Title

Sterol and genomic analyses validate the sponge biomarker hypothesis

Journal

Proceedings of the National Academy of Sciences of the United States of America, 113(10)

ISSN

0027-8424

Authors

Gold, David A
Grabenstatter, Jonathan
de Mendoza, Alex
et al.

Publication Date

2016-03-08

DOI

10.1073/pnas.1512614113

Peer reviewed

Sterol and genomic analyses validate the sponge biomarker hypothesis

David A. Gold[a], Jonathan Grabenstatter[a], Alex de Mendoza[b], Ana Riesgo[c], Iñaki Ruiz-Trillo[b,d], and Roger E. Summons[a,1]

[a]Department of Earth, Atmospheric and Planetary Science, Massachusetts Institute of Technology, Cambridge, MA 02139; [b]Institut de Biologia Evolutiva (Consejo Superior de Investigaciones Científicas-Universitat Pompeu Fabra), 08003 Barcelona, Spain; [c]Department of Life Sciences, Natural History Museum of London, London SW7 5BD, United Kingdom; and [d]Institut Català de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain

Edited by Katherine H. Freeman, Pennsylvania State University, University Park, PA, and approved January 21, 2016 (received for review June 26, 2015)

Molecular fossils (or biomarkers) are key to unraveling the deep history of eukaryotes, especially in the absence of traditional fossils. In this regard, the sterane 24-isopropylcholestane has been proposed as a molecular fossil for sponges, and could represent the oldest evidence for animal life. The sterane is found in rocks ∼650–540 million y old, and its sterol precursor (24-isopropylcholesterol, or 24-ipc) is synthesized today by certain sea sponges. However, 24-ipc is also produced in trace amounts by distantly related pelagophyte algae, whereas only a few close relatives of sponges have been assayed for sterols. In this study, we analyzed the sterol and gene repertoires of four taxa (*Salpingoeca rosetta*, *Capsaspora owczarzaki*, *Sphaeroforma arctica*, and *Creolimax fragrantissima*), which collectively represent the major living animal outgroups. We discovered that all four taxa lack C_{30} sterols, including 24-ipc. By building phylogenetic trees for key enzymes in 24-ipc biosynthesis, we identified a candidate gene (*carbon-24/28 sterol methyltransferase*, or *SMT*) responsible for 24-ipc production. Our results suggest that pelagophytes and sponges independently evolved C_{30} sterol biosynthesis through clade-specific *SMT* duplications. Using a molecular clock approach, we demonstrate that the relevant sponge *SMT* duplication event overlapped with the appearance of 24-isopropylcholestanes in the Neoproterozoic, but that the algal *SMT* duplication event occurred later in the Phanerozoic. Subsequently, pelagophyte algae and their relatives are an unlikely alternative to sponges as a source of Neoproterozoic 24-isopropylcholestanes, consistent with growing evidence that sponges evolved long before the Cambrian explosion ∼542 million y ago.

sponges | Porifera | sterols | steranes | Amorphea

Sterols represent a class of lipid molecules critical to the physiology of eukaryotic cells, thereby providing valuable insight into the evolution of life. With few exceptions (1, 2), sterols are exclusive to eukaryotes and are involved in diverse cellular functions, including membrane structure and fluidity, developmental regulation, and as precursors to signaling and hormone molecules. All sterols share a common structure consisting of a tetracyclic cyclopenta[a]phenanthrene nucleus and a side chain bound to carbon 17 (Fig. 1*A*). The basic sterol of animals—cholesterol—consists of 27 carbons (or C_{27}), but modifications to the nucleus and/or side chain allow for a diversity of structures, typically ranging from C_{26}–C_{31}. Some of the most exotic sterols are restricted to particular eukaryotic lineages, and because sterols (diagenetically altered into steranes) are stable through deep geological time, they can function as "molecular fossils," recording the evolution of organisms even in the absence of physically preserved fossils (3).

One presently debated molecular fossil is 24-isopropylcholestane, a sterane found in abundance in certain Neoproterozoic to Early Cambrian rocks ∼650–540 million y old (4). This sterane is a diagenetic product of the C_{30} sterol 24-isopropylcholesterol (24-ipc), one of many rare or unique sterols produced by a subset of modern sea sponges within the clade Demospongiae (4). This sterane has been widely accepted as a "sponge biomarker" and as the oldest evidence for animals in the geologic record. Subsequently, this biomarker is commonly used as a calibration point when estimating molecular clocks (5–8) and in the interpretation of Precambrian fossils and geology (9–11). However, several recent papers have challenged the sponge affinity of this biomarker (12, 13), arguing that (*i*) pelagophyte algae also produce 24-ipc, meaning they or their ancestors could be responsible for the sterane, and (*ii*) there is a general lack of information about the distribution of C_{30} sterols within the eukaryotes. Accordingly, the geological interpretation of this sterane deserves further investigation.

Resolving questions surrounding 24-ipc necessitates a better understanding of how the underlying sterol biosynthesis pathway has evolved, as well as the distribution of sterols in several critical, understudied eukaryote lineages. Pelagophyte algae and demosponges are distantly related, which suggests that they evolved the ability to produce 24-ipc independently. Additionally, bikonts (such as plants, diatoms, and algae) typically use cycloartenol as their biosynthetic protosterol, whereas most amorpheans (also known as unikonts, which include fungi and animals) use lanosterol, further suggestive of convergent evolution for 24-ipc (14). Within the Amorphea, sterols have been especially well-studied in the fungi, which primarily produce ergosterol (C_{28}), and animals, which—with the exception of sponges—primarily produce cholesterol (C_{27}) (14). But the clades nested between the fungi and animals remain poorly sampled. As sponges represent the earliest or one of the earliest branching animal lineages, elucidating the sterols of animal outgroups remains central to continued testing of the sponge biomarker hypothesis.

Significance

An unusual molecule is found in rocks ∼650–540 million y old, and its likely precursor, 24-isopropylcholesterol (24-ipc), is produced by some modern sea sponges. The sterane hydrocarbon analog of 24-ipc offers a potential "molecular fossil" for early animals, but certain algae also produce traces of this molecule, so it is unclear when and how frequently the ability to synthesize 24-ipc evolved. In this study, we connect 24-ipc production to a gene and conclude that algae and sponges independently evolved 24-ipc synthesis through unique gene duplication events. Although the timing of the sponge gene duplication overlaps with the geological record of the molecular fossil, the algal gene duplication occurs significantly later, supporting the connection of 24-ipc to sponges and providing the oldest evidence for animal life.

Author contributions: D.A.G. and R.E.S. designed research; D.A.G., J.G., A.d.M., A.R., and I.R.-T. performed research; D.A.G. and R.E.S. analyzed data; and D.A.G. and R.E.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: rsummons@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1512614113/-/DCSupplemental.

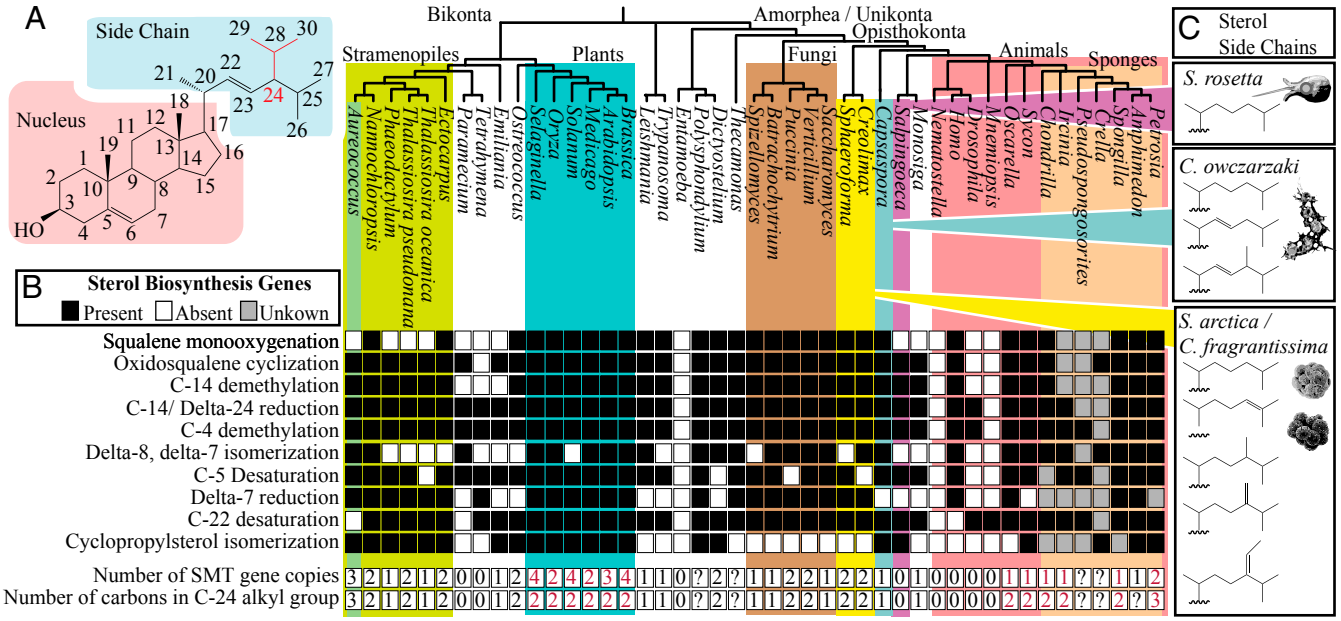

Fig. 1. (*A*) Structure of 24-ipc, illustrating the canonical carbon numbering system for sterols. The isopropyl group on carbon 24 is highlighted in red. (*B*) Distribution of sterol synthesis genes across eukaryotes. This phylogeny represents a consensus tree based on previous phylogenetic studies (6, 20, 35). In some cases, multiple genes have been combined into one category based on shared enzymatic function. *SMT* gene copy numbers and sterol alkylation are provided at the bottom. (*C*) Sterol side chains from taxa analyzed in this study. See Tables S1 and S2 for relevant data and references.

In this project, we investigated four amorpheans for their sterol repertoires: the choanoflagellate *Salpingoeca rosetta*, the filasterean *Capsaspora owczarzaki*, and the ichthyosporeans *Sphaeroforma arctica* and *Creolimax fragrantissima*. These taxa were chosen because they collectively represent the three major animal outgroups (Fig. 1*B*) and because each has a sequenced genome, allowing us to correlate sterol and gene repertoires. Consistent with lipid evidence, genetic data suggest that pelagophytes and sponges independently evolved the ability to produce C_{30} sterols using clade-specific duplications of the gene *carbon-24/28 sterol methyltransferase* (*SMT*; Enzyme Commission 2.1.1.41). Using a molecular clock approach, we demonstrate that the relevant sponge *SMT* duplication event overlaps with the appearance of 24-isopropylcholestanes in the Neoproterozoic, but that the algal *SMT* duplication event occurs in the Phanerozoic, at least 100 million y after the oldest evidence for 24-isopropylcholestanes.

Results and Discussion

Among Amorpheans, C_{30} Sterols Are Restricted to the Sponges. The four species analyzed posses a diverse array of sterols (see *SI Methods: Sterol Analysis*, Table S1, and Fig. S1 for full sterol analysis), but none contains any C_{30} sterols generated by side-chain modification (Fig. 1*C*). All side chain modifications to the sterols we identified occur at C-24. The icthyosporeans *S. arctica* and *C. fragrantissima* add both methyl and ethyl groups to C-24, generating a range of C_{28} and C_{29} sterols (*SI Methods: Sterol Analysis*). *C. owczarzaki* also generates C_{28} sterols through C-24 methylation. We found no evidence for C-24 modification in *S. rosetta*. Combined with previous data from the choanoflagellate *Monosiga brevicollis* (15), our results suggest that C_{30} sterols are absent from nonsponge amorpheans, and that their synthesis either represents a derived trait in the sponges or an ancient trait lost in at least six major amorphean lineages (amoebozoans, fungi, choanoflagellates, ichthyosporeans, filastereans, and nonsponge animals).

A Strong Correlation Exists Between C-24 Sterol Alkylation and *SMT* Gene Copy Number, Except for Sponges. Finding no evidence for C_{30} sterols among animal outgroups, we next analyzed a series of genomes for enzymes involved in sterol biosynthesis. The genome of one pelagophyte (*Aureococcus anophagefferens*) and one demosponge (*Amphimedon queenslandica*) are publically available and were included in our study. Like other pelagophytes, *A. anophagefferens* produces trace amounts of 24-ipc (<0.05%) during 24-propylidenecholesterol synthesis (16). In contrast, a recent study on *A. queenslandica* identified cholesterol (C_{27}), brassicasterol (C_{28}), and an unidentified sterol as its major components (17). Not all demosponges produce C_{30} sterols; in fact, sponges produce a wide array of common and exotic sterols, which only show moderate phylogenetic conservation (18, 19). Isopropylcholesterols, for example, are most common in the order Halichondrida, but can be found sporadically across all major demosponge clades (4). Because *A. queenslandica* lacks C_{30} sterols, we also analyzed the transcriptomes from six additional demosponges (*Crella elegans*, *Ircinia fasciculata*, *Chondrilla nucula*, *Pseudospongosorites suberitoides*, *Spongilla lacustris*, and *Petrosia ficiformis*) (20) as well as genomes from the homoscleromorph sponge *Oscarella carmela* and calcareous sponge *Sycon ciliatum*. This larger dataset includes sponges that produce rare C_{29} and C_{30} sterols (see Fig. 1*B*). In particular, the genus *Petrosia* generates a variety of exotic C_{30} sterols through the opening of a cyclopropyl ring on C-24, with *n*-propyl groups having previously been identified in *P. ficiformis* and isopropyl groups in *Petrosia weinbergi* (21). Thus, our collection of sponge data encompasses the diverse range of sterol production in this clade.

The most relevant discovery afforded by our survey of sterol biosynthesis genes (Fig. 1*B*; see *Results of Our Survey on Sterol Biosynthesis Gene Across Eukaryotes* and Fig. S2 for discussion) is the strong correlation between C-24 side chain alkylation and the number of copies of the gene *SMT* in the genome. In the four taxa studied here, we found that the number of carbons added during C-24 alkylation correlates with the number of *SMT* homologs. *S. rosetta* lacks an *SMT* gene or C-24 alkylation, *C. owczarzaki* has a single *SMT* and exhibits C-24 methylation, whereas *S. arctica* and *C. fragrantissima* each have two *SMTs* and contain both methylated and ethylated C-24s. This suggests that one *SMT* gene is responsible for the first round of methylation of

C-24, whereas the second copy controls methylation of C-28, generating the ethyl group. The correlation between *SMT* copy number and side chain modification appears broadly applicable across eukaryotes (Fig. 1*B*). In all cases, the number of *SMT* genes is at least as great as the number of C-24 alkylation steps needed to make the organism's largest sterol. In higher plants, the number of *SMT* genes is greater than the number of carbons added to C-24; the taxa we queried generally have three or four *SMT* copies, but only produce common methyl- and ethyl-group modifications. This finding is consistent with evidence for multiple rounds of whole-genome duplication in plants and is recapitulated in the high plant paralog counts for all genes we analyzed (*Supporting Information*). The overabundance of plant *SMT*s suggests that they exhibit overlapping functions that are not necessarily related to sterol synthesis, a hypothesis supported by evidence from *Arabidopsis thaliana* (22). However, the tight correlation holds for the rest of the bikonts, including a unique third *SMT* in the pelagophyte *A. anophagefferens*. This correlation strongly suggests that most eukaryotes—including pelagophytes—require separate *SMT* genes to perform each round of C-24 sterol modification.

Although this correlation between *SMT* copy number and C24 alkylation is strong across the eukaryotes, it does not hold for sponges. Instead, we consistently recovered fewer *SMT* copies than expected for sponges that produce C_{29} and C_{30} sterols. We did not recover *SMT*s from the demosponges *C. elegans* or *P. suberitoides*. This could represent insufficient depth of transcriptome sequencing but could also be genuine, as both species appear to be missing most other genes involved in sterol synthesis (Fig. 1*B*), their sterol repertoires have not been analyzed, and not all demosponges can synthesize sterols de novo (23). In contrast, the demosponges *C. nucula*, *I. fasciculata*, and *S. lacustris* all produce C_{29} sterols with C-24 ethyl moieties, yet we recovered a single *SMT* from each species. *P. ficiformis*, which produces C_{30} sterols, has two putative *SMT*s (hereafter referred to as *PfiSMT1* and *PfiSMT2*), which is again one less than predicted. There are several ways to interpret this discrepancy. One possibility is that our findings represent an artifact caused by the incompleteness of the sponge transcriptomes. However, genomic data from calcareous and homoscleromorph sponges support the pattern we find in the demosponges. For example, although the calcareous sponge *S. ciliatum* has a single *SMT*, species of *Sycon* tested for sterols contain methylated and ethylated C-24s (24). Similarly, although the homoslceromorph *O. carmela* has not been assayed for sterols, C-24 ethyl and methyl moieties have been identified in sterols from *Oscarella lobularis* (25). All sterols reported from this species contain an unusual 5α-hydroxy-6-keto-Δ^7 nucleus, which strongly suggests that they are modified by the animal (it is unclear, however, whether the sterol core is biosynthesized de novo or obtained from diet). An alternate interpretation for the discrepancy is that the gene counts are accurate, but exotic "sponge" sterols are actually produced by symbionts. But here the transcriptomes are informative, as *SMT* genes from putative eukaryotic symbionts living in sponge tissues should have been recovered alongside the sponge genes. This interpretation is also contradicted by significant research demonstrating that demosponges perform de novo sterol synthesis and side chain alkylation (23, 26) and that sterol patterns do not differ dramatically within species collected from different locations or times (19). The protocols used to generate the transcriptomes involved the enrichment of eukaryotic mRNAs (20), meaning that genes from prokaryotic symbionts could have been lost. This is relevant because it has been suggested that a bacterial symbiont of demosponges (phylum Poribacteria) has an *SMT* (27). However, this observation was the result of genome assembly error, and newer versions of the poribacterian genome no longer contain this sequence (see *A Note on the Putative Candidatus Poribacteria SMT Gene*), meaning there is no evidence for *SMT*s in prokaryotes. Ultimately, neither incomplete sequencing nor symbiont contamination can sufficiently explain the pattern we see in the sponge data.

A third interpretation of our discrepancy, and the one that we prefer, is that sponge *SMT*s are capable of catalyzing multiple rounds of C-24 methylation. This concept is known as enzyme promiscuity and has been demonstrated in other methyltransferases (28, 29). To test our hypothesis, we performed protein structure and function prediction using the I-TASSER server (30), comparing our sponge proteins against a typical *SMT* from the yeast *Saccharomyces cerevisiae* (*ERG6*) and a promiscuous *SMT2* from the plant *A. thaliana*. *ERG6* strictly performs a single C-24 methylation in yeast, producing the C_{28} ergosterol, but if *ERG6* is transgenically replaced with *SMT2*, C-24 methyl and ethyl sterols are produced, meaning *SMT2* is sufficient to perform both alkylation steps (31). Although there is no crystal structure of *SMT* to compare our models against, I-TASSER correctly predicts the presence of seven central parallel β-sheets surrounded by 10 α-helices (32) as well as ligand binding sites for S-adenosyl–modified amino acids and active enzyme sites for methyltransferase activity (Fig. 2*A*). In support of the yeast manipulation experiments cited above, protein modeling suggests that *ERG6* contains a single active methyltransferase site residue at isolucene-194, whereas *SMT2* has two predicted binding sites, isolucene-198 and glutamic acid-209 (Fig. 2 *A* and *B*). Similar to *SMT2*, *SMT*s from sponges also contain two predicted methyltransferase sites, and despite the fact that sponge *SMT*s are more closely related to *ERG6* (Fig. S3), their 3D structure is more similar to *SMT2* (data for *I. fasciculata* are shown in Fig. 2 *C* and *D*). This supports our hypothesis that a single sponge *SMT* should be capable of multiple rounds of C-24 alkylation and could be partially responsible for the unprecedented diversity of sterols found in this clade.

If our interpretation of the genomic data is correct, then we can present a working hypothesis for the evolution of 24-ipc production in eukaryotes. Sterol synthesis, including C-24 alkylation, was present in the last common ancestor of eukaryotes (as articulated in ref. 14). It is unclear from the distribution of SMT genes whether one or two copies existed in the last common ancestor of eukaryotes. Regardless, several additional events appear critical to 24-ipc evolution: (*i*) a third *SMT* gene evolved in the bikont lineage leading to pelagophytes, (*ii*) a

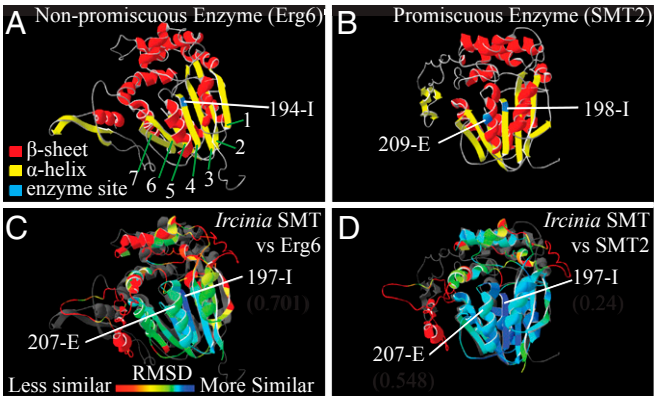

Fig. 2. Protein modeling of *SMT* genes. (*A* and *B*) Predicted structure of *S. cerevisiae Erg6* (*A*) and *A. thaliana SMT2* (*B*); α-helices are colored red, β-sheets yellow, and predicted methyltransferase sites blue. The seven central β-sheets are numbered in *A*. (*C* and *D*) *I. fasciculata*'s *SMT* superimposed on *Erg6* (*C*) and *SMT2* (*D*). The proteins were aligned in the Swiss-PdbViewer, using the isoleucine active site and its two surrounding peptides as a guide. Coloration signifies the root-mean-square distance (RMSD) between the superimposed structures.

single *SMT* gene promiscuously capable of generating C_{28} and C_{29} sterols evolved in the sponges, and (*iii*) a second *SMT* evolved in sponge lineages that produce C_{30} sterols, such as *P. ficiformis*. Knowing where and when these evolutionary events occurred in the eukaryote tree should resolve the question of which other organisms, if any, might be responsible for Neoproterozoic 24-isopropylcholestanes.

Molecular Clock Analyses Suggest That Sponge, but Not Algal, C_{30} Sterol Synthesis Evolved Before the Phanerozoic. To test the timing of these hypothesized *SMT* duplication events, we used a molecular clock approach to create a time-calibrated gene tree. Because of uncertainty in the eukaryote phylogeny as well as long-branch attractions in underrepresented clades, we restricted this analysis to sponges, fungi, archaeplastids (plants plus red and green algae), and stramenopiles (other algae—including pelagophytes—and diatoms). Species tree error correction was applied to the *SMT* gene tree, allowing us to differentiate orthologs from paralogs (33) (see *Supporting Information* and Figs. S3 and S4 for details). The error-corrected *SMT* tree suggests that the last common ancestor of eukaryotes had a single *SMT* and that independent duplication events have occurred in all major clades (Fig. S4). In this corrected tree, *Petrosia*'s second SMT (*PfiSMT2*) is placed as sister to the other *SMT*s in the "G3+G4" sponge clade (which includes *PfiSMT1* and *SMT*s from *A. queenslandica* and *S. lacustris*). However, 24-ipcs can also be found in sponges from the sister "G2" and "G1" clades (represented in our study by *C. nucula* and *I. fasciculata*, respectively) (4). Subsequently, a gene duplication event at the base of the G3+G4 clade cannot be solely responsible for demosponge 24-ipc synthesis. Assuming this second *SMT* is in fact necessary for demosponge C_{30} sterol production, the duplication event that led to *PfiSMT2* must predate the split of crown-group demosponges, or else independent gene duplication events occurred in multiple demosponge lineages. These two competing hypotheses of gene evolution are illustrated in Fig. 3*A*. Given this uncertainty, we tried a series of topologies for our molecular clock analysis, including the two topologies in Fig. 3*A*, as well as removing all topology priors for the sponges.

The results from the molecular clock analyses are provided in Table 1, with one analysis illustrated in Fig. 3*B*. Large error bars are present on most nodes—as expected given the limited amount of sequence data within the *SMT* gene—but all analyses support the same general conclusion. Consistent with species-level molecular clocks (8), demosponge *SMT*s diverged in the Neoproterozoic. In all analyses, the 95% confidence interval for the origin of *PfiSMT2* overlaps with the age range for Neoproterozoic 24-isopropylcholestanes (Table 1). By contrast, the gene duplication event that led to the algal *SMT2*/*SMT3* split fails to overlap with Neoproterozoic 24-isopropylcholestanes. All analyses suggest a Phanerozoic origin for this event, and even the oldest estimates postdate the oldest evidence of 24-isopropylcholestanes by over 100 million y.

Conclusion

Our study provides evidence against the two most pressing challenges to the sponge biomarker hypothesis: that pelagophytes or

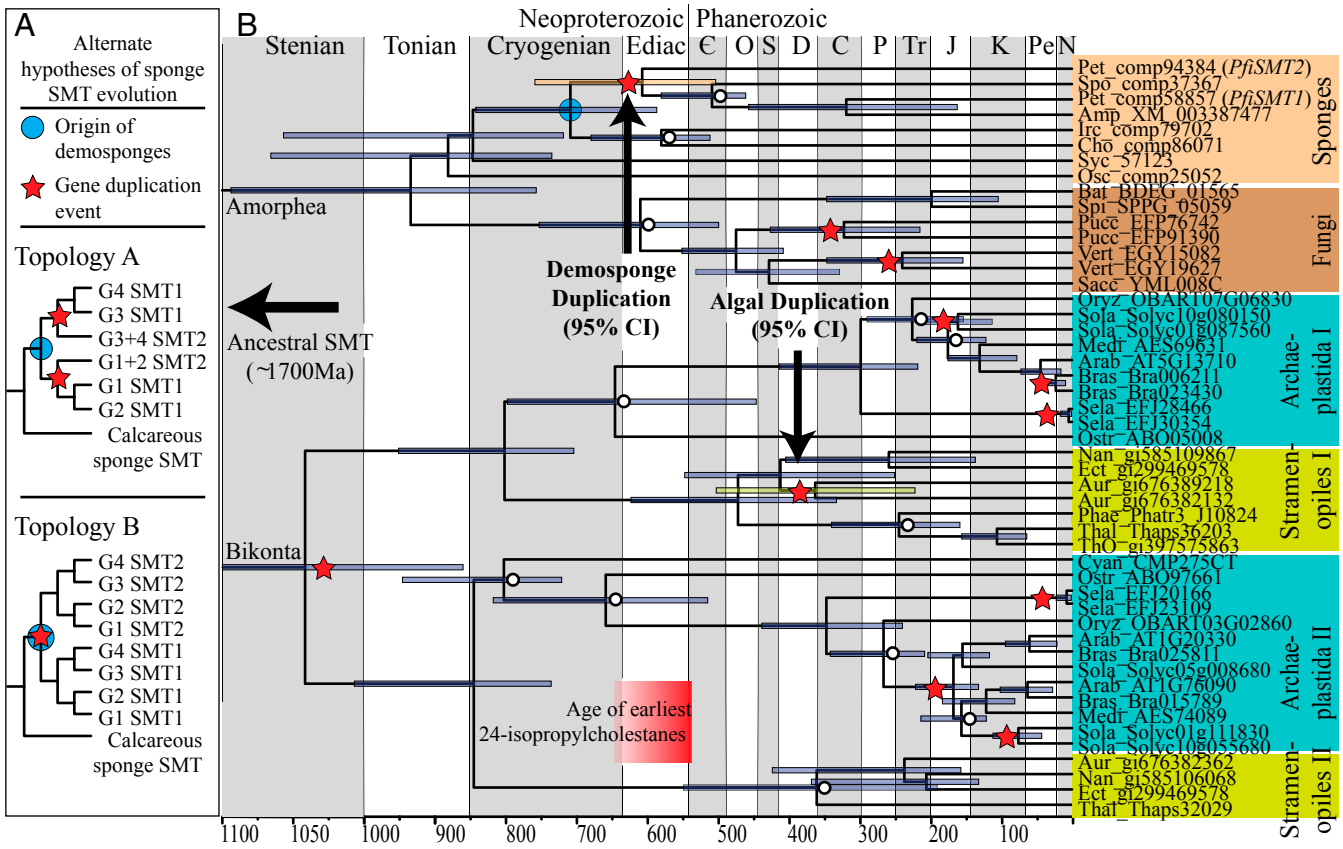

Fig. 3. (*A*) Two competing hypotheses for the origin of the second demosponge SMT. Topology A is the prefered methodology we used to discern orthologs from paralogs, but topology B is also consistent with our data. (*B*) Molecular clock of SMT divergence times, based on topology A with a long root prior (see Table 1 and Dataset S1 for all time-calibrated trees). Gene duplication events are labeled with stars. The gene duplication events relevant to sponge and pelagophyte C_{30} sterol synthesis are noted with arrows. The age estimate of Neoproterozoic 24-isopropylcholestane is provided in the red box. White circles indicate calibration points, which are described in Tables S3.

Table 1. Age ranges (in millions of years) for relevant gene duplication events according to different topology priors and the discrepancy between the origin of algal *SMT3* and the oldest biomarker evidence

Topology prior	Origin of demosponge *SMT*	Origin of demosponge SMT2, *PfiSMT2*	Origin of algal *SMT3*	Discrepancy between algal *SMT3* and oldest sterane
Topology A	560–747	499–668	264–520	162
Topology B	608–771	608–771	264–520	130
No topology prior	561–750	472–645	124–514	136
Topology A + long root	587–843	504–760	222–503	147
Topology B + long root	605–937	605–937	37–452	198

Ranges represent 95% confidence intervals. Topologies A and B refer to the evolutionary scenarios shown in Fig. 3*A*, and "no topology prior" means all sponge-related priors were removed from the analysis. "Long root" analyses fix the ancestral *SMT* at 1679–1866 Ma, consistent with estimates by ref. 7 of unikont/bikont divergence dates. Notice that in all scenarios, the origin of *PfiSMT2* falls within the estimated age range of the Neoproterozoic 24-isopropylcholestanes. In contrast, all predictions for the origin of algal *SMT3* occur more than 100 million y after the oldest evidence of 24-isopropylcholestanes.

their ancestors could be responsible for Neoproterozoic 24-isopropylcholestane and that C_{30} sterol synthesis evolved in amorpheans before the sponges. The tight linkage between *SMT* copy number and C-24 alkylation strongly suggests that pelagophytes such as *Aureococcus* require three *SMT* genes to generate C_{30} sterols. However, the evolution of the third algal *SMT* did not occur until the Phanerozoic, demonstrating that neither pelagophytes nor their ancestors are viable candidates for Neoproterozoic C_{30} steranes. A Phanerozoic origin for this gene duplication event is consistent with the absence of C_{30} sterols in all other studied stramenopiles (34) as well as molecular clock analyses that suggest a Mesozoic to Cenozoic origin for crown-group pelagophytes (35). Our results similarly demonstrate that pelagophytes are unlikely to be responsible for Neoproterozoic 24-*n*-propylcholestanes, a biomarker that has generally been attributed to this group. However, 24-*n*-propylidenecholesterol is also found in sponges as well as the foraminiferan *Allogromia laticollaris* (36), which is part of the "Monothalamid" clade thought to have Neoproterozoic origins (37, 38). The late emergence of algal 24-*n*-propylidenecholesterol (24-npc) is also consistent with the geological record. Although the C_{30} sterane contents of petroleum and source rocks remain very low in the Cambrian to Early Ordovician, the ratio of 24-isopropylcholestanes to 24-*n*-propylcholestanes remains high (39–41), consistent with a primary source from sponges. A rise in 24-*n*-propylcholestane contents thereafter suggests that algae sources appeared between the Early Ordovician and Devonian (40). We therefore conclude that sponges and/or foraminiferans are better candidates for 24-*n*-propylcholestane in the Neoproterozoic, which requires a reappraisal as to whether "npc/ipc" ratios are necessary for identifying sponge biomarkers.

Despite strong evidence that sponges are responsible for Neoproterozoic 24-isopropylcholestanes, the precise genetic history leading to sponge 24-ipc synthesis remains unclear and will require additional research. Our data suggest that a single sponge *SMT* is sufficient to generate a variety of C_{28}–C_{29} sterols. Such promiscuity could help explain the unparalleled diversity of sterols produced by sponges, although more work—particularly functional enzymatic studies—will be necessary to test this hypothesis. We have demonstrated that the correlation between C-24 alkylation and *SMT* count is sustained in animal outgroups such as choanoflagellates, filastereans, and ichthyosporeans, which suggests that *SMT* enzyme promiscuity is restricted to sponges. Our results similarly suggest that *Petrosia*, the one sponge in our study known to make C_{30} sterols, is also the only sponge to have a second copy of *SMT*, implicating a gene duplication in the evolution of sponge C_{30} sterol biosynthesis. However, even if future data complicates our interpretation of sponge *SMT* evolution, our primary hypothesis regarding the probable sponge origin of Neoproterozic biomarkers remains valid. In contrast to pelagophyte algae (35), molecular clocks indicate that crown-group demosponges radiated in the Neoproterozoic (this conclusion is robust even when 24-ipc is not used as a calibration point; see ref. 6 and figure S8 in ref. 8). Because 24-ipc is found in all major demosponge clades (4), their ability to produce 24-ipc most likely evolved by this time, regardless of the molecular mechanism required to perform it.

Taken collectively, our data provide strong evidence for the sponge biomarker hypothesis. We cannot rule out the possibility that 24-ipc convergence occurred in an independent, extinct branch of eukaryotes—an unavoidable consequence of studying the deep past—or that some living taxon might be discovered in the future that also produces 24-ipc. We can say that no other lineage of eukaryotes besides sponges is known to have the genetic toolkit necessary to generate such sterols and to have lived in the Neoproterozoic. Combined with recent rebuttals against putative sample contamination and alternate hypotheses of sterol diagenesis (42), there appears to be no current viable alternative to a sponge origin for Neoproterozoic 24-isopropylcholestanes. Perhaps the last major challenge against the sponge biomarker hypothesis is the "ghost lineage" linking this sterane to the oldest unambiguous fossils of siliceous sponges, which has previously been estimated at >100 million y (8). But the recent description of a 600-million-y-old putative sponge fossil dramatically shortens this gap (43). Such discoveries illustrate the significant incompleteness remaining in the early fossil record and reinforce the utility of biomarkers for understanding this critical period of biological evolution.

Methods

See *Supporting Information* for a detailed description of methods and Dataset S1 for all relevant input files, bioinformatics commands, and output trees. Briefly, lipid extractions were performed using a modified Bligh–Dyer method (44). The total lipid extracts (TLEs) were saponified by mild alkaline methanolysis, generating fatty-acid methyl-esters (45). Five-fraction column chromatography on SiO_2 was used to separate the sterols from other cellular lipids, which were derivatized to form trimethylsilyl ethers. GC–MS analyses were carried out on an Agilent GC-MSD 5975C equipped with a programmable temperature vaporizing injector (36). For comparative genomic analyses, proteomes were queried using candidate sequences (provided in Dataset S1) using BLASTp, with a *P* value cutoff of 10e-5. Proteins were vetted using tree-building algorithms and InterProScan (46). For the molecular clock analyses, an *SMT* gene tree was species tree error-corrected using NOTUNG (33). Molecular clock analyses were performed using the BEAST package (v.1.7.5) (47).

ACKNOWLEDGMENTS. We thank John Volkman and two anonymous referees for suggestions that significantly improved this report. We gratefully acknowledge funding from the Agouron Institute Geobiology Fellowship and the NASA Astrobiology Institute (NNA13AA90A) Foundations of Complex Life, Evolution, Preservation, and Detection on Earth and Beyond.

1. Pearson A, Budin M, Brocks JJ (2003) Phylogenetic and biochemical evidence for sterol synthesis in the bacterium Gemmata obscuriglobus. *Proc Natl Acad Sci USA* 100(26): 15352–15357.

2. Lamb DC, et al. (2007) Lanosterol biosynthesis in the prokaryote Methylococcus capsulatus: Insight into the evolution of sterol biosynthesis. *Mol Biol Evol* 24(8): 1714–1721.

3. Briggs DE, Summons RE (2014) Ancient biomolecules: Their origins, fossilization, and role in revealing the history of life. *BioEssays* 36(5):482–490.
4. Love GD, et al. (2009) Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* 457(7230):718–721.
5. Peterson KJ, Cotton JA, Gehling JG, Pisani D (2008) The Ediacaran emergence of bilaterians: Congruence between the genetic and the geological fossil records. *Philos Trans R Soc Lond B Biol Sci* 363(1496):1435–1443.
6. Sperling EA, Robinson JM, Pisani D, Peterson KJ (2010) Where's the glass? Biomarkers, molecular clocks, and microRNAs suggest a 200-Myr missing Precambrian fossil record of siliceous sponge spicules. *Geobiology* 8(1):24–36.
7. Parfrey LW, Lahr DJG, Knoll AH, Katz LA (2011) Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci USA* 108(33):13624–13629.
8. Erwin DH, et al. (2011) The Cambrian conundrum: Early divergence and later ecological success in the early history of animals. *Science* 334(6059):1091–1097.
9. Sperling EA, Peterson KJ, Laflamme M (2011) Rangeomorphs, Thectardis (Porifera?) and dissolved organic carbon in the Ediacaran oceans. *Geobiology* 9(1):24–33.
10. Maloof AC, et al. (2010) Possible animal-body fossils in pre-Marinoan limestones from South Australia. *Nat Geosci* 3(9):653–659.
11. Ghisalberti M, et al. (2014) Canopy flow analysis reveals the advantage of size in the oldest communities of multicellular eukaryotes. *Curr Biol* 24(3):305–309.
12. Antcliffe JB (2013) Questioning the evidence of organic compounds called sponge biomarkers. *Palaeontology* 56(5):917–925.
13. Antcliffe JB, Callow RHT, Brasier MD (2014) Giving the early fossil record of sponges a squeeze. *Biol Rev Camb Philos Soc* 89(4):972–1004.
14. Desmond E, Gribaldo S (2009) Phylogenomics of sterol synthesis: Insights into the origin, evolution, and diversity of a key eukaryotic feature. *Genome Biol Evol* 1:364–381.
15. Kodner RB, Summons RE, Pearson A, King N, Knoll AH (2008) Sterols in a unicellular relative of the metazoans. *Proc Natl Acad Sci USA* 105(29):9897–9902.
16. Giner J-L, Zhao H, Boyer GL, Satchwell MF, Andersen RA (2009) Sterol chemotaxonomy of marine pelagophyte algae. *Chem Biodivers* 6(7):1111–1130.
17. Watson JR, Brennan TC, Degnan BM, Degnan SM, Krömer JO (2014) Analysis of the biomass composition of the demosponge Amphimedon queenslandica on Heron Island Reef, Australia. *Mar Drugs* 12(6):3733–3753.
18. Schmitz FJ (1978) Uncommon marine steroids. *Marine Natural Products V1: Chemical and Biological Perspectives*, ed Schever P (Academic Press, New York), pp 241–297.
19. Bergquist PR, Karuso P, Cambie RC (1991) Sterol composition and classification of the Porifera. *Biochem Syst Ecol* 19(1):17–24.
20. Riesgo A, Farrar N, Windsor PJ, Giribet G, Leys SP (2014) The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. *Mol Biol Evol* 31(5):1102–1120.
21. Giner JL, Gunasekera SP, Pomponi SA (1999) Sterols of the marine sponge Petrosia weinbergi: Implications for the absolute configurations of the antiviral orthoesterols and weinbersterols. *Steroids* 64(12):820–824.
22. Carland F, Fujioka S, Nelson T (2010) The sterol methyltransferases SMT1, SMT2, and SMT3 influence Arabidopsis development through nonbrassinosteroid products. *Plant Physiol* 153(2):741–756.
23. Djerassi C, Silva CJ (1991) Biosynthetic studies of marine lipids. 41. Sponge sterols: Origin and biosynthesis. *Acc Chem Res* 24(12):371–378.
24. Hagemann A, Voigt O, Wörheide G, Thiel V (2008) The sterols of calcareous sponges (Calcarea, Porifera). *Chem Phys Lipids* 156(1-2):26–32.
25. Aiello A, Fattorusso E, Magno S, Menna M (1991) Isolation of five new 5 alpha-hydroxy-6-keto-delta 7 sterols from the marine sponge Oscarella lobularis. *Steroids* 56(6):337–340.
26. Silva CJ, Wünsche L, Djerassi C (1991) Biosynthetic studies of marine lipids. 35. The demonstration of de novo sterol biosynthesis in sponges using radiolabeled isoprenoid precursors. *Comp Biochem Physiol B* 99(4):763–773.
27. Siegl A, et al. (2011) Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. *ISME J* 5(1):61–70.
28. Cohen HM, Tawfik DS, Griffiths AD (2002) Promiscuous methylation of non-canonical DNA sites by HaeIII methyltransferase. *Nucleic Acids Res* 30(17):3880–3885.
29. Kopycki JG, et al. (2008) Biochemical and structural analysis of substrate promiscuity in plant Mg2+-dependent O-methyltransferases. *J Mol Biol* 378(1):154–164.
30. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: A unified platform for automated protein structure and function prediction. *Nat Protoc* 5(4):725–738.
31. Husselstein T, Gachotte D, Desprez T, Bard M, Benveniste P (1996) Transformation of Saccharomyces cerevisiae with a cDNA encoding a sterol C-methyltransferase from Arabidopsis thaliana results in the synthesis of 24-ethyl sterols. *FEBS Lett* 381(1-2):87–92.
32. Nes WD, et al. (2004) Sterol methyltransferase: Functional analysis of highly conserved residues by site-directed mutagenesis. *Biochemistry* 43(2):569–576.
33. Chen K, Durand D, Farach-Colton M (2000) NOTUNG: A program for dating gene duplications and optimizing gene family trees. *J Comput Biol* 7(3-4):429–447.
34. Volkman JK (2003) Sterols in microorganisms. *Appl Microbiol Biotechnol* 60(5):495–506.
35. Brown JW, Sorhannus U (2010) A molecular genetic timescale for the diversification of autotrophic stramenopiles (Ochrophyta): Substantive underestimation of putative fossil ages. *PLoS One* 5(9):12759.
36. Grabenstatter J, et al. (2013) Identification of 24-n-propylidenecholesterol in a member of the Foraminifera. *Org Geochem* 63:145–151.
37. Pawlowski J, et al. (2003) The evolution of early Foraminifera. *Proc Natl Acad Sci USA* 100(20):11494–11498.
38. Bosak T, et al. (2012) Possible early foraminiferans in post-Sturtian (716– 635 Ma) cap carbonates. *Geol* 40(1):67–70.
39. Moldowan JM, Jacobson SR (2000) Chemical signals for early evolution of major taxa: Biosignatures and taxon-specific biomarkers. *Int Geol Rev* 42(9):805–812.
40. Moldowan JM, et al. (1990) Sedimentary 12-n-propylcholestanes, molecular fossils diagnostic of marine algae. *Science* 247(4940):309–312.
41. Rohrssen M, Gill BC, Love GD (2015) Scarcity of the C_{30} sterane biomarker, 24-n-propylcholestane, in Lower Paleozoic marine paleoenvironments. *Org Geochem* 80:1–7.
42. Love GD, Summons RE (2015) The molecular record of Cryogenian sponges—A response to Antcliffe (2013). *J Paleo* 58(6):1131–1136.
43. Yin Z, et al. (2015) Sponge grade body fossil with cellular resolution dating 60 Myr before the Cambrian. *Proc Natl Acad Sci USA* 112(12):E1453–E1460.
44. Iverson SJ, Lang SL, Cooper MH (2001) Comparison of the Bligh and Dyer and Folch methods for total lipid determination in a broad range of marine tissue. *Lipids* 36(11):1283–1287.
45. Dittmer JC, Wells MA (1969) Quantitative and qualitative analysis of lipids and lipid components. *Methods in Enzymology*, eds Colowick SP, Kaplan NO (Academic Press, New York), pp 482–530.
46. Zdobnov EM, Apweiler R (2001) InterProScan–An integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17(9):847–848.
47. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29(8):1969–1973.
48. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 59(3):307–321.
49. Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57(5):758–771.
50. Brooks CJW, Horning EC, Young JS (1968) Characterization of sterols by gas chromatography-mass spectrometry of the trimethylsilyl ethers. *Lipids* 3(5):391–402.
51. Kates M, Tremblay P, Anderson R, Volcani BE (1978) Identification of the free and conjugated sterol in a non-photosynthetic diatom, *Nitzschia alba*, as 24-methylene cholesterol. *Lipids* 13(1):34–41.
52. Gerst N, Ruan B, Pang J, Wilson WK, Schroepfer GJ, Jr (1997) An updated look at the analysis of unsaturated C27 sterols by gas chromatography and mass spectrometry. *J Lipid Res* 38(8):1685–1701.
53. Knights BA (1967) Identification of plant sterols using combined GLC/mass spectrometry. *J Chromatogr Sci* 5(6):273–282.
54. Kamal-Eldin A, Määttä K, Toivo J, Lampi AM, Piironen V (1998) Acid-catalyzed isomerization of fucosterol and delta5-avenasterol. *Lipids* 33(11):1073–1077.
55. Weete JD, Gandhi SR (1997) Sterols of the phylum zygomycota: Phylogenetic implications. *Lipids* 32(12):1309–1316.
56. Patterson GW, Nes WD (1992) *Physiology and Biochemistry of Sterols* (American Oil Chemists' Society, Champaign, IL).
57. Weete JD, Abril M, Blackwell M (2010) Phylogenetic distribution of fungal sterols. *PLoS One* 5(5):e10899.
58. Bergmann W, McTigue FH (1948) Contributions to the study of marine products; chondrillasterol. *J Org Chem* 13(5):738–741.
59. Hase A (1981) Fatty acid composition and sterol content of Dictyostelium discoideum cells at various stages of development. *J Fac Sci Hokkaido Univ* 12(36):183–194.
60. Volkman JK, et al. (2009) Sterol and fatty acid composition of four marine haptophycean algae. *J Mar Biol Ass* 61(2):509–527.
61. Latour NG, Reeves RE, Guidry MA (1965) Steroid requirement of Entamoeba histolytica. *Exp Parasitol* 16(1):18–22.
62. Venkateswarlu Y, Reddy MVR, Rao MR (1996) A new epoxy sterol from the sponge Ircinia fasciculata. *J Nat Prod* 59(9):876–877.
63. Goad LJ, Holz GG, Jr, Beach DH (1984) Sterols of Leishmania species. Implications for biosynthesis. *Mol Biochem Parasitol* 10(2):161–170.
64. Nelson MM, Phleger CF, Mooney BD, Nichols PD (2000) Lipids of gelatinous Antarctic zooplankton: Cnidaria and Ctenophora. *Lipids* 35(5):551–559.
65. Méjanelle L, Sanchez-Gargallo A, Bentaleb I, Grimalt JO (2003) Long chain n-alkyl diols, hydroxy ketones and sterols in a marine eustigmatophyte, Nannochloropsis gaditana, and in Brachionus plicatilis feeding on the algae. *Org Geochem* 34:527–538.
66. Whitaker BD, Nelson DL (1987) Growth support and metabolism of phytosterols in Paramecium tetraurelia. *Lipids* 22(6):386–396.
67. Lund ED, Chu F-LE, Soudant P, Harvey E (2007) Perkinsus marinus, a protozoan parasite of the eastern oyster, has a requirement for dietary sterols. *Comp Biochem Physiol A Mol Integr Physiol* 146(1):141–147.
68. Fabris M, et al. (2014) Tracking the sterol biosynthesis pathway of the diatom Phaeodactylum tricornutum. *New Phytol* 204(3):521–535.
69. Manconi R, Piccialli V, Pronzato R, Sica D (1988) Steroids in porifera, sterols from freshwater sponges Ephydatia fluviatilis (L.) and Spongilla lacustris (L.). *Comp Biochem Physiol B* 91(2):237–245.
70. Najle SR, Nusblat AD, Nudel CB, Uttaro AD (2013) The Sterol-C7 desaturase from the ciliate Tetrahymena thermophila is a Rieske Oxygenase, which is highly conserved in animals. *Mol Biol Evol* 30(7):1630–1643.
71. Volkman JK, Hallegraeff GM (1988) Lipids in marine diatoms of the genus Thalassiosira: Predominance of 24-methylenecholesterol. *Phytochemistry* 27(5):1389–1394.
72. Nes CR, et al. (2012) Novel sterol metabolic network of *Trypanosoma brucei* procyclic and bloodstream forms. *Biochem J* 443(1):267–277.
73. Ehrlich H, et al. (2013) Discovery of 505-million-year old chitin in the basal demosponge Vauxia gracilenta. *Sci Rep* 3:3497.
74. Botting JP (2005) Exceptionally well-preserved middle Ordovician sponges from the Llandegley rocks lagerstätte, Wales. *Palaeontology* 48(3):577–617.
75. Taylor TN, Remy W, Hass H (1992) Fungi from the Lower Devonian Rhynie chert: Chytridiomycetes. *Am J Bot* 79(11):1233–1241.