogen and stored at −80°C until further processing.

565   A single worm was transferred using featherweight forceps onto a sodium

566   carboxymethyl cellulose (CMC; Sigma-Aldrich) stamp which was glued onto a cryostat

567   specimen disc (Leica). Gelatin solution (8 % (β = 80 g L^{-1}), 10-20 µL) was used to coat

568   the worm. The specimen disc with the sample was transferred to the cryostat (Leica

569   CM3050 S, Leica Biosystems) and kept for 30 min at -22°C before sectioning into

570   sections of 12 µm thickness (chamber temperature -22°C, object temperature -22°C).

571   Sections were thaw-mounted on IntelliSlides (Bruker Daltonics), and their quality

572   determined microscopically. Sections of sufficient quality were stored in a desiccator at

573   room temperature (RT) until further analysis.

574   **Matrix application and data acquisition.** The matrix 2′,5′-dihydroxyacetophenone

575   (DHAP; Merck) was applied by sublimation and the data were acquired on a modified

576   timsTOF fleX instrument (Bruker Daltonics, Bremen, Germany) (Soltwisch et al., 2020)

577   as described in (Bien et al., 2021). The mass resolving power of this hybrid QTOF-type

578   instrument is about 40,000 (fwhm) in the investigated *m/z* range of 300–1500. For

579   material ejection, a scan range of 1 µm of the laser spot on the target was used,

580   resulting in an ablated area of 5 µm in diameter. The step size of the stage during the

581   MSI run was set to 5 µm. The laser power was set to 40%, with 50 laser shots/pixel.

582   **Data analysis.** SCiLS lab (Bruker Daltonics, version 2021a) was used for data analysis

583   and to produce the ion images shown in the figures. For image visualization, an interval

584   width of 15 ppm was used. The ion images represent the data after root mean square

585   normalization and without denoising.

TOF-SIMS

587   **Sample preparation.** *O. algarvensis* specimens were fixed with 4% paraformaldehyde

588   (PFA) at 4°C for 4 h. The fixative was removed by washing three times with marine

589   phosphate buffer. The washed samples were then stored at -20°C in MeOH.

590   PFA-fixed samples were embedded in paraffin. The MeOH was exchanged with pure

591   EtOH by three successive incubations of 60 min in pure EtOH at RT. The samples were

592   then incubated in RothiHistol (30 min, 60 min, and overnight at RT). The samples were

593   then infiltrated with paraffin at 60°C, they were placed in fresh paraffin three times for

594   30, 60 and 60 min and then left to incubate overnight. For embedding, two-thirds of the

595   embedding mold was filled with paraffin. The sample was placed in the mold and the

596   mold was filled completely with paraffin. The sample was aligned and left to polymerize

597   for one week. After polymerization, a microtome was used to cut 4 µm sections. The

598   sections were placed on poly-L-lysine-coated glass slides (Sigma-Aldrich), left to air dry

599   overnight and baked at 60°C for 2 h to improve adherence to the slide. Finally, the

600   sections were de-waxed, first in three baths of 10 min in RothiHistol, followed by an

601   EtOH series (96%, 80%, 70%, 50%), and the slides were dipped in ultra-pure water and

602   left to air dry. Once dried they were wrapped in aluminum foil and stored in a desiccator

603   (Roth, Desiccator ROTILABO® Glass, DN 250, 8.0 l) until TOF-SIMS analysis.

604   **Data acquisition and analysis:** SIMS data were acquired on an IONTOF TOF-SIMS 5

605   instrument (IONTOF GmbH) using a 25 keV Bi_3^+ LMIG analytical beam in positive

606   mode. To obtain high-resolution mass spectra, high current bunch mode was used with

607   a beam current of ~1 pA. The analytical area was typically 100 µm^2. Secondary ion

608   maps were collected using the Bi_3^+ LMIG in burst alignment mode for greater lateral

609   resolving power (raster size: 512 x 512, FoV 210 x 210). The samples were also sputter

610   pre-cleaned using a 5 keV Ar_{1000}^+ cluster ion beam. The data were analyzed in

611   SurfaceLab 7 (IONTOF). We analyzed standards to determine the most abundant ions

612 produced by each sterol: cholesterol ([M-H_2O+H]$^+$ $C_{27}H_{45}$, at *m/z* 369.38; [M-H]$^-$

613 $C_{27}H_{45}O$, at *m/z* 385.34; [M-H+O]$^-$ $C_{27}H_{45}O_2$, at *m/z* 401.35) and sitosterol ([M-H_2O+H]$^+$

614 $C_{29}H_{49}$, at *m/z* 397.47; [M-CH_3O]$^-$ $C_{28}H_{47}$, at *m/z* 383.37; [M-H]$^-$ $C_{29}H_{49}O$, at *m/z* 413.45).

615 To determine the distribution of each sterol in *O. algarvensis* sections, the intensity of

616 the three most abundant ions was combined.

Sterol identification

618 Sterols were identified by comparison to chemical standards and MS database. Matching

619 mass spectra and retention time with sterol standards confirmed sterol identification. In

620 addition, tandem mass spectra were acquired with high-mass resolution and accuracy on

621 all sterols. Each sterol was identified on using a combination of two different

622 chromatography types (GC-MS and LC-MS), including different ionization methods

623 (electrospray ionization, electron impact ionization). For sterol identification with MSI we

624 used chemical standards for sterols and measured them with MALDI-2 and SIMS in

625 parallel to the tissue sections

Identification of genes involved in sterol biosynthesis in *O. algarvensis* transcriptomes

628 Transcriptomes generated in a previous study (Wippler et al., 2016) were analyzed for

629 genes involved in sterol biosynthesis Protein sequences from humans, *Arabidopsis*

630 *thaliana* and *Saccharomyces cerevisiae* were used as queries (**Supplementary Table**

631 **12**) to search the transcriptomic assemblies with TBLASTN (e-value 1e-10). The identity

632 of the hits was confirmed by BLASTP search against the NCBI nr and Swiss-Prot

633 database as well as by INTERPROSCAN domain prediction. *O. algarvensis* sequences

634 were aligned with reference sequences (Desmond & Gribaldo, 2009) using Clustalw

635 (Larkin et al., 2007), trimmed with trimAl (Capella-Gutiérrez et al., 2009). Alignments

636 were used to calculate maximum-likelihood trees with ultrafast bootstrap support values

637 using IQ-TREE (Minh et al., 2020). The resulting trees were visualized using iTOL

638 (Letunic & Bork, 2019). The trees are shown in **Figure 3** and **Supplementary Figures**

639 **5** to **14.**

Gutless annelid nucleic acid extraction, sequencing and analysis

Extraction and sequencing. Genomic DNA was extracted from fresh specimens of two gutless annelid species (*O. algarvensis* and *O. ilvae*, one individual each). High molecular weight genomic DNA was isolated with the MagAttract HMW DNA Kit (Qiagen). Quality was assessed by the Agilent FEMTOpulse and DNA quantified by the Quantus dsDNA kit (Promega). DNA was processed to obtain a PacBio Sequencing-compatible library following the recommendations outlined in "Procedure & Checklist – Preparing HiFi Libraries from Low DNA Input Using SMRTbell Express Template Prep Kit 2.0". Libraries were sequenced on a Sequel II instrument at the Max-Planck Genome-Centre Cologne (MP-GC) with sequencing chemistry 2.0, binding kit 2.0 on one 8M SMRT cell for 30 h applying continuous long read (CLR) sequencing mode.

Assembly and identification of genes involved in sterol biosynthesis. The CLR reads were assembled using Flye (v 2.8) (Kolmogorov et al., 2019). The completeness of the assembly was assessed with QUAST (v. 5.0.0) (Gurevich et al., 2013) and BUSCO (version 5.2.2) (Seppey et al., 2019). *O. algarvensis* and *O. ilvae* sequences were retrieved from the PacBio assembly with BLAT (Kent, 2002) and SCIPIO (version 1.4) (Keller et al., 2008) using *O. algarvensis* transcripts as queries.

Metaproteomics

Protein identification and quantification. We re-analyzed data produced by Jensen (Jensen et al., 2021) using a customized database containing 1,439,433 protein sequences including host and symbiont proteins as well as a cRAP protein sequence database (http://www.thegpm.org/crap/) of common laboratory contaminants. We performed searches of the MS/MS spectra against this database with the Sequest HT node in Proteome Discoverer version 2.3.0.523 (Thermo Fisher Scientific). The following parameters were used: trypsin (full), maximum two missed cleavages, 10 ppm precursor mass tolerance, 0.1 Da fragment mass tolerance and maximum of 3 equal dynamic modifications per peptide, namely: oxidation on M (+ 15.995 Da), carbamidomethyl on C (+ 57.021 Da) and acetyl on the protein N terminus (+ 42.011

668 Da). False discovery rates (FDRs) for peptide spectral matches (PSMs) were calculated

669 and filtered using the Percolator Node in Proteome Discoverer (Spivak et al., 2009).

670 Percolator was run with a maximum delta Cn 0.05, a strict target FDR of 0.01, a relaxed

671 target FDR of 0.05 and validation based on q-value. We used the Protein FDR Validator

672 Node in Proteome Discoverer to calculate q-values for inferred proteins based on the

673 results from a search against a target-decoy database. Proteins with a q-value <0.01

674 were categorized as high-confidence identifications and proteins with a q-value of 0.01–

675 0.05 were categorized as medium-confidence identifications. We combined search

676 results for all samples into a multi-consensus report in Proteome Discoverer and only

677 proteins identified with medium or high confidence were retained, resulting in an overall

678 protein-level FDR of 5%.

C_{24}-SMT distribution in animals

680 **Gutless annelid search.** Total RNA was extracted from nine gutless annelid species.

681 RNA was quality and quantity assessed by capillary electrophoresis (Agilent

682 Bioanalyser PicoChip) and Illumina-compatible libraries were generated with the

683 NEBNext® Single Cell/Low Input RNA Library Prep Kit for Illumina®. Libraries were

684 again quality and quantity controlled followed by sequencing on a HiSeq 3000 device

685 with 2 × 150 bp paired-end read mode. The raw reads were trimmed and corrected, the

686 symbiont reads were mapped out and the host read assembled using Trinity. The

687 resulting assemblies were screened for C_{24}-SMT homologs using the model C_{24}-SMTs

688 (P25087, Q9LM02, Q39227) as subject. Only hits with e-values smaller than 1e-30 and

689 spanning at least half the subject sequences were kept for further analysis. The contigs

690 of interest were isolated and TransDecoder (v 5.5.0) (https://github.com/TransDecoder)

691 was used to identified candidate coding region TransDecoder, homology searches

692 (blastp and PFAM search) were used as additional retention criteria.

693 **Search in public databases.** To assess the presence of C_{24}-SMT homologs in other

694 animals, protein searches against NCBI (National Center for Biotechnology Information

695 (NCBI), available from: https://www.ncbi.nlm.nih.gov/) databases (nr, tsa_nr,

696 refseq_prot, env_nr and tsa) as well as the proteomes predicted by Ensembl metazoan

697   (Howe et al., 2020) and Compagen (Hemmrich & Bosch, 2008) were performed. The

698   search was restricted to metazoan sequences to avoid hits from fungi and viridiplantae.

699   Only hits with e-values smaller than 1e-30 and a coverage > 60% were retained for

700   further analysis.

701   **Phylogenetic tree.** A phylogenetic tree was constructed from selected C_{24}-SMT protein

702   sequences. Briefly, published C_{24}-SMT sequences (Desmond & Gribaldo, 2009; Gold

703   et al., 2016) were downloaded from UniProt and JGI Genome Portal and used as

704   references. Other SAM-dependent methyltransferases (ubiquinone biosynthesis O-

705   methyltransferase, phosphoethanolamine N-methyltransferase, tocopherol O-

706   methyltransferase and C_4-sterol methyltransferase) were used as outgroups. Together

707   with the animal C_{24}-SMTs these sequences were clustered (95% ID), aligned using

708   Clustalw 2.1 (Larkin et al., 2007) and trimmed with TrimAI v1.2 (Capella-Gutiérrez et al.,

709   2009). IQ-TREE (v1.6.12) (Minh et al., 2020) was used to predict the best-fit models of

710   evolution and to calculate a maximum-likelihood tree with ultrafast bootstrap support

711   values from the concatenated alignment. The resulting tree was visualized using iTOL

712   v6 (Letunic & Bork, 2019).

713   # C_{24}-SMT heterologous gene expression and enzyme assay

714   **Heterologous gene expression.** To determine the activity of the putative C_{24}-SMT

715   enzymes, we overexpressed OalgSMT, OclaSMT and ArathSMT1 (**Supplementary**

716   **Table 13**) in *E. coli* and performed enzymatic assays. GenScript (Genscript®)

717   generated pet28a(+) (NheI/XhoI) plasmid containing the sequence of interest. For

718   expression, the pet28(a)-OalgSMT and pet28(a)-OclaSMT vectors were transformed in

719   Lemo21(DE3) *E. coli* competent cells (New England Biolabs (NEB)). The pet28(a)-

720   ArathSMT1 vector was transformed in Overexpress C43(DE3) pLysS *E. coli* competent

721   cells (Lucigen). A single colony of transformed cells was grown in 3 mL lysogeny broth

722   (LB) supplemented with the appropriate antibiotics for 8 h (37°C, 150 rpm). 1 mL pre-

723   culture was used to inoculate 1 L ZYP-5052-Rich-Autoinduction-Medium (Studier, 2005)

724   supplemented with antibiotics and 1500 µM rhamnose. The cultures were grown for 72

725   h at 20°C, with rotation at 150 rpm. Cells were harvested by centrifugation at 4,500 × g

726   for 25 min at 4°C, the supernatant discarded and the resulting pellet stored at -20°C

727   until further use.

Protein extraction/cell lysis. The frozen pellets were thawed on ice. They were then resuspended in 15 mL sucrose solution (750 mM sucrose solution in 20 mM phosphate buffer at pH 7.5). Once the pellets were dissolved in the sucrose solution, 5 mg lysozyme was added and the tubes shaken for 10 min at RT. Cells were lysed by addition of 30 mL lysis buffer (100 mM NaCl, 15% glycerol, 3 mM EDTA in 20 mM phosphate buffer at pH 7.5), 0.2 mL $MgSO_4$ stock solution (120 g L^{-1}) and 0.3 mL Triton X-100. The mix was vigorously shaken and incubated on ice until it reached a gelatinous consistency. The DNA was fragmented by addition of 2 mL DNase stock solution (50 mg DNase I in 35 mL buffer (20 mM Tris-HCl pH8, 0.5 M NaCl) and 15 mL glycerol). Finally, the cell fragments and inclusion bodies were pelleted by centrifugation (45min, 16,000 × g, 4°C). The supernatant, containing soluble proteins, was aliquoted and stored at -80°C until further use. The overexpression yielded a target protein migrating on SDS-PAGE with the expected size of ~40 kDa.

Enzymatic assay. We tested nine sterol substrates (cycloartenol, lanosterol, zymosterol, lathosterol, 7-dehydrocholesterol, desmosterol, campesterol, cholesterol and 24-methylene-cholesterol) (**Supplementary Table 14**). The assay was performed in 600 μL total volume. 100 μL crude soluble protein extract was mixed with 400 μL phosphate buffer (20 mM, 5% glycerol, pH 7.5) in a 15 mL tube containing a sterol substrate (final concentration 100 μM) dispersed in Tween 80 (0.1% v/v). The reaction was initiated with 100 μL SAM working solution (0.6 mM). The reaction was performed in a water bath at 35°C for 16 h. The reaction was terminated with 600 μL 10% methanolic KOH. The products were extracted three times with 2.5 mL hexane and mixed on a vortex for 30 s. The resulting organic layer was evaporated to dryness in a concentrator at 30°C, V-AL for 1.5 h. Two internal standards, 5α-cholestane (100 μL, 1 mM solution) and ribitol (40 μL 200 mg/L solution), were added to the tubes and evaporated to dryness. The samples were derivatized and analyzed on an Agilent GC-MS as described above or directly analyzed on a QExactive Plus Orbitrap (Thermo Fisher Scientific) equipped with a HESI probe and a Vanquish Horizon UHPLC system

756   (Thermo Fisher Scientific). The sterols were separated on a C18 column (Accucore

757   Vanquish C18+, 100 x 2.1, 1.5 μm, Thermo Fischer Scientific), for method details see

758   **Supplementary Table 11 and 15**.

Data availability

760   The metaproteomic mass spectrometry data have been deposited at the

761   ProteomeXchange Consortium via the PRIDE partner repository (Vizcaíno et al., 2016)

762   with the following dataset identifiers: PXD014881.

763   The upload of metabolomics and sequencing data on public platforms is in progress.

Author Contributions

765   D.M., M.L. and N.D. conceived the study. D.M. collected, processed and analyzed the

766   metabolomic, metatranscriptomic and metagenomic data. T.B. and S.F. collected mass

767   spectrometry imaging data. D.B. performed the GC-IRMS measurements. M.J. and

768   M.K. collected and analyzed the proteomics data. C.Z. and D.M. performed the

769   heterologous gene expression. D.M. performed the enzyme assay and analyzed the

770   data. D.M. and M.L. wrote the manuscript together with N.D. and contributions from

771   T.B., D.B., M.K., and C.Z..

Acknowledgments

References

Bien, T., Hambleton, E. A., Dreisewerd, K., & Soltwisch, J. (2021). Molecular insights into symbiosis—Mapping sterols in a marine flatworm-algae-system using high spatial resolution MALDI-2-MS imaging with ion mobility separation. *Analytical and Bioanalytical Chemistry*, *413*(10), 2767–2777. https://doi.org/10.1007/s00216-020-03070-0

Bin Sayeed, M. S., & Ameen, S. S. (2015). Beta-Sitosterol: A Promising but Orphan Nutraceutical to Fight Against Cancer. *Nutrition and Cancer*, *67*(8), 1216–1222. https://doi.org/10.1080/01635581.2015.1087042

Birgel, D., Elvert, M., Han, X., & Peckmann, J. (2008). 13C-depleted biphytanic diacids as tracers of past anaerobic oxidation of methane. *Organic Geochemistry*, *39*(1), 152–156. https://doi.org/10.1016/j.orggeochem.2007.08.013

Bouvier-Navé, P., Husselstein, T., & Benveniste, P. (1998). Two families of sterol methyltransferases are involved in the first and the second methylation steps of plant sterol biosynthesis. *European Journal of Biochemistry*, *256*(1), 88–96. https://doi.org/10.1046/j.1432-1327.1998.2560088.x

Breitkopf, S., Ricoult, S., Yuan, M., Xu, Y., Peake, D., Manning, B., & Asara, J. (2017). A relative quantitative positive/negative ion switching method for untargeted lipidomics via high resolution LC-MS/MS from any biological source. *Metabolomics*, *13*. https://doi.org/10.1007/s11306-016-1157-8

Bright, M., & Giere, O. (2005). Microbial symbiosis in Annelida. *Symbiosis*, *38*(1), 1–45.

Canuel, E. A., Freeman, K. H., & Wakeham, S. G. (1997). Isotopic compositions of lipid biomarker compounds in estuarine plants and surface sediments. *Limnology and Oceanography*, *42*(7), 1570–1583. https://doi.org/10.4319/lo.1997.42.7.1570

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chiang, J. Y. L., & Ferrell, J. M. (2020). Bile Acid Biology, Pathophysiology, and Therapeutics. *Clinical Liver Disease*, *15*(3), 91–94. https://doi.org/10.1002/cld.861

Chitwood, D. J. (1991). Nematode sterol biochemistry. *Physiology and Biochemistry of Sterols.*, 257–293.

Chitwood, D. J., Hutzell, P. A., & Lusby, W. R. (1985). Sterol Composition of the Corn Cyst Nematode, *Heterodera zeae*, and Corn Roots. *Journal of Nematology*, *17*(1), 64–68.

Chitwood, D. J., McClure, M. A., Feldlaufer, M. F., Lusby, W. R., & Oliver, T. E. (1987). Sterol Composition and Ecdysteroid Content of Eggs of the Root-knot Nematodes *Meloidogyne incognita* and *M. arenaria*. *Journal of Nematology*, *19*(3), 352–360.

Cole, R. J., & Krusberg, L. R. (1967). Sterol composition of the Nematodes *Ditylenchus triformis* and *Ditylenchus dipsaci*, and host tissues. *Experimental Parasitology*, *21*(2), 232–239. https://doi.org/10.1016/0014-4894(67)90085-9

Darnet, S., Fliesler, S. J., & Schaller, H. (2020). Worming our way toward multiple evolutionary origins of convergent sterol pathways. *Journal of Lipid Research*, *61*(2), 129–132. https://doi.org/10.1194/jlr.C119000600

Desmond, E., & Gribaldo, S. (2009). Phylogenomics of Sterol Synthesis: Insights into the Origin, Evolution, and Diversity of a Key Eukaryotic Feature. *Genome Biology and Evolution*, *1*, 364–381. https://doi.org/10.1093/gbe/evp036

Germer, J., Cerveau, N., & Jackson, D. J. (2017). The Holo-Transcriptome of a Calcified Early Branching Metazoan. *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00081

Giere, O. (1981). The Gutless Marine Oligochaete *Phallodrilus leukodermatus*. Structural Studies on an Aberrant Tubificid Associated with Bacteria. *Marine Ecology Progress Series*, *5*, 353–357. https://doi.org/10.3354/meps005353

Giere, O. (1985). The Gutless Marine Tubificid Phallodrilus planus, a Flattened Oligochaete with Symbiotic Bacteria. Results from Morphological and Ecological Studies. *Zoologica Scripta*, *14*(4), 279–286. https://doi.org/10.1111/j.1463-6409.1985.tb00198.x

Gold, D. A., Grabenstatter, J., Mendoza, A. de, Riesgo, A., Ruiz-Trillo, I., & Summons, R. E. (2016). Sterol and genomic analyses validate the sponge biomarker hypothesis. *Proceedings of the National Academy of Sciences*, *113*(10), 2684–2689. https://doi.org/10.1073/pnas.1512614113

Guan, H., Birgel, D., Feng, D., Peckmann, J., Liu, L., Liu, L., & Tao, J. (2021). Lipids and their δ13C values reveal carbon assimilation and cycling in the methane-seep tubeworm *Paraescarpia echinospica* from the South China Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, *174*, 103556. https://doi.org/10.1016/j.dsr.2021.103556

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hartmann, M.-A. (2004). 5 Sterol metabolism and functions in higher plants. In G. Daum (Ed.), *Lipid Metabolism and Membrane Biogenesis* (pp. 183–211). Springer. https://doi.org/10.1007/978-3-540-40999-1_6

Haubrich, B. A., Collins, E. K., Howard, A. L., Wang, Q., Snell, W. J., Miller, M. B., Thomas, C. D., Pleasant, S. K., & Nes, W. D. (2015). 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. https://doi.org/10.1016/j.phytochem.2014.07.019

Hayes, J. M. (2018). 3. Fractionation of Carbon and Hydrogen Isotopes in Biosynthetic Processes. In *Stable Isotope Geochemistry* (pp. 225–278). De Gruyter. https://www.degruyter.com/document/doi/10.1515/9781501508745-006/html

Hemmrich, G., & Bosch, T. C. G. (2008). Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation. *BioEssays*, *30*(10), 1010–1018. https://doi.org/10.1002/bies.20813

Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D. M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., George, N., … Flicek, P. (2020). Ensembl Genomes 2020—Enabling non-vertebrate genomic research. *Nucleic Acids Research*, *48*(D1), D689–D695. https://doi.org/10.1093/nar/gkz890

Jayasimha, P., & Nes, W. D. (2008). Photoaffinity Labeling and Mutational Analysis of 24-C-Sterol Methyltransferase Defines the AdoMet Binding Site. *Lipids*, *43*(8), 681–693. https://doi.org/10.1007/s11745-008-3198-x

Jensen, M., Wippler, J., & Kleiner, M. (2021). Evaluation of RNAlater™ as a field-compatible preservation method for metaproteomic analyses of bacteria-animal symbioses (p. 2021.06.16.448770). https://doi.org/10.1101/2021.06.16.448770

Kadesch, P., Quack, T., Gerbig, S., Grevelding, C. G., & Spengler, B. (2019). Lipid Topography in Schistosoma mansoni Cryosections, Revealed by Microembedding and High-Resolution Atmospheric-Pressure Matrix-Assisted Laser Desorption/Ionization (MALDI) Mass Spectrometry Imaging. *Analytical Chemistry*, *91*(7), 4520–4528. https://doi.org/10.1021/acs.analchem.8b05440

Keller, O., Odronitz, F., Stanke, M., Kollmar, M., & Waack, S. (2008). Scipio: Using protein sequences to determine the precise exon/intron structures of genes and their orthologs in closely related species. *BMC Bioinformatics*, *9*(1), 278. https://doi.org/10.1186/1471-2105-9-278

Kent, W. J. (2002). BLAT—The BLAST-Like Alignment Tool. *Genome Research*, *12*(4), 656–664. https://doi.org/10.1101/gr.229202

Kleiner, M., Wentrup, C., Holler, T., Lavik, G., Harder, J., Lott, C., Littmann, S., Kuypers, M. M. M., & Dubilier, N. (2015). Use of carbon monoxide and hydrogen by a bacteria–animal symbiosis from seagrass sediments. *Environmental Microbiology*, *17*(12), 5023–5035. https://doi.org/10.1111/1462-2920.12912

Kleiner, M., Wentrup, C., Lott, C., Teeling, H., Wetzel, S., Young, J., Chang, Y.-J., Shah, M., VerBerkmoes, N. C., Zarzycki, J., Fuchs, G., Markert, S., Hempel, K., Voigt, B., Becher, D., Liebeke, M., Lalk, M., Albrecht, D., Hecker, M., … Dubilier, N. (2012). Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proceedings of the National Academy of Sciences*, *109*(19), E1173–E1182. https://doi.org/10.1073/pnas.1121198109

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*(5), 540–546. https://doi.org/10.1038/s41587-019-0072-8

Lagarda, M. J., García-Llatas, G., & Farré, R. (2006). Analysis of phytosterols in foods. *Journal of Pharmaceutical and Biomedical Analysis*, *41*(5), 1486–1496. https://doi.org/10.1016/j.jpba.2006.02.052

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, *23*(21), 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, *47*(W1), W256–W259. https://doi.org/10.1093/nar/gkz239

Li, H., Zhao, X., Wang, J., Dong, Y., Meng, S., Li, R., Niu, X., & Deng, X. (2015). β-sitosterol interacts with pneumolysin to prevent *Streptococcus pneumoniae* infection. *Scientific Reports*, *5*(1), 17668. https://doi.org/10.1038/srep17668

Luo, J., Yang, H., & Song, B.-L. (2020). Mechanisms and regulation of cholesterol homeostasis. *Nature Reviews Molecular Cell Biology*, *21*(4), 225–245. https://doi.org/10.1038/s41580-019-0190-7

McCutchan, J. H., Lewis, W. M., Kendall, C., & McGrath, C. C. (2003). Variation in trophic shift for stable isotope ratios of carbon, nitrogen, and sulfur. *Oikos*, *102*(2), 378–390. https://doi.org/10.1034/j.1600-0706.2003.12098.x

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mouritsen, O. G., & Zuckermann, M. J. (2004). What's so special about cholesterol? *Lipids*, *39*(11), 1101–1113. https://doi.org/10.1007/s11745-004-1336-x

Najle, S. R., Molina, M. C., Ruiz-Trillo, I., & Uttaro, A. D. (2016). Sterol metabolism in the filasterean *Capsaspora owczarzaki* has features that resemble both fungi and animals. *Open Biology*, *6*(7), 160029. https://doi.org/10.1098/rsob.160029

Neelakandan, A. K., Song, Z., Wang, J., Richards, M. H., Wu, X., Valliyodan, B., Nguyen, H. T., & Nes, W. D. (2009). Cloning, functional expression and phylogenetic analysis of plant sterol 24C-methyltransferases involved in sitosterol biosynthesis. *Phytochemistry*, *70*(17), 1982–1998. https://doi.org/10.1016/j.phytochem.2009.09.003

Nes, W. D., & Heupel, R. C. (1986). Physiological requirement for biosynthesis of multiple 24β-methyl sterols in *Gibberella fujikuroi*. *Archives of Biochemistry and Biophysics*, *244*(1), 211–217. https://doi.org/10.1016/0003-9861(86)90110-4

Nes, W. D., Jayasimha, P., & Song, Z. (2008). Yeast sterol C24-methyltransferase: Role of highly conserved tyrosine-81 in catalytic competence studied by site-directed mutagenesis and thermodynamic analysis. *Archives of Biochemistry and Biophysics*, *477*(2), 313–323. https://doi.org/10.1016/j.abb.2008.05.016

Nes, W. D., Jayasimha, P., Zhou, W., Kanagasabai, R., Jin, C., Jaradat, T. T., Shaw, R. W., & Bujnicki, J. M. (2004). Sterol Methyltransferase: Functional Analysis of Highly Conserved Residues by Site-Directed Mutagenesis. *Biochemistry*, *43*(2), 569–576. https://doi.org/10.1021/bi035257z

Phleger, C. F., Nelson, M. M., Groce, A. K., Cary, S. C., Coyne, K. J., & Nichols, P. D. (2005). Lipid composition of deep-sea hydrothermal vent tubeworm *Riftia pachyptila*, crabs *Munidopsis subsquamosa* and *Bythograea thermydron*, mussels *Bathymodiolus sp.* And limpets *Lepetodrilus spp. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *141*(2), 196–210. https://doi.org/10.1016/j.cbpc.2005.03.001

Radhakrishnan, A., Rohatgi, R., & Siebold, C. (2020). Cholesterol access in cellular membranes controls Hedgehog signaling. *Nature Chemical Biology*, *16*(12), 1303–1313. https://doi.org/10.1038/s41589-020-00678-2

Rieley, G., Dover, C. L. van, Hedrick, D. B., White, D. C., & Eglinton, G. (1995). Lipid characteristics of hydrothermal vent organisms from 9°N, East Pacific Rise. *Geological Society, London, Special Publications*, *87*(1), 329–342. https://doi.org/10.1144/GSL.SP.1995.087.01.25

Saeidnia, S., Manayi, A., Gohari, A. R., & Abdollahi, M. (2014). The story of beta-sitosterol-a review. *European Journal of Medicinal Plants*, *4*(5), 590.

Sarkar, P., & Chattopadhyay, A. (2022). Cholesterol in GPCR Structures: Prevalence and Relevance. *The Journal of Membrane Biology*, *255*(1), 99–106. https://doi.org/10.1007/s00232-021-00197-8

Schaller, H. (2004). New aspects of sterol biosynthesis in growth and development of higher plants. *Plant Physiology and Biochemistry*, *42*(6), 465–476. https://doi.org/10.1016/j.plaphy.2004.05.012

Seppey, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods in Molecular Biology (Clifton, N.J.)*, *1962*, 227–245. https://doi.org/10.1007/978-1-4939-9173-0_14

Shi, C., Wu, F., & Xu, J. (2013). Incorporation of β-sitosterol into mitochondrial membrane enhances mitochondrial function by promoting inner mitochondrial membrane fluidity. *Journal of Bioenergetics and Biomembranes*, *45*(3), 301–305. https://doi.org/10.1007/s10863-012-9495-3

Simons, K., & Ikonen, E. (1997). Functional rafts in cell membranes. *Nature*, *387*(6633), 569–572. https://doi.org/10.1038/42408

Simons, K., & Toomre, D. (2000). Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology*, *1*(1), 31–39. https://doi.org/10.1038/35036052

Soltwisch, J., Heijs, B., Koch, A., Vens-Cappell, S., Höhndorf, J., & Dreisewerd, K. (2020). MALDI-2 on a Trapped Ion Mobility Quadrupole Time-of-Flight Instrument for Rapid Mass Spectrometry Imaging and Ion Mobility Separation of Complex Lipid Profiles. *Analytical Chemistry*, *92*(13), 8697–8703. https://doi.org/10.1021/acs.analchem.0c01747

Spivak, M., Weston, J., Bottou, L., Käll, L., & Noble, W. S. (2009). Improvements to the Percolator Algorithm for Peptide Identification from Shotgun Proteomics Data Sets. *Journal of Proteome Research*, *8*(7), 3737–3745. https://doi.org/10.1021/pr801109k

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, *41*(1), 207–234. https://doi.org/10.1016/j.pep.2005.01.016

Summons, R. E., Bradley, A. S., Jahnke, L. L., & Waldbauer, J. R. (2006). Steroids, triterpenoids and molecular oxygen. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1470), 951–968. https://doi.org/10.1098/rstb.2006.1837

Summons, R. E., Welander, P. V., & Gold, D. A. (2022). Lipid biomarkers: Molecular tools for illuminating the history of microbial life. *Nature Reviews Microbiology*, *20*(3), 174–185. https://doi.org/10.1038/s41579-021-00636-2

Thummel, C. S., & Chory, J. (2002). Steroid signaling in plants and insects—Common themes, different pathways. *Genes & Development*, *16*(24), 3113–3129. https://doi.org/10.1101/gad.1042102

Tiunov, A. V. (2007). Stable isotopes of carbon and nitrogen in soil ecological studies. *Biology Bulletin*, *34*(4), 395–407. https://doi.org/10.1134/S1062359007040127

van der Meer-Janssen, Y. P. M., van Galen, J., Batenburg, J. J., & Helms, J. B. (2010). Lipids in host–pathogen interactions: Pathogens exploit the complexity of the host cell lipidome. *Progress in Lipid Research*, *49*(1), 1–26. https://doi.org/10.1016/j.plipres.2009.07.003

Veeramachaneni, P. P. (2005). Active site mapping of yeast recombinant sterol methyltransferase (Y81 W) from *Saccharomyces cerevisiae* and inhibition of ergosterol biosynthesis in *Candida albicans*. [Master Thesis]. Texas Tech University Lubbock.

Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.-W., Wang, R., & Hermjakob, H. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Research*, *44*(D1), D447–D456. https://doi.org/10.1093/nar/gkv1145

Volkman, J. K. (2005). Sterols and other triterpenoids: Source specificity and evolution of biosynthetic pathways. *Organic Geochemistry*, *36*(2), 139–159. https://doi.org/10.1016/j.orggeochem.2004.06.013

Wippler, J., Kleiner, M., Lott, C., Gruhl, A., Abraham, P. E., Giannone, R. J., Young, J. C., Hettich, R. L., & Dubilier, N. (2016). Transcriptomic and proteomic insights into innate immunity and adaptations to a symbiotic lifestyle in the gutless marine worm *Olavius algarvensis*. *BMC Genomics*, *17*(1), 942. https://doi.org/10.1186/s12864-016-3293-y

Wollam, J., & Antebi, A. (2011). Sterol Regulation of Metabolism, Homeostasis, and Development. *Annual Review of Biochemistry*, *80*(1), 885–916. https://doi.org/10.1146/annurev-biochem-081308-165917

Woyke, T., Teeling, H., Ivanova, N. N., Huntemann, M., Richter, M., Gloeckner, F. O., Boffelli, D., Anderson, I. J., Barry, K. W., Shapiro, H. J., Szeto, E., Kyrpides, N. C., Mussmann, M., Amann, R., Bergin, C., Ruehland, C., Rubin, E. M., & Dubilier, N. (2006). Symbiosis

insights through metagenomic analysis of a microbial consortium. *Nature*, *443*(7114), 950–955. https://doi.org/10.1038/nature05192

Yeagle, P. L. (1985). Cholesterol and the cell membrane. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, *822*(3), 267–287. https://doi.org/10.1016/0304-4157(85)90011-5