

Lipidomics of the sea sponge *Amphimedon queenslandica* and implication for biomarker geochemistry

D. A. Gold¹ | S. S. O'Reilly¹ | J. Watson² | B. M. Degnan² | S. M. Degnan² |
J. O. Krömer³ | R. E. Summons¹

¹Department of Earth, Atmospheric and Planetary Science, Massachusetts Institute of Technology, Cambridge, MA, USA

²School of Biological Sciences, University of Queensland, Brisbane, QLD, Australia

³Department for Solar Materials, Helmholtz Center for Environmental Research GmbH - UFZ, Leipzig, Germany

Correspondence

D. A. Gold, Caltech, Pasadena, CA, USA.

Email: dagold@caltech.edu

and

R. E. Summons, MIT, Cambridge, MA, USA.

Email: rsummons@mit.edu

Present address

D. A. Gold, Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA

S. O'Reilly, School of Earth Sciences, University College Dublin, Belfield, Ireland

Funding information

NASA Astrobiology Institute, Grant/Award Number: NNA13AA90A; Agouron Institute Geobiology Fellowship; Marie Skłodowska-Curie Actions Programme; Irish Research Council, Grant/Award Number: ELEVATEPD/2014/47

Abstract

Demosponges are a rich natural source of unusual lipids, some of which are of interest as geochemical biomarkers. Although demosponges are animals, they often host dense communities of microbial symbionts, and it is therefore unclear which lipids can be synthesized by the animal *de novo*, and which require input from the microbial community. To address this uncertainty, we analyzed the lipids of *Amphimedon queenslandica*, the only demosponge with a published genome. We correlated the genetic and lipid repertoires of *A. queenslandica* to identify which biomarkers could potentially be synthesized and/or modified by the sponge. The fatty acid profile of *A. queenslandica* is dominated by an unusual $\Delta^{5,9}$ fatty acid (cis-5,9-hexacosadienoic acid)—similar to what has been found in other members of the *Amphimedon* genus—while the sterol profile is dominated by C_{27} - C_{29} derivatives of cholesterol. Based on our analysis of the *A. queenslandica* genome, we predict that this sponge can synthesize sterols *de novo*, but it lacks critical genes necessary to synthesize basic saturated and unsaturated fatty acids. However, it does appear to have the genes necessary to modify simpler products into a more complex “algal-like” assemblage of unsaturated fatty acids. Ultimately, our results provide additional support for the poriferan affinity of 24-isopropylcholestanes in Neoproterozoic-age rocks (the “sponge biomarker” hypothesis) and suggest that some algal proxies in the geochemical record could also have animal contributions.

1 | INTRODUCTION

Sea sponges—particularly those in the class Demospongiae—are the greatest natural source of diverse lipid compounds, many of which have been studied for their potential medicinal properties and utility as geochemical biomarkers (Thiel et al., 1999, 2002). Unusual lipids discovered in demosponges include C_{23} to C_{31} fatty acids, often with $\Delta^{5,9}$ desaturations—sometimes collectively referred to as demospongiic acids (Kornprobst & Barnathan, 2010)—as well as sterols with atypical side-chain modifications such as 24-isopropylcholesterol (Hofheinz & Oesterhelt, 1979). However, ongoing skepticism exists in treating these lipids as “sponge biomarkers.” Part of this reluctance has

to do with the fact that these lipids, while rare in nature, are not taxonomically restricted to demosponges (Giner, Zhao, Boyer, Satchwell, & Andersen, 2009; Gold, O'Reilly, Luo, Briggs, & Summons, 2016; Kornprobst & Barnathan, 2010). But the greater concern is that sea sponges host dense and complex communities of microbes. It is therefore unclear how many of the lipids recovered in demosponges are synthesized by the animal *de novo*, and which require input from the broader holobiont (a term for the animal and its resident microbes).

In principle, this question could be resolved by comparing the lipids found in demosponges with their repertoire of lipid biosynthesis genes. The demosponge with the best assembled and annotated genome is *Amphimedon queenslandica* from the Great Barrier Reef, Australia

(Srivastava et al., 2010). The fatty acid profiles of several sponges from the genus *Amphimedon* have been analyzed previously and are generally dominated by long-chain 5,9 diunsaturated or 5,9,x triunsaturated fatty acids (Carballeira & Alicea, 2001; Carballeira, Negrón, & Reyes, 1992; Carballeira & Shalabi, 1994; Garson et al., 1993). Unfortunately, genetic data have thrown the taxonomy of *Amphimedon* and its relatives into flux (until recently, *A. queenslandica* was known as *Reniera* sp.), meaning these earlier studies may not be informative for understanding the fatty acid repertoire of *A. queenslandica* (Redmond et al., 2011). So while *Amphimedon* is considered a sponge with a conventional lipid repertoire (Lawson, Stoilov, Thompson, & Djerassi, 1988), recent phylogenetic analyses suggest that some "*Amphimedon*" species might actually be more closely related to genera that make exotic lipids, such as *Petrosia* and *Haliclona* (Redmond et al., 2011). The most recent analysis of fatty acids in definitive specimens of *A. queenslandica* (Watson, Degnan, Degnan, & Krömer, 2014) is from a study that broadly quantified protein, lipid, carbohydrate, and nucleic acid content. However, the structure of many fatty acids and sterols was left unresolved, in part because lipids were not derivatized prior to analysis. Given this uncertainty, a detailed assessment of the genes and lipids of *A. queenslandica* will provide critical insight into the fatty acids that can be synthesized by a sponge de novo.

2 | METHODS

Specimens of *A. queenslandica* were collected from Shark Bay, Heron Island, Australia (23°26'37.92"S, 151°55'8.81"E) on 10/7/2014 using previously established protocols (Leys et al., 2008). Biomass samples were cut from four individual sponges spaced at least 50 m apart, and snap-frozen with liquid nitrogen. These samples were stored at -80°C and transported back to The University of Queensland (St Lucia campus) on dry ice. Total lipids were extracted using hexane:isopropanol (Watson, Brennan, et al., 2014; Watson, Degnan, et al., 2014) and stored in sterile glass sample vials under nitrogen gas at -80°C. Four total lipid extracts from this original study were sent to MIT on 5/17/2015, and subjected to new lipid analyses.

2.1 | Lipid analyses

26 µg of 3-methylheneicosane (ULTRA Scientific, North Kingstown, RI) and 5α-androstan-3β-ol-17-one (epi-androsterone) (Sigma-Aldrich, St Louis, MO) was added to each total lipid extract as internal standards. Total fatty acids were analyzed after methylation/trans-esterification of a 10% aliquot of the total lipid extract using 0.5 M methanolic HCl (Supleco, Bellefonte, PA). Total sterols were analyzed after mild alkaline methanolysis (0.5 M KOH in methanol; 45 min, 50°C) of another 10% portion of the total lipid extract, followed by silylation with 1:1 (v/v) N,O-bis-trifluoroacetamide (containing 1% trimethylchlorosilane) and pyridine (2 h, 70°C). One microliter (internal standard concentration ~13 ng µL⁻¹) of the derivatized extract was injected in splitless mode onto an Agilent 7890A gas chromatograph interfaced with an Agilent 5975C mass selective detector (MSD). The

inlet was equipped with a programmable temperature vaporization (PTV) inlet (Gerstel, Linthicum, MD), which was ramped from 70 to 360°C at a rate of 720°C min⁻¹. Separation was achieved on a J&W DB-5MS fused silica capillary column (Agilent: 60 m × 0.25 mm I.D. and film thickness of 0.25 µm). The MSD source and quadrupole were set at 230 and 150°C, respectively. The column temperature program was as follows: 70°C injection and hold for 2 min, ramp at 10°C min⁻¹ to 130°C, followed directly by a ramp at 4°C min⁻¹ to 310°C, and finally an isothermal hold at 310°C for 20 min. The MS was operated in electron impact mode with an ionization energy of 70 eV and a mass scan range set from *m/z* 50 to 600. All reported compounds were confirmed using a combination of mass spectral libraries, interpretation of mass fragmentation patterns, compound retention times and by comparison with literature. In the full-scan mode we employed, the on-column detection limits of our instrument are routinely in the high to low picogram range. We used 13 ng on-column injections of 3-methylheneicosane and epi-androsterone as internal standards to monitor reproducibility and performance. From our derivatized sample extracts, we have calculated the LOD (3.S/N) to be about 0.15 ng and LOQ (10.S/N) to be about 0.5 ng. The relative abundances of lipids were calculated from peak areas after gas chromatography-flame ionization detection (Agilent 5975C) analysis using the same conditions described above. Double bond locations of monounsaturated fatty acids were confirmed using dimethyl disulfide adduction (Nichols, Guckert, & White, 1986) by treating a portion of the extracts with dimethyl disulfide and iodine solution (6% w/v in diethyl ether) at 50°C for 48 h. The detailed results of lipid analysis are provided in Additional File 1.

2.2 | Genomic analyses

Data on metabolic gene pathways were collected from the Kyoto Encyclopedia of Genes and Genomes (KEGG PATHWAY) (Kanehisa & Goto, 2000). This included *A. queenslandica*-specific data on fatty acid biosynthesis (pathway entry aqu00061), steroid biosynthesis (aqu00100), and biosynthesis of unsaturated fatty acids (aqu01040). Representatives of proteins purportedly missing from *A. queenslandica* were downloaded from KEGG, and compared to the *A. queenslandica* proteome in the National Center for Biotechnology Information (NCBI) refseq database using BLASTp (e-value cutoff 10e-5). Putative *A. queenslandica* genes missing from KEGG PATHWAY were verified using reciprocal BLAST and PFAM vetting of conserved domains.

To determine the homology of *A. queenslandica* desaturase and elovl-like proteins, phylogenetic trees were generated. We queried the PFAM protein family database (<http://pfam.xfam.org/>) for members of the Fatty acid desaturases (Accession: PF00487; ID: FA_desaturase), as well as the GNS1/SUR4 family (Accession: PF01151; ID: ELO), which includes the elovl-like proteins. All peptide sequences from these families were collected for the following organisms: *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Helobdella robusta*, *Lottia gigantea*, *Nematostella vectensis*, *Trichoplax adhaerens*, *Amphimedon queenslandica*, *Salpingoeca rosetta*, *Monosiga brevicollis*, *Sphaeroforma arctica*, *Capsaspora owczarzaki*, *Saccharomyces cerevisiae*, *Aureococcus*

anophagefferens, and *Arabidopsis thaliana*. In both datasets, the resulting sequences were vetted for conserved domains using HMMER (v3.1b2) and the Pfam-A database (Eddy, 2008). The conserved domains were aligned using MUSCLE (Edgar, 2004), and edited by hand to remove identicle sequences and sequence fragments. Maximum likelihood trees were generated from the alignments using RaxML (Stamatakis, 2006) with an LG amino acid substitution model, 4 substitution categories, and 100 bootstraps. The resulting trees created several evolutionarily impossible topologies where *A. queenslandica* proteins were present as both ancestral and daughter sequences, making it difficult to determine the precise homology of these sequences. We therefore optimized the trees using Notung DL (v 2.9; Chen, Durand, & Farach-Colton, 2000), which rearranges poorly supported branches in a gene/protein tree to produce the most parsimonious result based on the known evolutionary relationships between species. All alignments and trees are provided in Additional File 2.

3 | RESULTS

3.1 | *Amphimedon queenslandica* contains fatty acids and sterols common to demosponges, as well as lipids typically identified with algae and bacteria

We identified many fatty acids not described in the initial analysis by Watson, Brennan, et al. (2014), including a number of supposed "algal" and "bacterial" biomarkers. The original study identified hexadecanoic (palmitic) acid ($C_{16:0}$), an unknown fatty acid (UU4-FA), and octadecanoic acid ($C_{18:0}$) as the most abundant fatty acids in *A. queenslandica*. Our analyses suggest that UU4-FA is most likely cis-5,9-hexacosadienoic acid ($C_{26:2}\Delta^{5,9}$), a "demospongiac" acid that represents ~25% of all fatty acids (see Table 1). Dienoic demospongiac acid methyl esters have characteristic ions at m/z 109 (cleavage between carbon 7 and 8 in addition to loss of 32 Da from methanol), 141 (cleavage between carbon 7 and 8) and $[M-49]^+$. Additional "demospongiac" acids recovered include cis-5,9-pentacosadienoic ($C_{25:2}\Delta^{5,9}$), cis-5,9,14-hexacosatrienoic ($C_{26:3}\Delta^{5,9,17}$), heptacosadienoic ($C_{27:2}\Delta^{5,9}$), and octacosadienoic ($C_{28:2}\Delta^{5,9}$) acid. On average, these "demospongiac" acids make up ~31% of total fatty acid abundance in *A. queenslandica*. In addition, we recovered a number of unsaturated fatty acids often considered algal/protist biomarkers, including eicosatetraenoic (arachidonic) ($C_{20:4}\Delta^{5,8,11,14}$), eicosapentaenoic (EPA; $C_{20:5}\Delta^{5,8,11,14,17}$), docosapentaenoic (DPA; $C_{22:5}\Delta^{4,7,10,14,16}$), and docosahexaenoic (DHA; $C_{22:6}\Delta^{4,7,10,14,16,19}$) acid. These compounds collectively represent ~1% of all fatty acids. We also recovered multiple branched fatty acids with a putative bacterial origin, including 9-methyl-hexadecanoic (9Me- $C_{16:0}$), 7-methyl-5-hexadecenoic (7Me- $C_{16:1}\Delta^5$), 15-methyl-hexadecanoic (i $C_{17:0}$), 14-methyl-hexadecanoic (a $C_{17:0}$), 14-methyl-pentadecanoic (i $C_{16:0}$), 13-methyl-tetradecanoic (i $C_{15:0}$), and 12-methyl-tetradecanoic (a $C_{15:0}$) acid. These compounds make up ~7.1% of all the fatty acids. The fatty acid profiles for *A. queenslandica* contrast with those reported for *A. viridis*, *A. complanata*, and *A. compressa*, which were found to have $C_{29:2}\Delta^{5,9,23}$ and $C_{30:2}\Delta^{5,9,23}$ as major long-chain fatty acids (Carballeira & Alicea, 2001; Carballeira &

TABLE 1 Occurrence and average relative abundance (n specimens = 4) of fatty acids from *Amphimedon queenslandica*

Fatty acids		Average \pm range (n = 4)
Tetradecanoic acid	C14:0	1.0 \pm 0.5
13-methyl-tetradecanoic acid	iC15:0	1.4 \pm 1.0
12-methyl-tetradecanoic acid	aC15:0	0.2 \pm 0.2
Pentadecanoic acid	C15:0	0.2 \pm 0.1
14-methyl-pentadecanoic acid	iC16	3.0 \pm 8.4
cis-7-Hexadecenoic acid	C16:1 Δ^7	1.1 \pm 0.1
cis-9-Hexadecenoic acid	C16:1 Δ^9	1.6 \pm 0.4
Hexadecanoic acid	C16:0	14.9 \pm 9.3
9-methyl-hexadecanoic acid	9Me-C16:0	1.3 \pm 0.6
7-methyl-5-hexadecenoic acid	7Me-C16:1 $\Delta^{5,*}$	0.3 \pm 0.3
15-methyl-hexadecanoic acid	iC17:0	0.2 \pm 0.1
14-methyl-hexadecanoic acid	aC17:0	0.6 \pm 0.4
Heptadecanoic fatty acid	C17:0	0.1 \pm <0.1
9-octadecenoic acid	C18:1 Δ^9	2.1 \pm 0.3
11-octadecenoic acid	C18:1 Δ^{11}	1.0 \pm 0.1
Octadecanoic acid	C18:0	8.9 \pm 4.4
Nonadecanoic acid	C19:0	0.3 \pm 0.2
Eicosatetraenoic acid (Arachidonic)	C20:4 $\Delta^{5,8,11,14}$	0.4 \pm 0.3
Eicosapentaenoic acid (EPA)	C20:5 $\Delta^{5,8,11,14,17}$	0.1 \pm <0.1
cis-11-eicosenoic acid	C20:1 Δ^{11c}	0.8 \pm 0.3
trans-11-eicosenoic acid	C20:1 Δ^{11t}	0.2 \pm 0.2
Eicosanoic acid	C20:0	0.2 \pm 0.1
Docosapentaenoic acid	C22:5 $\Delta^{4,7,10,14,16}$	0.3 \pm 0.3
Docosahexaenoic acid (DHA)	C22:6 $\Delta^{4,7,10,14,16,19}$	0.2 \pm 0.1
Docosanoic acid	C22:0	7.1 \pm 2.3
methyl-branched tricosanoic acid	brC22	0.8 \pm 0.2
Tricosanoic acid	C23:0	3.4 \pm 1.1
cis-17-tetracosenoic acid	C24:1 Δ^{17}	3.0 \pm 1.8
Tetracosanoic acid	C24:0	5.1 \pm 1.3
5,9-pentacosadienoic acid	C25:2 $\Delta^{5,9}$	1.6 \pm 0.3
5,9,14-hexacosatrienoic acid	C26:3 $\Delta^{5,9,17}$	0.8 \pm 0.7
5,9-hexacosadienoic acid	C26:2 $\Delta^{5,9}$	25.3 \pm 5.4
15-hexacosenoic acid	C26:1 $\Delta^{15,*}$	1.1 \pm 0.2
17-hexacosenoic acid	C26:1 Δ^{17}	5.9 \pm 1.5
19-hexacosenoic acid	C26:1 Δ^{19}	1.1 \pm 0.6
5,9-heptacosadienoic acid	C27:2 $\Delta^{5,9}$	2.9 \pm 0.7
5,9-octacosadienoic acid	C28:2 $\Delta^{5,9}$	0.1 \pm <0.1
Unknown		1.0 \pm 1.2

Shalabi, 1994; Carballeira et al., 1992). We also failed to detect brominated long-chain fatty acids, which have been reported as major fatty acids for *A. terpenensis* (Garson, Zimmermann, Battershill, Holden, & Murphy, 1994; Garson et al., 1993).

Our analysis of sterols in *A. queenslandica* also includes newly identified compounds for this organism (Table 2). The sterol profile is dominated by cholesterol ($C_{27}\Delta^5$; 36.0% of total sterols on average), as reported in Watson, Brennan, et al. (2014). We identified β -sitosterol ($C_{29}\Delta^5$) as the second-most abundant sterol (~26.3% of total sterols on average), followed by 24-methylcholesta-5,24(28)-dien-3 β -ol ($C_{28}\Delta^{5,24(28)}$; 11.4% of total sterols). Other identified sterols were brassicasterol ($C_{28}\Delta^{5,22}$; ~7.9%), campesterol ($C_{28}\Delta^5$), stigmasterol ($C_{29}\Delta^{5,22}$), fucosterol ($C_{29}\Delta^{5,24(28)}$), and 22-cholest-5,22-dien-3 β -ol ($C_{26}\Delta^{5,22}$). The C_{28} - C_{29} sterols identified here are typically considered biomarkers of plants and algae as opposed to animals like *A. queenslandica* (Volkman et al., 1998). No C_{30} sterols were identified in this study, although we cannot discount the possibility of trace amounts being present, but falling below instrumental detection limits (Love et al., 2009).

3.2 | Genome analysis suggests that “demospongiac” and algal biomarkers—but not bacterial biomarkers—can be synthesized by *Amphimedon queenslandica*

To determine whether *A. queenslandica* has the genes necessary to synthesize the putative algal and bacterial lipids identified in our study, we utilized the KEGG PATHWAY database, which organizes current knowledge on metabolic pathways and records gene presence/absence data across species. KEGG PATHWAY suggests that *A. queenslandica* is missing genes necessary for the biosynthesis of saturated fatty acids (Figure 1A), sterols (Figure 1B), and unsaturated fatty acids (Figure 1C). However, the KEGG database makes its presence/absence determinations based on gene homology assignments against model organisms and does not necessarily account for non-homologous genes with similar enzymatic functions, or for paralog/ortholog nuances that arise during gene duplication events. For example, we previously demonstrated that *A. queenslandica* appears to contain the genes necessary to synthesize complex sterols (Gold, Grabenstatter, et al., 2016) but the KEGG PATHWAY map fails to identify many of these genes because it uses distantly related plant genes as models (Figure 1B). We therefore looked for each putative missing enzyme in the *A. queenslandica* genome to verify its presence or absence.

Regarding saturated fatty acids (Figure 1A), we found no evidence for the required fatty acid synthase-type enzyme in *A. queenslandica*, either the animal-homolog (FASN) or the equivalent enzymes in fungi (Fas1/Fas2) or bacteria (fas). *A. queenslandica* also appears to lack the BCKA decarboxylase enzyme (EC 4.1.1.72) necessary to synthesize the branched-chain fatty acids we recovered that are considered bacterial proxies (Kaneda, 1991; O'Reilly et al., 2016). However, *A. queenslandica* does contain some enzymes necessary for downstream fatty acid elongation, including 3-oxoacyl-[acyl-carrier-protein] synthase (FabF), 3-oxoacyl-[acyl-carrier protein] reductase (FabG), and enoyl-[acyl-carrier protein] reductase (FabI). BLASTp searches against the NCBI protein database (accessed 06/27/2017) provide insight into the evolutionary history of these proteins. The *Amphimedon* FabF (NCBI accession: XP_003384093.1) has high sequence similarity with FabF-like proteins in other animals, while FabI and FabG are most similar to bacterial proteins. The *Amphimedon* FabI sequence (NCBI accession: XP_011407379.2) appears most similar to various Gammaproteobacteria, with the best hit coming from *Thioalkalivibrio*, followed by multiple taxa within the Methylococcales. *Amphimedon* FabG (NCBI accession: XP_019858191.1) is also most similar to *Thioalkalivibrio*, with the next best hits coming from various Acidobacteria and Alphaproteobacteria. These results suggest that FabF was laterally inherited, while FabI and FabG were obtained through horizontal gene transfer with bacteria. However, we caution that these later proteins could also represent contamination in the *A. queenslandica* genome assembly; we initially found a putative FabZ-like protein in *A. queenslandica*, only to have it removed from NCBI after a recent update to the gene models.

Interpreting the genomic data for unsaturated fatty acids is more complicated (Figure 1C), as some of the missing enzymes are members of the large “fatty acid desaturase” and “elongation of very long-chain fatty acids (elovl)” gene families. The *A. queenslandica* genome contains multiple copies of fatty acid desaturase and ELOVL-type proteins, so we constructed phylogenetic trees to determine their evolutionary relationships (Figure 2). *A. queenslandica* has putative homologs to ELOVL3, ELOVL6, and the ELOVL1/2/4/5/7 supergroup (Figure 2B). The divergence of ELOVL1,2,4,5, and 7 appear to represent gene duplications that occurred in bilaterian animals, following their divergence from the sponge lineage. It is therefore plausible that the single *A. queenslandica* protein (Uniprot ID: I1FDN3;

TABLE 2 Occurrence and average relative abundance (n specimens = 4) of sterols from *Amphimedon queenslandica*

Sterols		Average \pm range (n = 4)
cholest-5,22-dien-3 β -ol	cis22- $C_{26}\Delta^{5,22}$	3.9 \pm 1.1
cholest-5-en-3 β -ol (cholesterol)	$C_{27}\Delta^5$	36.0 \pm 8.4
24-methylcholest-5,22-dien-3 β -ol (brassicasterol)	$C_{28}\Delta^{5,22}$	7.9 \pm 2.2
24-methylcholest-5,24(28)-dien-3 β -ol	$C_{28}\Delta^{5,24(28)}$	11.4 \pm 1.7
24-methylcholest-5-en-3 β -ol (campesterol)	$C_{28}\Delta^5$	4.9 \pm 1.0
24-ethylcholest-5,22-dien-3 β -ol (stigmasterol)	$C_{29}\Delta^{5,22}$	4.3 \pm 1.0
24-ethylcholest-5-en-3 β -ol (β -sitosterol)	$C_{29}\Delta^5$	26.3 \pm 2.7
24-ethylcholesta-5,24(28)-dien-3 β -ol (fucosterol)	$C_{29}\Delta^{5,24(28)}$	4.0 \pm 1.5
Unknown C_{29} sterols (n = 3)		1.3 \pm 0.2

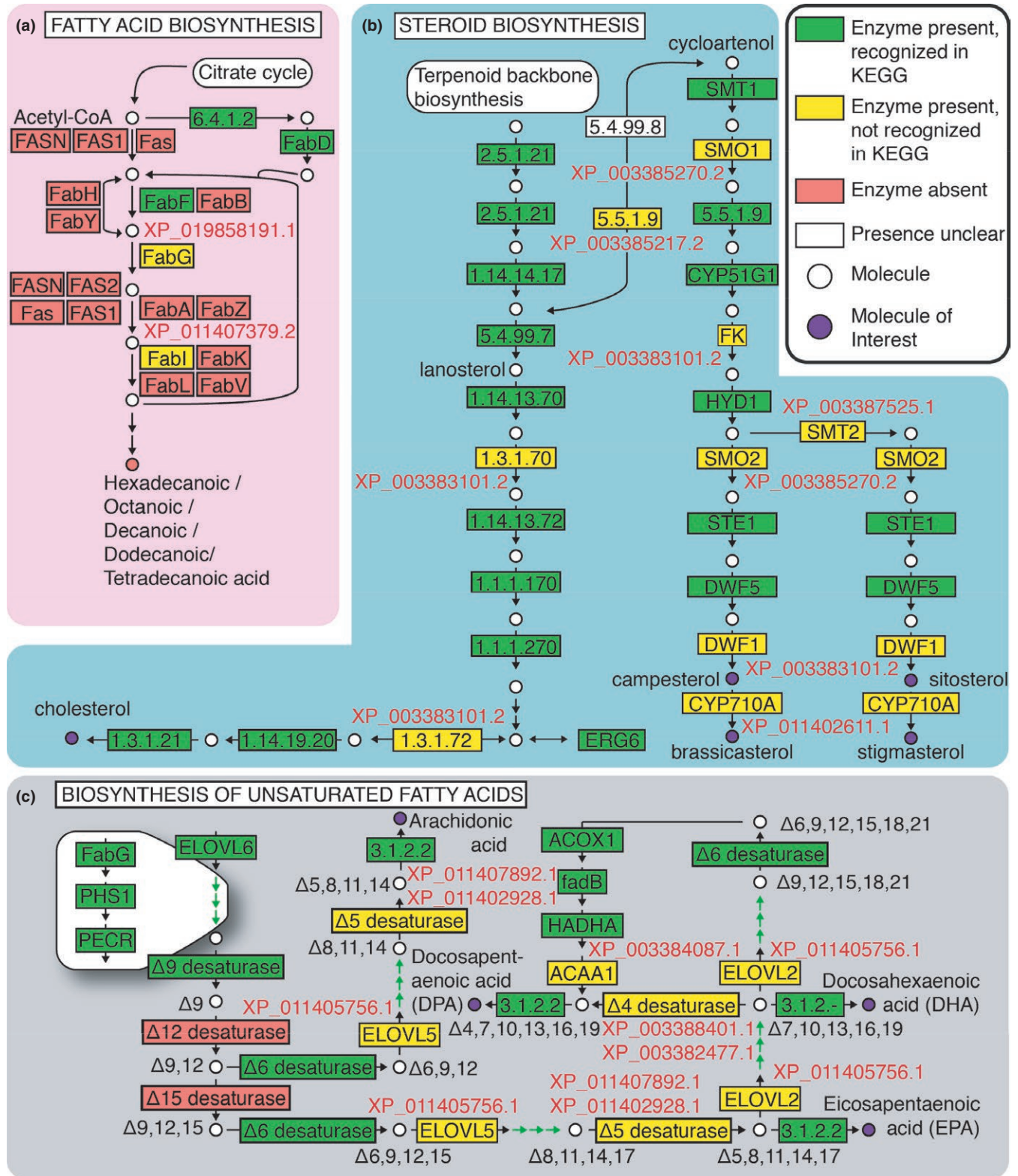


FIGURE 1 The presence and absence of *Amphimedon queenslandica* proteins involved in canonical lipid biosynthesis (based on the KEGG PATHWAY database). In cases where we recovered a putative protein from the *A. queenslandica* genome that is reported as "absent" in KEGG PATHWAY (yellow boxes), the NCBI accession number for the protein(s) is provided in red text. Pathways in this figure are redrawn from the original KEGG pathways: (a) fatty acid biosynthesis (pathway entry aqu00061), (b) steroid biosynthesis (aqu00100), (c) biosynthesis of unsaturated fatty acids (aqu01040) [Colour figure can be viewed at wileyonlinelibrary.com]

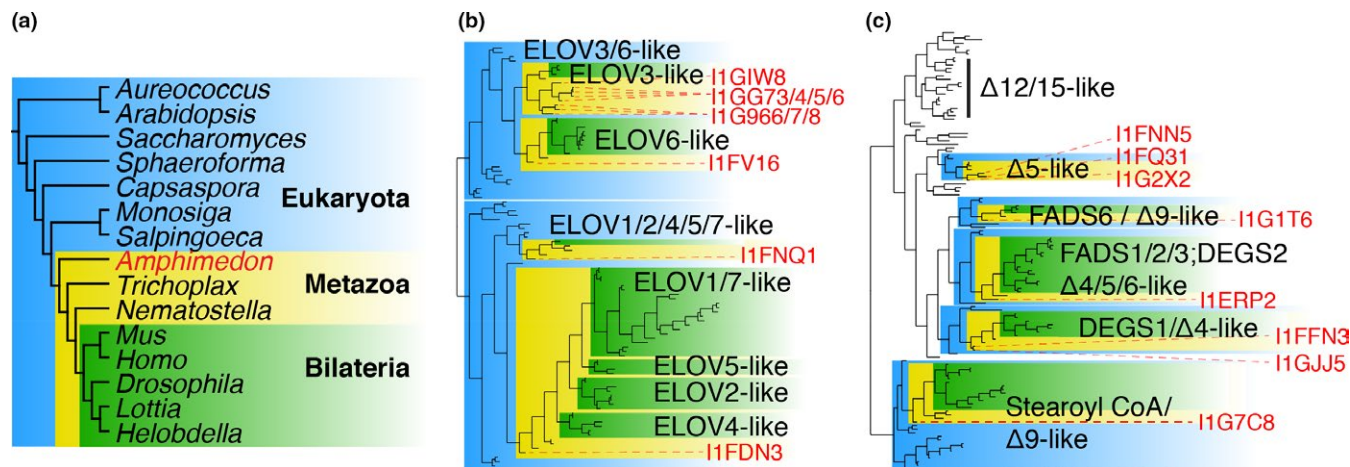


FIGURE 2 The evolution of *Amphimedon queenslandica* fatty acid desaturase and ELOVL proteins. (a) A phylogenetic tree showing the taxa used in this analysis. The colors distinguishing bilaterian animals (green) from non-bilaterian animals (yellow) and other eukaryotes (blue) are used in the protein trees to denote the taxonomy of relevant nodes. (b) A phylogenetic tree showing the evolutionary history of fatty acid desaturases, with *A. queenslandica* proteins noted with red UNIPROT IDs. (c) A phylogenetic tree showing the evolutionary history of ELOVL genes. Note that the trees in (b) and (c) represent maximum likelihood trees that have been optimized using species-tree guidance, as described in the Materials and Methods [Colour figure can be viewed at wileyonlinelibrary.com]

NCBI accession: XP_011405756.2) has functions related to both ELOVL2 and ELOVL5, which play separate roles in bilaterians during the biosynthesis of unsaturated fatty acids (Figure 1C). For the desaturase family (Figure 2C), we recovered putative delta-4, delta-5, delta-6, and delta-9-like desaturases; however, the species lacks a member of the delta-12 fatty acid desaturase superfamily, which includes delta-12 and delta-15 desaturases. Ultimately, most of the genes reported as missing in the KEGG PATHWAY for unsaturated fatty acid biosynthesis do have one or more putative homologs in *A. queenslandica*, with the notable exception of those necessary for delta-12/15 desaturation.

4 | DISCUSSION

Based on our re-vetting of the genome data, we predict that *A. queenslandica* cannot produce basic fatty acids de novo, but that it should be able to modify certain fatty acids into the more complex “demospongiac” and “algal” lipids recovered in our analysis. *A. queenslandica* lacks the enzymes necessary to initiate fatty acid synthesis (i.e., catalyze acetyl-CoA/malonyl-CoA into palmitic acid; Figure 1A), or create the branched fatty acids typically considered indicative of bacteria. It also lacks candidate delta-12 and delta-15 desaturases, suggesting it cannot synthesize basic $\Delta^{9,12}$ or $\Delta^{9,15}$ unsaturated fatty acids de novo (Figure 1C). But if given palmitic acid or $\Delta^{9,12/15}$ unsaturated fatty acids, *A. queenslandica* appears to have the additional enzymes necessary to generate delta-9 and delta-5 modifications typical of demospongiac acids, as well as putative “algal” proxies such as arachidonic acid, EPA, DPA, and DHA.

Amphimedon queenslandica similarly appears to have all genes necessary to synthesize C_{28} and C_{29} sterols de novo. Our analysis cannot definitively say what portion of the sterols found in *A. queenslandica*

is generated by the sponge de novo, and what portion is obtained through diet or association with eukaryotic symbionts. However, the strong uniformity of sterol ratios across four individuals supports the hypothesis that sterol concentrations are actively regulated. This study supports several arguments regarding sponge biomarker genetics proposed in Gold, Grabenstatter, et al. (2016). Firstly, the presence of proteins such as CYP51G1, SMO2, STE1, and DWF1/5 suggests that demosponges actively produce sterols in the cycloartenol pathway, which is largely absent in most animals and fungi. The enzymes used to produce lanosterol (5.4.99.7 in Figure 1C) and cycloartenol (5.4.99.8) are orthologs of the same protein (oxidosqualene cyclase or OSC) so it is unclear if the single enzyme in *A. queenslandica* can produce both products. The amino acid composition of OSC in *A. queenslandica* supports lanosterol synthesis, although the presence of the CPL1 enzyme (5.5.1.9) suggests that this species can convert cycloartenol into lanosterol (Gold, Grabenstatter, et al., 2016). Additionally, the identification of diverse and abundant C_{28} and C_{29} sterols in *A. queenslandica* supports the hypothesis that conventional side-chain modifications can be synthesized by demosponges through the use of a promiscuous 24-sterol *c-methyltransferase* (*smt*) gene (Gold, Grabenstatter, et al., 2016; Gold, O'Reilly, et al., 2016). As *A. queenslandica* lacks a second *smt* paralog, we previously predicted that it cannot produce C_{30} sterols such as the “sponge biomarker” 24-isopropylcholesterol (Gold, Grabenstatter, et al., 2016), and our prediction has been borne out in this analysis.

Some caution and circumspection are necessary in the interpretation of our results. Although *A. queenslandica* appears to have the genes necessary to synthesize C_{28}/C_{29} sterols and unsaturated fatty acids, this does not necessarily mean that the species actually produces these compounds in abundance. Fatty acid biosynthesis in *A. queenslandica* likely involves a combination of biosynthesis and the accumulation of lipid precursors taken up through diet

and/or association with microbial symbionts, as has been found in other sponges (Carballeira, Thompson, Ayanoglu, & Djerassi, 1986; Djerassi & Lam, 1991; Hahn et al., 1988; Hochmuth et al., 2010; Raederstorff, Shu, Thompson, & Djerassi, 1987). The genetic data must also be interpreted with caution: While *A. queenslandica* has enzymes broadly similar in sequence identity to those necessary for sterol and fatty acid production, it is possible that these genes actually represent unidentified paralogs, or have unique functions unrelated to lipid biosynthesis. It is also possible that sponges have evolved unique genes to perform certain enzymatic steps in lipid biosynthesis, as has been found in other eukaryotes (Desmond & Gribaldo, 2009). Ultimately, the ability of *A. queenslandica* to synthesize the lipids identified in this study should be followed up with radiolabeling experiments, and additional demosponges should be analyzed in this manner to determine the enzymatic capabilities of demosponges more broadly.

Despite these uncertainties, this study supports the idea that many “algal” biomarkers in the geological record contain contributions from animal sources, although primary producer inputs are expected to dominate. Our results add to a growing body of evidence suggesting that marine invertebrates are an important and understudied contributor to the marine polyunsaturated fatty acid pool (Monroig, Tocher, & Navarro, 2013). The strong correlation between certain demosponges lineages and unusual sterols (Love et al., 2009) or fatty acids (Schreiber, Wörheide, & Thiel, 2006) suggests that these animals play a critical role in the biosynthesis of such compounds, even if they require precursor input from their diet or the holobiont. It is important to note that the presence of conventional C_{28}/C_{29} steranes in the geologic record will never provide sufficient evidence for sponges; such inputs can only be recognized through unconventional steroid detection, with appropriate self-consistency checks using abundance ratios and demonstration of syngeneity (Love & Summons, 2015). However, our results also suggest that the contribution of sponges to the conventional steroid pool cannot be discounted either.

Ultimately, the implications of this study for understanding sponge contributions to the geochemical record depends on the biomarker class. Polyunsaturated fatty acids rapidly oxidize in the environment, and finding these in environmental samples almost certainly indicates biomass or very fresh planktonic detritus. In other words, these diagnostic fatty acids have little use as biomarkers beyond modern settings. By contrast, C_{28}/C_{29} steranes (the diagenetic alteration products of precursor sterols) in rocks are overwhelmingly algal, as microalgae together with photosynthetic prokaryotes (which do not produce sterols) are the major contributors to modern marine organic matter. However, community ecology and structure have changed significantly over deep time, and sponges might have provided a significant contribution to the environmental lipid signal following mass extinction events (Botting et al., 2017; Ritterbush, Rosas, Corsetti, Bottjer, & West, 2015), or in Neoproterozoic/early Phanerozoic rocks prior to the evolution of some algal lineages. In these scenarios, the possible contribution of sponges to the geochemical record should be considered.

ACKNOWLEDGMENTS

We gratefully acknowledge funding from the NASA Astrobiology Institute (NNA13AA90A) Foundations of Complex Life, Evolution, Preservation, and Detection on Earth and Beyond. DAG acknowledges support from an Agouron Institute Geobiology Fellowship, and SSOR acknowledges support from the Marie Skłodowska-Curie Actions Programme and the Irish Research Council (ELEVATE Postdoctoral Fellowship ELEVATEPD/2014/47).

CONFLICT OF INTEREST

The authors declare no conflict of interests.

REFERENCES

- Botting, J. P., Muir, L. A., Zhang, Y., Ma, X., Ma, J., Wang, L., ... Fang, X. (2017). Flourishing sponge-based ecosystems after the End-Ordovician mass extinction. *Current Biology*, *27*, 556–562.
- Carballeira, N. M., & Alicea, J. (2001). The first naturally occurring α -methoxylated branched-chain fatty acids from the phospholipids of *Amphimedon complanata*. *Lipids*, *36*, 83–87.
- Carballeira, N. M., Negrón, V., & Reyes, E. D. (1992). Novel Monounsaturated Fatty Acids from the Sponges *Amphimedon compressa* and *Mycale laevis*. *Journal of Natural Products*, *55*, 333–339.
- Carballeira, N. M., & Shalabi, F. (1994). Unusual lipids in the Caribbean sponges *Amphimedon viridis* and *Desmaysamma anchorata*. *Journal of Natural Products*, *57*, 1152–1159.
- Carballeira, N., Thompson, J. E., Ayanoglu, E., & Djerassi, C. (1986). Biosynthetic studies of marine lipids. 5. The biosynthesis of long-chain branched fatty acids in marine sponges. *The Journal of Organic Chemistry*, *51*, 2751–2756.
- Chen, K., Durand, D., & Farach-Colton, M. (2000). NOTUNG: a program for dating gene duplications and optimizing gene family trees. *Journal of Computational Biology*, *7*, 429–447.
- 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.
- Djerassi, C., & Lam, W. K. (1991). Phospholipid studies of marine organisms. Part 25. Sponge phospholipids. *Accounts of Chemical Research*, *24*, 69–75.
- Eddy, S. R. (2008). A probabilistic model of local sequence alignment that simplifies statistical significance estimation. *PLOS Computational Biology*, *4*, e1000069.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797.
- Garson, M. J., Zimmermann, M. P., Battershill, C. N., Holden, J. L., & Murphy, P. T. (1994). The distribution of brominated long-chain fatty acids in sponge and symbiont cell types from the tropical marine sponge *Amphimedon terpenensis*. *Lipids*, *29*, 509–516.
- Garson, M. J., Zimmermann, M. P., Hoberg, M., Larsen, R. M., Battershill, C. N., & Murphy, P. T. (1993). Isolation of brominated long-chain fatty acids from the phospholipids of the tropical marine sponge *Amphimedon terpenensis*. *Lipids*, *28*, 1011–1014.
- Giner, J. L., Zhao, H., Boyer, G. L., Satchwell, M. F., & Andersen, R. A. (2009). Sterol chemotaxonomy of marine pelagophyte algae. *Chemistry & Biodiversity*, *6*, 1111–1130.
- Gold, D. A., Grabenstatter, J., de Mendoza, A., 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*, 2684–2689.

- Gold, D. A., O'Reilly, S. S., Luo, G., Briggs, D. E. G., & Summons, R. E. (2016). Prospects for sterane preservation in sponge fossils from museum collections and the utility of sponge biomarkers for molecular clocks. *Bulletin of the Peabody Museum of Natural History*, 57, 181–189.
- Hahn, S., Stoilov, I. L., Ha, T. B. T., Raederstorff, D., Doss, G. A., Li, H. T., & Djerassi, C. (1988). Biosynthetic studies of marine lipids. 17. The course of chain elongation and desaturation in long-chain fatty acids of marine sponges. *Journal of the American Chemical Society*, 110, 8117–8124.
- Hochmuth, T., Niederkrüger, H., Gernert, C., Siegl, A., Taudien, S., Platzer, M., ... Piel, J. (2010). Linking chemical and microbial diversity in marine sponges: possible role for Poribacteria as producers of Methyl-branched fatty acids. *ChemBioChem*, 11, 2572–2578.
- Hofheinz, W., & Oesterhelt, G. (1979). 24-Isopropylcholesterol and 22-Dehydro-24-isopropylcholesterol, novel sterols from a sponge. *Helvetica Chimica Acta*, 62, 1307–1309.
- Kaneda, T. (1991). Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiology and Molecular Biology Reviews*, 55, 288–302.
- Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28, 27–30.
- Kornprobst, J.-M., & Barnathan, G. (2010). Demospongiac acids revisited. *Marine Drugs*, 8, 2569–2577.
- Lawson, M. P., Stoilov, I. L., Thompson, J. E., & Djerassi, C. (1988). Cell membrane localization of sterols with conventional and unusual side chains in two marine demosponges. *Lipids*, 1, 750–754.
- Leys, S. P., Larroux, C., Gauthier, M., Adamska, M., Fahey, B., Richards, G. S., ... Degnan, B. M. (2008). Isolation of Amphimedon developmental material. *Cold Spring Harbor Protocols*, 2008(13), pdb.prot5095–pdb.prot5095.
- Love, G. D., Grosjean, E., Stalvies, C., Fike, D. A., Grotzinger, J. P., Bradley, A. S., ... Summons, R. E. (2009). Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature*, 457, 718–721.
- Love, G. D., & Summons, R. E. (2015). The molecular record of Cryogenian sponges—a response to Antcliffe (2013). *Palaeontology*, 58, 1131–1136.
- Monroig, Ó., Tocher, R. D., & Navarro, C. J. (2013). Biosynthesis of polyunsaturated fatty acids in marine invertebrates: recent advances in molecular mechanisms. *Marine Drugs*, 11, 3998–4018.
- Nichols, P. D., Guckert, J. B., & White, D. C. (1986). Determination of monosaturated fatty acid double-bond position and geometry for microbial monocultures and complex consortia by capillary GC-MS of their dimethyl disulphide adducts. *Journal of Microbiological Methods*, 5, 49–55.
- O'Reilly, S. S., Pentlavalli, P., Flanagan, P. V., Allen, C. C. R., Monteys, X., Szpak, M. T., ... Kelleher, B. P. (2016). Abundance and diversity of sedimentary bacterial communities in a coastal productive setting in the Western Irish Sea. *Continental Shelf Research*, 113, 1–9.
- Raederstorff, D., Shu, A. Y. L., Thompson, J. E., & Djerassi, C. (1987). Biosynthetic studies of marine lipids. 11. Synthesis, biosynthesis, and absolute configuration of the internally branched demospongiac acid, 22-methyl-5,9-octacosadienoic acid. *The Journal of Organic Chemistry*, 52, 2337–2346.
- Redmond, N. E., Raleigh, J., van Soest, R. W. M., Kelly, M., Travers, S. A. A., Bradshaw, B., ... McCormack, G. P. (2011). Phylogenetic relationships of the Marine Haplosclerida (Phylum Porifera) employing Ribosomal (28S rRNA) and Mitochondrial (cox1, nad1) gene sequence data. *PLoS ONE*, 6, e24344.
- Ritterbush, K. A., Rosas, S., Corsetti, F. A., Bottjer, D. J., & West, A. J. (2015). Andean sponges reveal long-term benthic ecosystem shifts following the end-Triassic mass extinction. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 420, 193–209.
- Schreiber, A., Wörheide, G., & Thiel, V. (2006). The fatty acids of calcareous sponges (Calcarea, Porifera). *Chemistry and Physics of Lipids*, 143, 29–37.
- Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T., ... Rokhsar, D. S. (2010). The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature*, 466, 720–726.
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22, 2688–2690.
- Thiel, V., Blumenberg, M., Hefter, J., Pape, T., Pomponi, S., Reed, J., ... Michaelis, W. (2002). A chemical view of the most ancient metazoa-biomarker chemotaxonomy of hexactinellid sponges. *Naturwissenschaften*, 89, 60–66.
- Thiel, V., Jenisch, A., Wörheide, G., Löwenberg, A., Reitner, J., & Michaelis, W. (1999). Mid-chain branched alkanolic acids from "living fossil" demosponges: a link to ancient sedimentary lipids? *Organic Geochemistry*, 30, 1–4.
- Volkman, J. K., Barrett, S. M., Blackburn, S. I., Mansour, M. P., Sikes, E. L., & Gelin, F. (1998). Microalgal biomarkers: a review of recent research developments. *Organic Geochemistry*, 29, 1163–1179.
- Watson, J., Brennan, T., Degnan, B., Degnan, S., & Krömer, J. O. (2014). Analysis of the biomass composition of the Demosponge *Amphimedon queenslandica* on Heron Island Reef, Australia. *Marine Drugs*, 12, 3733–3753.
- Watson, J., Degnan, B., Degnan, S., & Krömer, J. O. (2014). Determining the biomass composition of a sponge holobiont for flux analysis. In J. O. Krömer, L. K. Nielsen, & L. M. Blank (Eds.), *Methods in molecular biology*, Vol. 1191 (pp. 107–125). New York, NY: Springer, New York.

How to cite this article: Gold DA, O'Reilly SS, Watson J, et al. Lipidomics of the sea sponge *Amphimedon queenslandica* and implication for biomarker geochemistry. *Geobiology*. 2017;15:836–843. <https://doi.org/10.1111/gbi.12253>