

RESEARCH ARTICLE

Open Access



Sulfur, sterol and trehalose metabolism in the deep-sea hydrocarbon seep tubeworm *Lamellibrachia luymesii*

Hong Shi^{1*}, Lingwei Ruan^{1,2,3*}, Zimeng Chen¹, Yifei Liao^{1,4}, Wenhao Wu^{1,5}, Linmin Liu¹ and Xun Xu¹

Abstract

Background *Lamellibrachia luymesii* dominates cold sulfide-hydrocarbon seeps and is known for its ability to consume bacteria for energy. The symbiotic relationship between tubeworms and bacteria with particular adaptations to chemosynthetic environments has received attention. However, metabolic studies have primarily focused on the mechanisms and pathways of the bacterial symbionts, while studies on the animal hosts are limited.

Results Here, we sequenced the transcriptome of *L. luymesii* and generated a transcriptomic database containing 79,464 transcript sequences. Based on GO and KEGG annotations, we identified transcripts related to sulfur metabolism, sterol biosynthesis, trehalose synthesis, and hydrolysis. Our in-depth analysis identified sulfation pathways in *L. luymesii*, and sulfate activation might be an important detoxification pathway for promoting sulfur cycling, reducing byproducts of sulfide metabolism, and converting sulfur compounds to sulfur-containing organics, which are essential for symbiotic survival. Moreover, sulfide can serve directly as a sulfur source for cysteine synthesis in *L. luymesii*. The existence of two pathways for cysteine synthesis might ensure its participation in the formation of proteins, heavy metal detoxification, and the sulfide-binding function of haemoglobin. Furthermore, our data suggested that cold-seep tubeworm is capable of de novo sterol biosynthesis, as well as incorporation and transformation of cycloartenol and lanosterol into unconventional sterols, and the critical enzyme involved in this process might have properties similar to those in the enzymes from plants or fungi. Finally, trehalose synthesis in *L. luymesii* occurs via the trehalose-6-phosphate synthase (TPS) and trehalose-6-phosphate phosphatase (TPP) pathways. The TPP gene has not been identified, whereas the TPS gene encodes a protein harbouring conserved TPS/OtsA and TPP/OtsB domains. The presence of multiple trehalases that catalyse trehalose hydrolysis could indicate the different roles of trehalase in cold-seep tubeworms.

Conclusions We elucidated several molecular pathways of sulfate activation, cysteine and cholesterol synthesis, and trehalose metabolism. Contrary to the previous analysis, two pathways for cysteine synthesis and the cycloartenol-C-24-methyltransferase gene were identified in animals for the first time. The present study provides new insights into particular adaptations to chemosynthetic environments in *L. luymesii* and can serve as the basis for future molecular studies on host-symbiont interactions and biological evolution.

*Correspondence:

Hong Shi

shihong@tio.org.cn

Lingwei Ruan

ruanlingwei@tio.org.cn

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Adaptation to deep-sea chemosynthetic environment, Cold seep, Transcriptome, Sulfate, Cysteine, Cholesterol, Cycloartenol-C-24-methyltransferase, Trehalase, Fungi

Background

Cold seeps are areas of the ocean floor where leakage of hydrogen sulfide, methane, and other hydrocarbon-rich fluids occurs. These chemically and geologically diverse vent fields provide habitats for various vent fauna, including a single species of tubeworm, *Lamellibrachia luymesii* (Annelida: Siboglinidae). These tubeworms live only at shallow-water hydrocarbon seep vents in the Atlantic Ocean at depths of less than 1000 m [1].

These tubeworms lack a digestive tract and rely on symbiosis with chemoautotrophic bacteria [2]. They exchange gas with deep red plumes and take up sulfide from the environment by using extensions of their tubes that penetrate the seafloor sediment. Then, the sulfide is transported to the chemoautotrophic bacterial endosymbionts, belonging to Gammaproteobacteria, which live inside the bacteriocytes (specialized cells) of the trophosome (an organ produced by the host to house and protect its microbial partner). In return, these endosymbionts utilize the energy from sulfide oxidation to fuel the Calvin-Benson cycle. This respiratory process results in the production of organic carbon, which is provided to *L. luymesii*. This symbiosis has received much attention, and studies have examined the energy metabolism, host-microbe interactions, and adaptation to deep-sea chemosynthetic environments to reveal unique mechanisms associated with chemical ecology and metabolic evolution.

To date, most studies of tubeworms and their endosymbionts have focused on the hydrothermal vent tubeworm *Riftia pachyptila* (Vestimentifera), including studies on the metabolism of sulfide [3–7], hydrogen [8], nitrogen [9–12], and carbon [13]. The *L. luymesii* genomic sequences were recently published [14], and the mechanisms related to nutrition mode, haemoglobin evolution, immunity function, longevity, and the cell cycle were discussed. Moreover, we have completed the *L. luymesii* transcriptome sequencing and assembly analysis. The availability of this large set of transcriptomic and genomic data from *L. luymesii* facilitates further exploration of the adaptation of deep-sea fauna to chemosynthetic environments.

Of particular interest has been the ability of hydrocarbon seep tubeworms to cope with toxic levels of hydrogen sulfide and eliminate sulfate, the primary waste product of chemoautotrophic sulfide oxidation. Tubeworms rely on sulfide for metabolism; hydrogen sulfide and oxygen are delivered to chemoautotrophic

bacteria by haemoglobin [3, 14–17]. Moreover, more haemoglobin B1 gene copies were recently found in the genome of *L. luymesii* than in other siboglinids, which might indicate increased sulfide-binding capacity to fulfil symbiosis [18]. On the other hand, symbionts oxidize sulfide to produce energy for carbon fixation and release sulfate and hydrogen ions as byproducts [6, 7]. *L. luymesii* was predicted to eliminate these byproducts across its roots to conserve energy and ensure sulfide supply. The elimination of metabolic waste products is vital for *L. luymesii*. However, the oxidized oxyanion sulfate is also the most available form of sulfur in nature, and detailed knowledge of how *L. luymesii* activates sulfate and incorporates it into diverse metabolites is unknown.

Deep-sea animals live under high pressure. One of the ways in which they adapt to their harsh environment might result from the unique properties of the cellular membrane. Sterols are essential components of the membranes of all eukaryotic organisms, controlling membrane fluidity and permeability [19]. Moreover, some cyclic triterpenoid lipids, such as steranes (degraded and saturated derivatives of sterols), are essential and have remained stable through deep geological time. They have been proposed as molecular fossils, recording the evolution of organisms even in the absence of physically preserved fossils [20]. Indeed, sterol biosynthesis is a multistep process catalysed by a series of enzymes. Animals (e.g., mammals) can de novo synthesize cholesterol from low molecular weight precursors such as acetate [21]. Nevertheless, several marine invertebrates (e.g., *Bathymodiolus* mussels) are exceptional in their ability to synthesize cholesterol de novo by themselves, converting exogenous sterols to cholesterol [22]. Interestingly, a recent study revealed that *L. luymesii* lacks many genes essential for amino acid biosynthesis and might depend mainly on endosymbionts for nutrition [14]. However, information on the sterol biosynthetic pathway in *L. luymesii* is lacking.

The sugar trehalose (α -D-glucopyranosyl (1, 1)- α -D-glucopyranoside), a nonreducing disaccharide of glucose, is found widely but not abundantly in nature [23, 24]. Its functions proposed by Colaço and Roser in 1995 are myriad and include water replacement, glass transformation, and chemical stability [25]. Thus, this sugar can act as a stress protectant in biological systems as it interacts with and directly protects membranes and proteins from the damage caused by physiological and

environmental changes [26]. For this reason, many organisms accumulate trehalose to adapt to environmental extremes [27]. Additionally, trehalose also acts as an energy reserve in nematodes and insects. It appears to function as a significant circulating blood sugar [28, 29]. Based on the importance of trehalose metabolism for extreme adaptation in animals, we aimed to characterize enzymes involved in the synthesis and hydrolysis of trehalose.

Here, we report the assembly of the cold-seep *Lamellibrachia luymesii* transcriptome. By combining transcriptomic and genomic data, we further investigated the adaptation to deep-sea chemosynthetic environments in *L. luymesii* with special attention to sulfur, sterol, and trehalose metabolism pathways. The novel candidate genes identified in these pathways could be associated with the extraordinary adaptation and symbiotic associations of tubeworms. These results provide new insights into adaptations to chemosynthetic environments in *L. luymesii*.

Results and discussion

De novo assembly and functional annotation of the *L. luymesii* transcriptome

The mitochondrial cytochrome c oxidase subunit I (COI) gene sequence of the specimen shared 100% identity with the corresponding *L. luymesii* sequences from the Gulf of Mexico (GenBank accession no. GU059168.1). We produced 6.86 Gbp of clean data from the tissue sample. Over 96% of the clean Illumina reads exceeded Q20, indicating the high quality of the sequencing data. The raw sequencing data have been submitted to the Short Read Archive (SRA) of NCBI under accession number SRR17999638. The final transcriptome generated from Trinity de novo assembly contained 79,464 transcript sequences. The lengths of the transcripts ranged from 200 to 13,725 bp, with an average length of 1047 bp and an N50 value of 1906 bp. The statistics for the data output and de novo assemblies are summarized in Table 1. The length distribution of all the transcripts and unigenes is shown in Figure S1. In total, 21,300 (43.24%) unigenes had at least one significant hit in the NCBI nonredundant (nr), Swiss-Prot, KO, PFAM, GO, and KOG databases and the NCBI nucleotide database. Over half of the dominant transcripts had no match among proteins in any public database. The functional annotation results are shown in Table 1. The KOG functional categories and KEGG pathways for the annotated unigenes are shown in Fig. 1A and B, respectively. There were 16,887 (34.28%) nr-annotated unigenes assigned to major Gene Ontology (GO) categories, i.e., “Biological Process”, “Cellular Component”, and “Molecular Function” (Fig. 1C). EST-scan and BLAST search of all the databases mentioned above resulted in the prediction of 30,781 protein-coding

Table 1 Summary of assembling and functional annotation of *L. luymesii* transcriptome

De novo assembly by Trinity	
Total base (bp)	68,590,831
Total number of transcripts	79,464
Number of unigenes	49,254
Mean length of transcripts (bp)	1,047
N50 (bp)	1,906
Transcripts size range (bp)	200–13,725
Functional annotation	
Total number of Unigenes annotated by public databases	49,254
NCBI non-redundant (nr) database (e-value < 1e-5)	16,887
SwissProt (e-value < 1e-5)	13,654
PFAM	15,753
Gene Ontology	16,254
KOG	9,211
Coding sequence prediction	
CDS predicted from BLAST results	17,055
CDS predicted by ESTScan	13,726
Total CDS predicted	30,781
Mean length of CDS (aa)	196

transcripts (Table 1). The best hit for most of the annotated unigenes (4412 out of 16,887, 26.1%) was *Crassostrea gigas* in the nr database (Fig. 1D). This result might be attributed to annelids and molluscs sharing a sister phylogenetic relationship [30] and to the availability of molecular resources for marine annelids being scarcer than that for other phyla.

Sulfur metabolism and cysteine biosynthesis

L. luymesii acquires sulfide from the sediment and then supplies sulfide to endosymbiotic bacteria to produce energy for carbon fixation. While *L. luymesii* obtains its nutrition from symbionts, they must eliminate or transform sulfate byproducts.

The present study found that sulfation pathways are present in *L. luymesii*. We have depicted the *L. luymesii* sulfation pathways in Fig. 2A and Table S1. The FPKM (fragments per kilobase million) values of genes in this pathway are shown in Fig. 2B. Several sulfate transporters and sulfite oxidases are responsible for cellular sulfate uptake, followed by the two-step enzymatic sulfate activation performed by the bifunctional 3'-phosphoadenosine-5'-phosphosulfate synthase (PAPSS). PAPSS in *L. luymesii* possesses both ATP sulfurylase (SUL) and APS kinase (KIN) activities (Fig. 3). This fusion probably increases the catalytic efficiency [31, 32]. The product PAPS, as the activated sulfate group donor, is either used directly by cytoplasmic and nuclear sulfotransferases

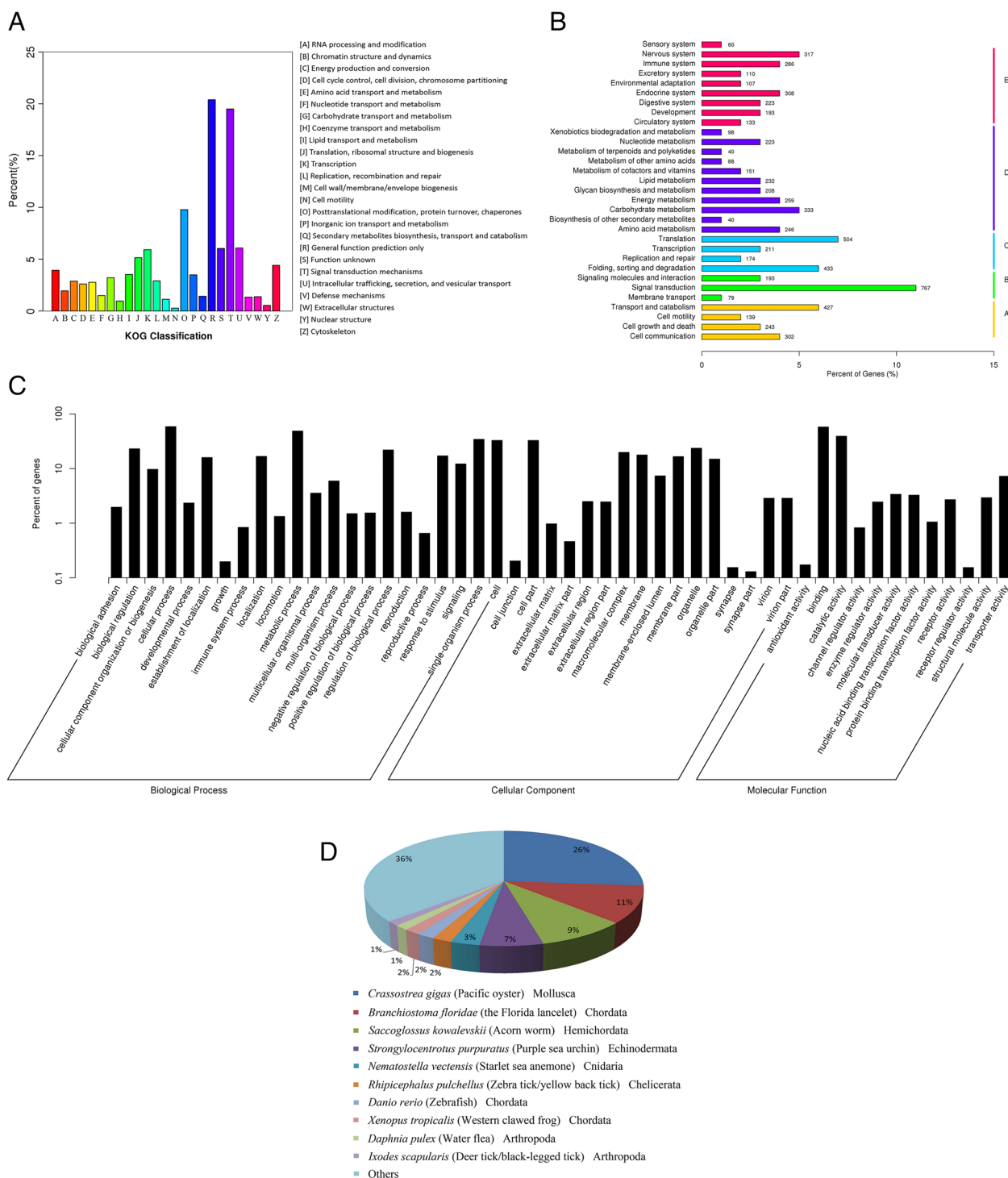


Fig. 1 A summary of the functional annotation of the *L. lyymesii* transcriptome. Functional classification of transcripts annotated by **A** KOG; **B** KEGG; **C** GO. **D** Species distribution of annotated transcripts

or shuttled to the Golgi apparatus to serve a multitude of Golgi-residing carbohydrate and protein sulfotransferases. Due to the substrate diversity, sulfotransferases

and sulfatases form large multigene families. These enzymes, which participate in sulfation and desulfation of bioorganic compounds, are essential for cellular activity

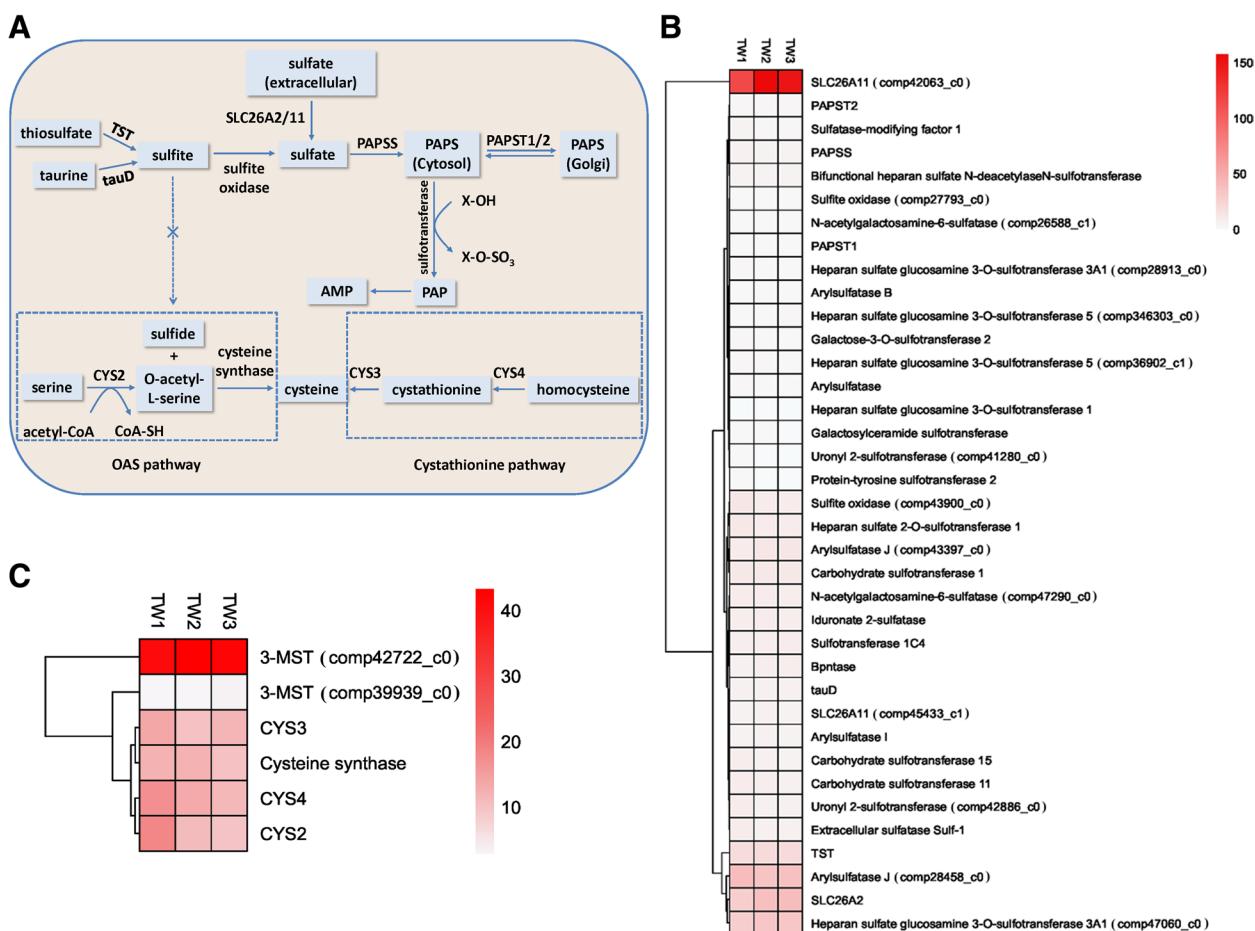


Fig. 2 **A** Scheme of *L. luymesi* sulfation pathways and biosynthetic pathways for cysteine. TST, thiosulfate sulfurtransferase; tauD, alpha-ketoglutarate-dependent taurine dioxygenase; SLC26A2, sulfate transporter; SLC26A11, sodium-independent sulfate anion transporter; PAPSS, bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase; PAPST1/2, adenosine 3'-phospho 5'-phosphosulfate transporter 1/2; CYS2, probable serine-O-acetyltransferase cys2; CYS3, putative cystathionine gamma-lyase 2; CYS4, cystathionine beta-synthase. **B** Heatmap of FPKM expression values for the annotated genes in *L. luymesi* sulfation pathway from three individual samples (TW1, TW2 and TW3). **C** Heatmap of FPKM expression values for the annotated genes in *L. luymesi* biosynthetic pathways for cysteine. FPKM values were listed in Table S4

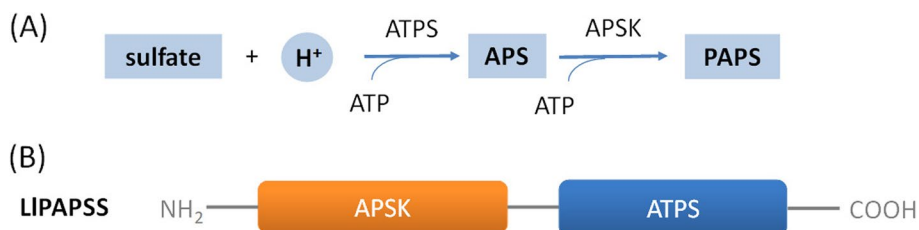


Fig. 3 PAPS synthesis in *L. luymesi*. **A** PAPS synthesis pathway. **B** The domain formation of *L. luymesi* PAPSS, a fused gene encoding a protein with two conserved domains. APS, adenosine 5'-phosphosulfate; PAPS, bifunctional 3'-phosphoadenosine-5'-phosphosulfate; PAPSS, PAPS synthase; APSK, APS kinase; ATPS, ATP sulfurylase

in animal cells, such as extracellular communication, inflammation, and lymphocyte homing, and play a vital role in a functional endocrine system [33–36]. Furthermore, cytoplasmic adenosine 3',5'-bisphosphate (PAP),

the otherwise toxic byproduct of sulfation, is removed and degraded to AMP by phosphate 3'(2'),5'-bisphosphate nucleotidase. Therefore, sulfate activation could be an important detoxification pathway in *L. luymesi* to

promote sulfur cycling, reduce sulfate, and convert sulfur compounds to sulfur-containing organics essential for the tubeworm.

Additionally, the heat map displays the relative expression levels of two sulfate transporter genes (sulfate transporter and sodium-independent sulfate anion transporter, also known as SLC26A2 and SLC26A11 in human, respectively) are particularly high, compared to that of other genes in this pathway (Fig. 2B). SLC26A11 was sensitive to the anion exchange inhibitor DIDS and can specifically mediate $\text{Cl}^-/\text{HCO}_3^-$ exchange, remarkable co-localized with H^+ -ATPase in human cells [37, 38]. On the other hand, Dattagupta et al. previously proposed a model to indicate *L. luymesii* primarily eliminated the sulfate mainly with transporters coupled with bicarbonate uptake [6]. This model depends on an anion exchanger that is sensitive to DIDS and exchanges sulfate ions for either Cl^- or HCO_3^- . Thus we speculated that SLC26A2 and SLC26A11 might be the potential sulfate-bicarbonate exchangers which are highly expressed in *L. luymesii*. This result provided evidence and further support that sulfation pathway might be involved in the detoxification for promoting sulfur cycling and reducing byproducts.

Sulfite reductase is a flavoprotein complex that catalyses the reduction of sulfite to sulfide in the cysteine and methionine biosynthesis pathway. Four transcripts of sulfite reductase (three alpha and one beta component) were found in the cold seep mussel *Bathymodiolus platifrons* [39]. It is worth noting that we failed to find any gene that could encode a sulfite reductase homologue. The transcriptome and genome sequence analyses did not have 100% coverage, so we cannot completely discount the possibility that the gene is present. However, because *L. luymesii* lives in sulfide-rich subsurface sediment zones, the existence of sulfite reductase might not be necessary for symbiotic survival.

In addition, sulfur is also a constituent of the amino acid *cysteine*. Interestingly, in the present study, the transcriptome analysis revealed four genes encoding proteins possibly involved in two pathways for cysteine synthesis in *L. luymesii*, namely, cystathionine gamma-lyase (CYS3), cystathionine beta-synthase (CYS4), serine-O-acetyltransferase (CYS2), and cysteine synthase (Table S1). Thus, sulfide can be a direct sulfur source for cysteine synthesis in *L. luymesii*.

In general, two different routes for cysteine biosynthesis have been described in animals and plants/bacteria. Cysteine biosynthesis begins with the amino acid serine in animals. Sulfur is derived from methionine, which is converted to homocysteine. CYS4 combines homocysteine and serine to form the asymmetrical

thioether cystathionine. CYS3 further converts cystathionine to cysteine and alpha-ketobutyrate. This pathway is called the cystathionine pathway. Moreover, in plants and bacteria, cysteine biosynthesis also starts from serine, which is converted to O-acetylserine (OAS) by CYS2. Then, cysteine synthase, using sulfide sources, converts this ester to cysteine, releasing acetate [40, 41] (Fig. 2).

To the best of our knowledge, this is the first report showing that two pathways for cysteine synthesis are present in animals. In the OAS pathway, *cysteine synthase* (also known as O-acetylserine sulfhydrylase, *cysK*) is a pyridoxal 5'-phosphate-dependent enzyme that catalyses the final step. Cysteine synthase of *L. luymesii* (LICs) is predicted to encode a protein of 374 amino acids, with a subunit molecular mass of 40.37 kDa. Phylogenetic analysis showed that the LICs separated from other marine invertebrate to form an independent evolutionary branch and aligned more closely with CSs of fungi than CSs from bacteria, plants, and archaea (Fig. 4). To date, the existence of two pathways for cysteine synthesis has been demonstrated in fungi, such as the filamentous fungus *Aspergillus nidulans*, which has been reported to have the richest repertoire of sulfur metabolic options [42]. Analysis of the sulfur amino acid metabolism in *L. luymesii* demonstrated that it shares features of plants, bacteria and animal systems in cysteine interconversion. This characteristic highlights that the cysteine synthesis pathway plays a vital role in *L. luymesii*. Cysteine is involved in the structure, stability, and catalytic functions of many proteins and contributes to the antioxidant activity of glutathione [43]. Cysteine biosynthesis is critical for heavy metal and metalloid resistance, such as arsenic and cadmium resistance [44]. More importantly, cysteine residues in haemoglobin determine the sulfur binding capabilities that permit tubeworms to live in chemosynthetic conditions [18, 45].

Whether *serine-O-acetyltransferase* is present in *L. luymesii* might be controversial. Published genome analysis showed that serine-O-acetyltransferase (also known as *cysE*) was present in symbionts but absent in the Lamellibrachia host [18]. This study found that a probable serine-O-acetyltransferase (LICYS2) was annotated in the *L. luymesii* transcriptome and genome. Moreover, sequence analysis revealed that LICYS2 shares significant similarity with *Cys2* in *Schizosaccharomyces pombe*. As noted previously, *S. pombe* *Cys2* is a serine O-acetyltransferase specifically essential for cysteine biosynthesis [44, 46]. Therefore, LICYS2 is more likely a serine O-acetyltransferase involved in the OAS pathway for cysteine biosynthesis in *L. luymesii*.

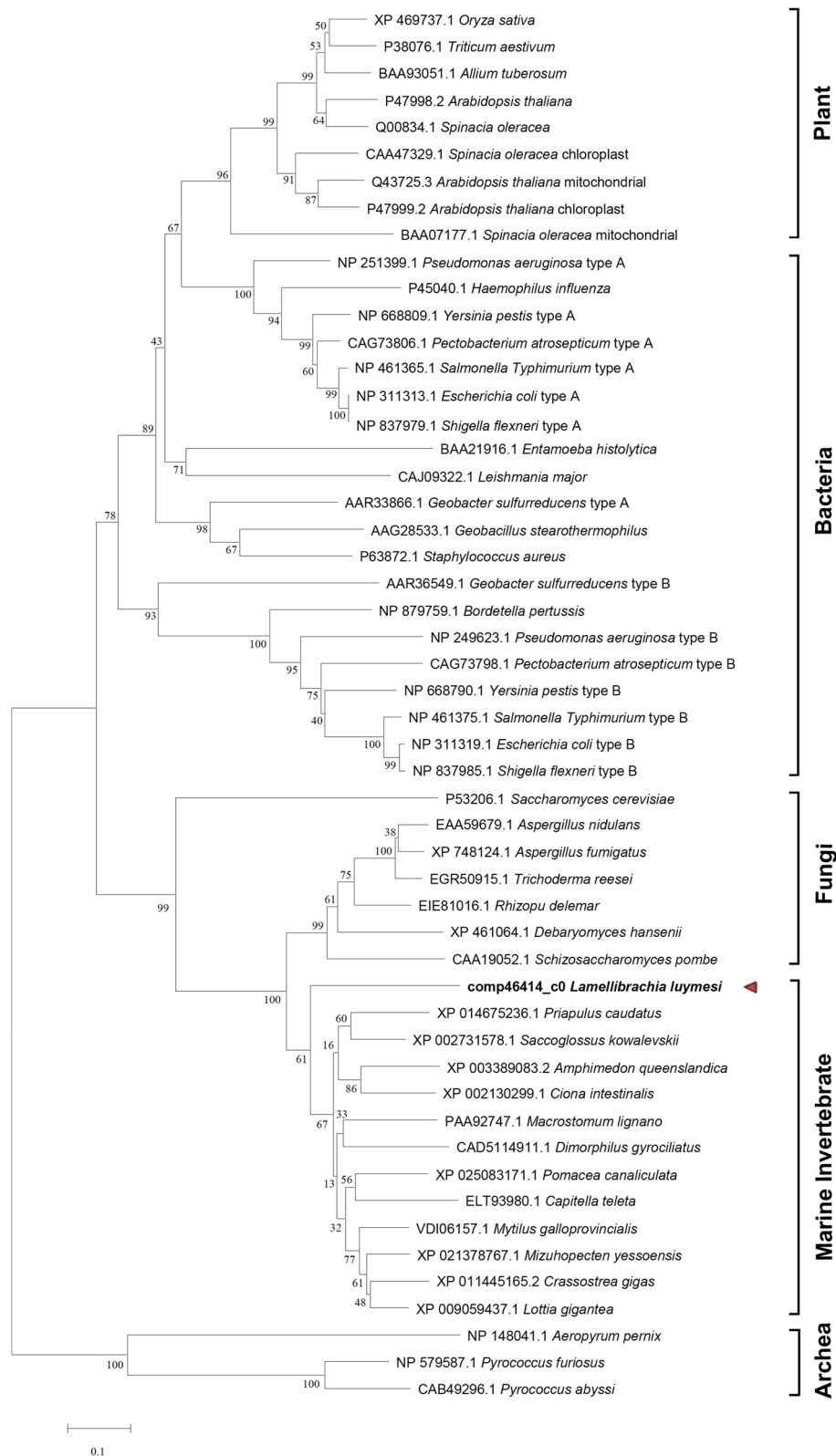


Fig. 4 Phylogenetic tree of CS sequences. *L. luymesii* CS is highlighted in bold

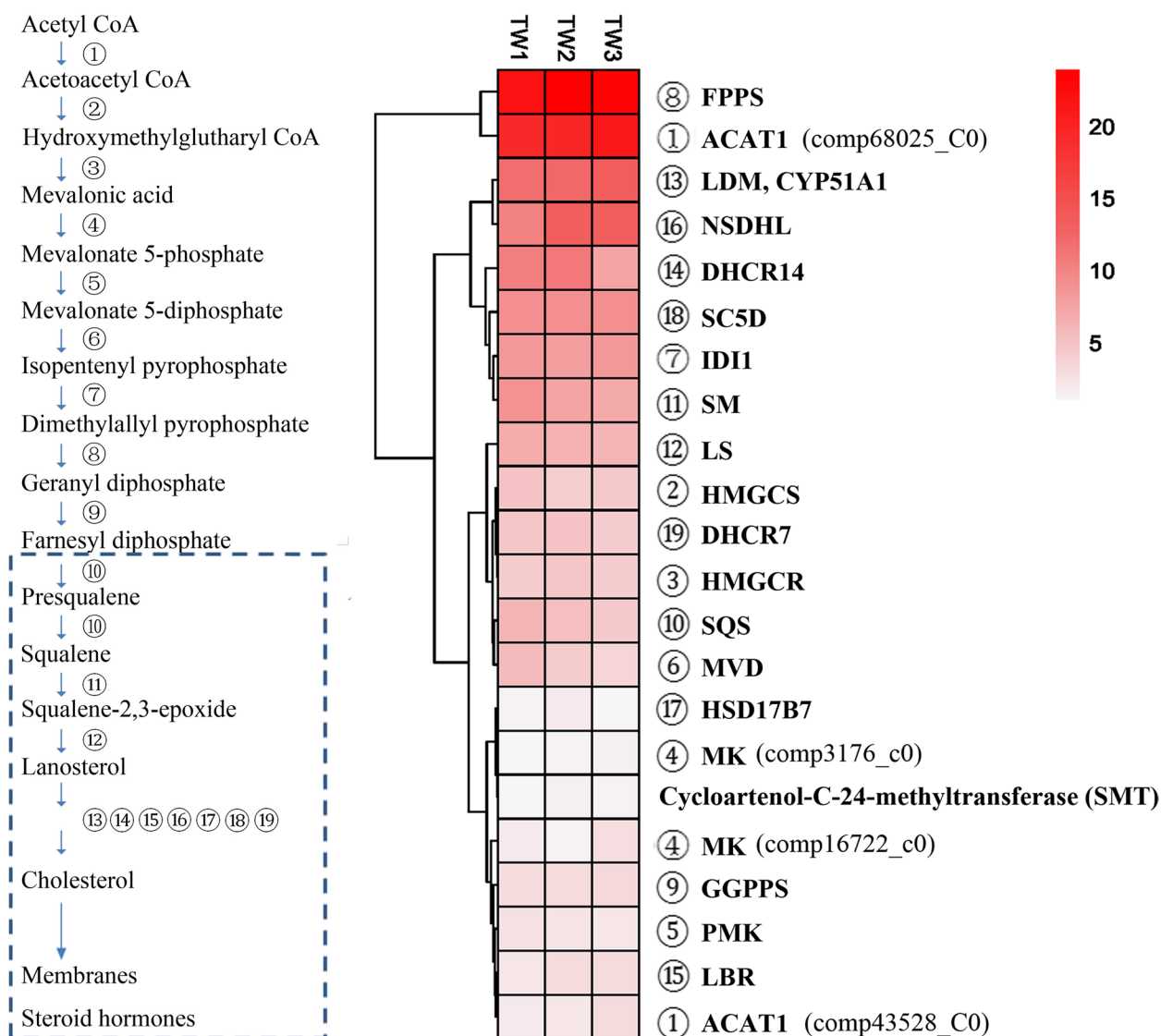


Fig. 5 **A** Overview of cholesterol synthesis pathway in *L. lyngbyi*. Circles represent the following enzymes: 1, acetyl-CoA acetyltransferase; 2, hydroxymethylglutaryl-CoA synthase; 3, 3-hydroxy-3-methylglutaryl-coenzyme A reductase; 4, mevalonate kinase; 5, promyelvalonate kinase; 6, diphosphomevalonate decarboxylase; 7, isopentenyl-diphosphate delta-isomerase; 8, farnesyl pyrophosphate synthase; 9, geranylgeranyl pyrophosphate synthase; 10, squalene synthase; 11, squalene monooxygenase; 12, lanosterol synthase; 13, lanosterol 14-alpha demethylase; 14, delta(14)-sterol reductase; 15, lamin-B receptor; 16, sterol-4-alpha-carboxylate 3-dehydrogenase; 17, 3-keto-sterol reductase; 18, lathosterol oxidase; 19, 7-dehydrocholesterol reductase. *C. elegans* lacks the branch inside the dashed box. **B** Heatmap of FPKM expression values for the annotated genes in *L. lyngbyi* cholesterol synthesis pathway from three individual samples (TW1, TW2 and TW3). FPKM values were listed in Table S4

Biosynthetic pathways for sterols

Sterols are crucial for the functioning of most eukaryotic cells and the structure of their membranes and serve as precursors to signalling and hormone molecules to control developmental processes [47, 48].

Mammals synthesize squalene and various sterols from low-molecular-weight precursors such as acetate via mevalonate. However, other animals exclusively depend on dietary sterols due to the complete or partial absence

of a cholesterol biosynthetic pathway, such as insects, nematodes [49, 50], the annelid *Lumbricus terrestris* [51, 52], and some marine invertebrate crustaceans [53–56]. For example, *C. elegans* expresses predicted homologues of the enzymes that produce the initial intermediates of the mammalian sterol biosynthetic pathway up to farnesyl diphosphate but cannot synthesize either squalene or lanosterol [49]. Deep-sea *Bathymodiulus* mussels by themselves cannot synthesize cholesterol de

novo and depend mainly on endosymbionts for nutrition [22]. There are no genes encoding enzymes associated with upstream steps of the cholesterol biosynthesis pathway in the Pacific oyster *Crassostrea gigas* [57]. In the present study, 19 enzymes involved in biosynthetic pathways for sterols were identified (Fig. 5) (Table S2). The annotated enzymes in this pathway are activated (Fig. 5B), and farnesyl pyrophosphate synthase (FPPS) displays the

highest expression. This may be attributed to its substrate farnesyl diphosphate which is a branching point to the biosynthesis of vital classes of molecules, such as ubiquinone or dolichol. These results indicated that *L. luymesii* could synthesize cholesterol from acetate or mevalonate via squalene and lanosterol by themselves and suggested that *L. luymesii* might possess cholesterol as its principal sterol.

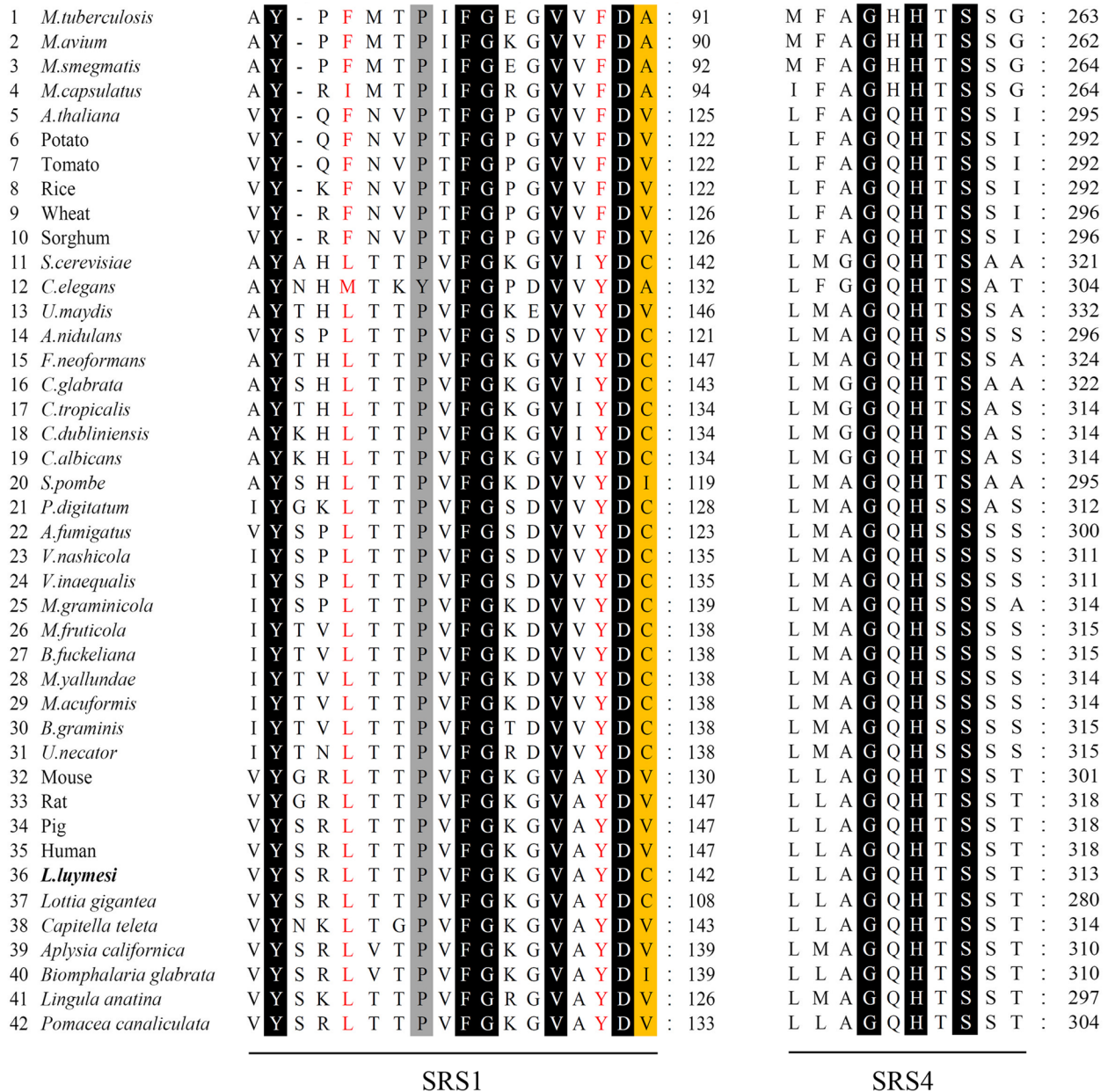


Fig. 6 Sequence alignment of 42 CYP51 family members from different biological kingdoms [bacteria (1–4), plants (5–10), fungi (11, 13–31) and animals (12, 32–42)] in CYP51 substrate recognition sites (SRS) 1 and 4; 100% and more than 95% conserved residues are shaded in black and grey, respectively. Phyla-specific residues are highlighted in red or shaded in yellow. The accession number of sequences can be found in Supplemental data file 1

In this pathway, a *lanosterol 14- α demethylase* (also known as cytochrome P450 family 51 subfamily A member 1, CYP51A1) transcript which has relatively high expression was detected in the transcriptome of *L. luymesii* (Table S2). It is involved in the conversion of lanosterol to 4,4-dimethylcholesta-8(9),14,24-trien-3 β -ol. This demethylation step is the initial checkpoint in the transformation of lanosterol to other sterols widely used within the cell [58]. CYP51A1 is studied primarily in fungi, where it plays an essential role in mediating membrane permeability [59]. CYP51 is an essential enzyme in sterol biosynthetic pathways and the only and evolutionarily oldest P450 gene family with catalytically identical orthologues in different biological kingdoms [58, 60]. The CYP51 proteins have low sequence similarity across phyla but a highly conserved catalytic function. This requires CYP51s have a typical configuration of their substrate-binding pockets and specific amino acid conservation. Most of the conserved CYP51 residues are clustered into six regions, representing substrate recognition sites (SRSs) 1–5 and surrounding the heme-coordinating Cys [60, 61]. The most conserved, SRS1 and SRS4, are regarded as the CYP51 signature. SRS1 forms the upper surface of a P450 substrate-binding cavity. SRS4 is located in the C-terminal part of the P450 I-helix, forming the right wall of the distal surface of the substrate-binding cavity.

SRS1 and SRS4 were found in *L. luymesii* CYP51A1 (LCYP51A1) (Fig. 6). LCYP51A1 SRS1 shares high identity with the human orthologue with only a one-residue difference (residue C). Interestingly, this different residue, C141, was ubiquitous in fungi, and in animals, V is always present at this position, except in one sequence from *Biomphalaria glabrata* (I). Additionally, the residue F/L in SRS1 is phylum-specific (F in plant and L in animal/fungal CYP51). The phylum-specific residue is used to predict preferred substrates (mono- or dimethyl at C4) of newly identified sterol 14 α -demethylases. Thus, we speculated that *L. luymesii* CYP51A1 probably has the physiological substrate C-4 dimethylated and can metabolize four substrates (lanosterol, 24, 25-dihydrolanosterol, 24-methylenedihydrolanosterol, and obtusifoliol) [62].

Phylogenetic analysis suggested that *L. luymesii* CYP51A1 is homologous to other CYP51s. It shared 62% amino acid identity with the human CYP51A1 sequence. CYP51 proteins in fungi and animals clustered within an independent clade, while those in bacteria and plants clustered in another clade on the phylogenetic tree. *L. luymesii* CYP51A1 was assigned to the animal group, and *Capitella teleta* CYP51 formed an independent branch (Fig. 7).

Unexpectedly, a transcript of *cycloartenol-C-24-methyltransferase*, which catalyses the transfer of the methyl group from S-adenosyl-methionine to the C-24 of cycloartenol to form 24-methylene cycloartenol, was identified in *L. luymesii*. There are two copies of the gene in the *L. luymesii* genome. Cycloartenol is an important triterpenoid of the sterol class, which is found in plants. It is the starting point for the synthesis of almost all plant steroids [48], making them chemically distinct from the steroids of fungi and animals that are instead produced from lanosterol. Cycloartenol-C-24-methyltransferase is a sterol C24-methyltransferase (SMT), which are unique to fungi, plants, and protozoa but are not synthesized by animals, as reported previously [63]. Thus, this sequence was identified in animals for the first time.

Alignment against the NCBI database showed that the genes from only *Capitella teleta* and *Thecamonas trahens* ATCC 50,062 were homologous to *L. luymesii* SMT, and they shared 52% and 47% sequence identity, respectively. Most *L. luymesii* SMT orthologues were derived from *Ostreococcus* and higher plants. Plant sterols in animal tissues were detected previously. The analysis of sterol composition obtained by gas chromatography and mass spectrometry showed the presence of plant sterols in the tissues of shallow tropical sponges and cnidarians, as well as their deep-sea counterparts [64]. Significantly, it was proven that most sponges could synthesize sterols de novo and transform cycloartenol and lanosterol, the sterol precursors of photosynthetic and nonphotosynthetic organisms, respectively, into 4,4,14-demethyl sterols. Symbiotic cyanobacteria found in many sponges are not involved in the observed biosynthesis [65, 66]. Thus, our data suggested that *L. luymesii* should be capable of incorporating and transforming cycloartenol to unconventional sterols, while the critical enzyme involved in this process might have similar properties to those in plants. However, the source of cycloartenol in *L. luymesii* is unclear. It might be derived from some photosynthetic symbiotic organism and ingested by tubeworm. As another possibility, the *L. luymesii* SMT arose through lateral gene transfer from the symbiont also deserves consideration. It was predicted that the lipids found in *L. luymesii* might differ significantly from those found in other living organisms. The unique sterol metabolism in *L. luymesii* might play a vital role in membrane structure fluidity and developmental regulation associated with adaptation to a unique environment. This finding might also indicate a candidate gene for research on biological evolution.

Trehalose metabolism

Trehalose is found in various organisms, including bacteria, yeast, fungi, insects, invertebrates, and plants, but it

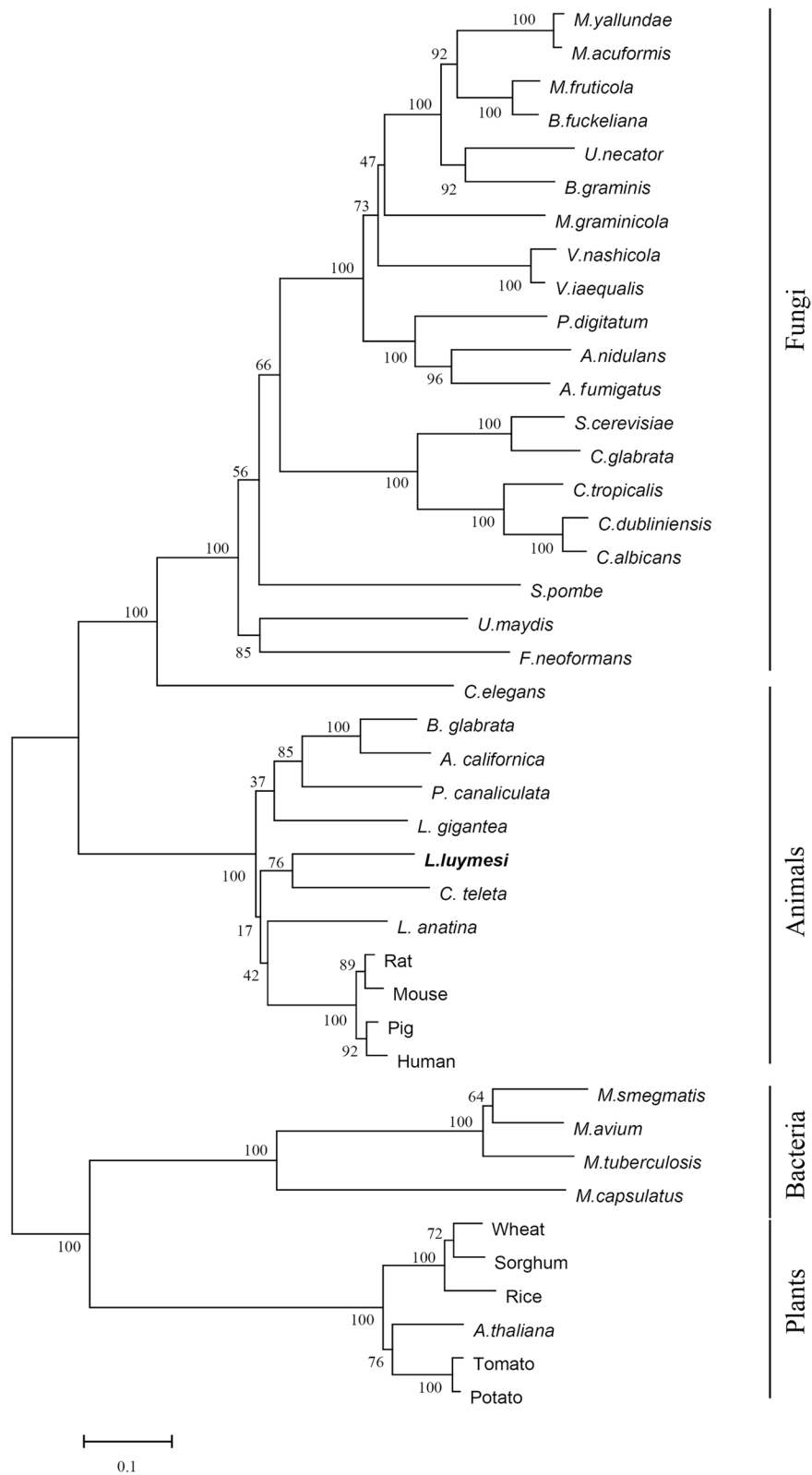


Fig. 7 Phylogenetic tree of 42 CYP51 family members with 1000 bootstrap replications. *L. luymesii* CYP51A1 is highlighted in bold. The accession number of sequences can be found in Supplemental data file 1

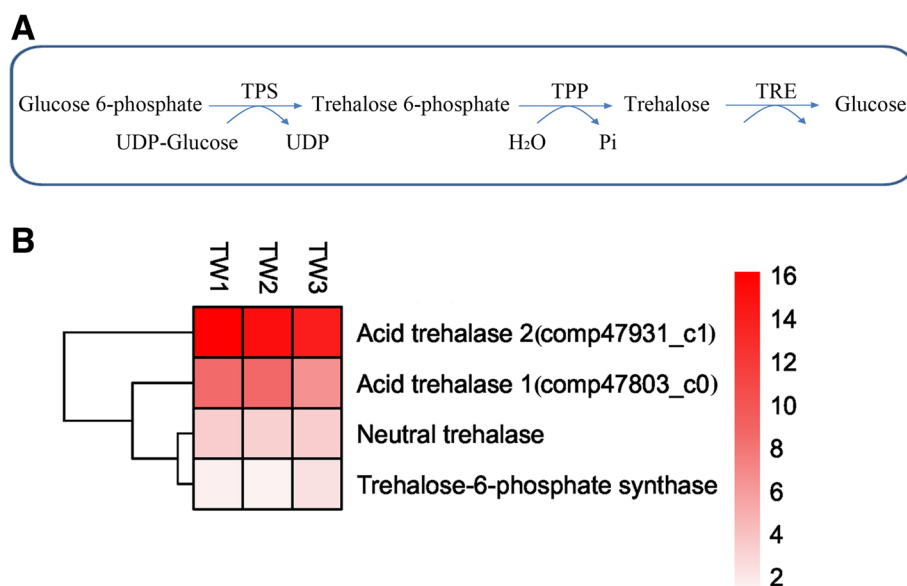


Fig. 8 **A** Trehalose metabolic pathway. TPS, trehalose 6-phosphate synthase; TPP, trehalose 6-phosphate phosphatase; TREH, trehalase. **B** Heatmap of FPKM expression values for the annotated genes in *L. luymesi* trehalose metabolic pathway from three individual samples (TW1, TW2 and TW3). FPKM values were listed in Table S4

is absent in vertebrates. It plays several important physiological roles as a carbon source, reserve carbohydrate, compatible solute, nutritional and environmental stress protector, metabolic regulator, and even signalling transducer or virulence factor, among others [67].

This study identified four putative genes involved in the trehalose metabolic pathway in *L. luymesi* (Fig. 8, Table S3). One putative *trehalose-6-phosphate synthase* (TPS) gene encoded the enzyme that catalyses trehalose synthesis. The deduced *L. luymesi* TPS contains a GT1_TPS domain and a trehalose_PPase domain. At least five different pathways have been described for trehalose biosynthesis in different organisms [23, 68]. However, trehalose is mainly synthesized in eukaryotes by trehalose-6-phosphate synthase (TPS) and trehalose-6-phosphate phosphatase (TPP). TPS catalyses the transfer of glucose from uridine diphosphate glucose to glucose 6-phosphate to generate trehalose 6-phosphate (T6P), whereas TPP catalyses the dephosphorylation of T6P to form trehalose [69–71]. In the present study, we found that the biosynthesis of trehalose in *L. luymesi* is carried out through the TPS/TPP pathway but in a different manner; the process catalysed by a fused gene encoding a protein harbouring conserved TPS/OtsA and TPP/OtsB domains [72], while no TPP gene was identified.

Compared to trehalose synthesis, hydrolysis of trehalose in *L. luymesi* might be more complex. At

least three putative *trehalase* genes (TREs) encode enzymes catalysing the hydrolysis of the sugar. Database searches revealed that the three putative TRE genes encode two forms, namely, neutral (NTH) and acid (ATH) enzymes. Phylogenetic analysis showed that *L. luymesi* NTH was homologous to trehalases from other animals. Moreover, both *L. luymesi* ATHs and fungal ATHs fell into a significant group. The transcript comp47803_c0 (ATH1) with enzymes from other worms further formed a branch, while the unigene comp47931_c1 (ATH2) was assigned another independent branch (Fig. 9).

Neutral trehalase is known to hydrolyse the intracellular pool of trehalose in response to physiological responses, whereas the acid enzyme carries out the hydrolysis of exogenous trehalose [73, 74] and might have “scavenger activity” in fungi. Can the presence of multiple trehalase genes be explained by adaptation to a trehalose-rich diet or by the occurrence of different roles of trehalase? Further research is needed to answer this question. However, our sequence analysis found that ATHs in *L. luymesi* might have different enzyme functions. Previous studies have shown that protein-glucosylgalactosylhydroxylysine glucosidase (PGGHG), which effectively releases glucose from type IV collagen, is encoded by the *ATHL1* gene (acid trehalase-like protein), which has been classified into

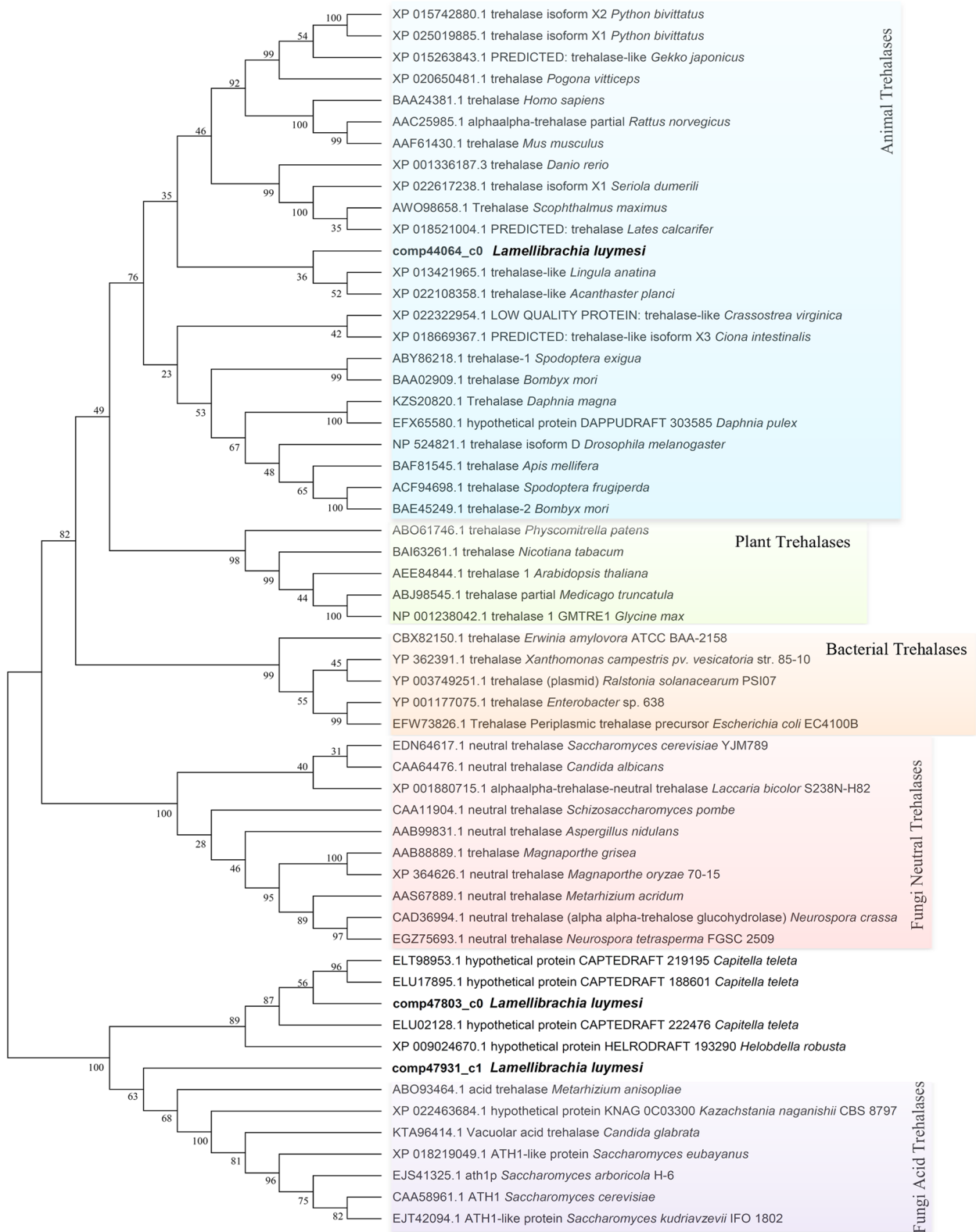


Fig. 9 Phylogenetic tree of trehalase genes and protein isoform diversity in *L. luymesii* and other taxa

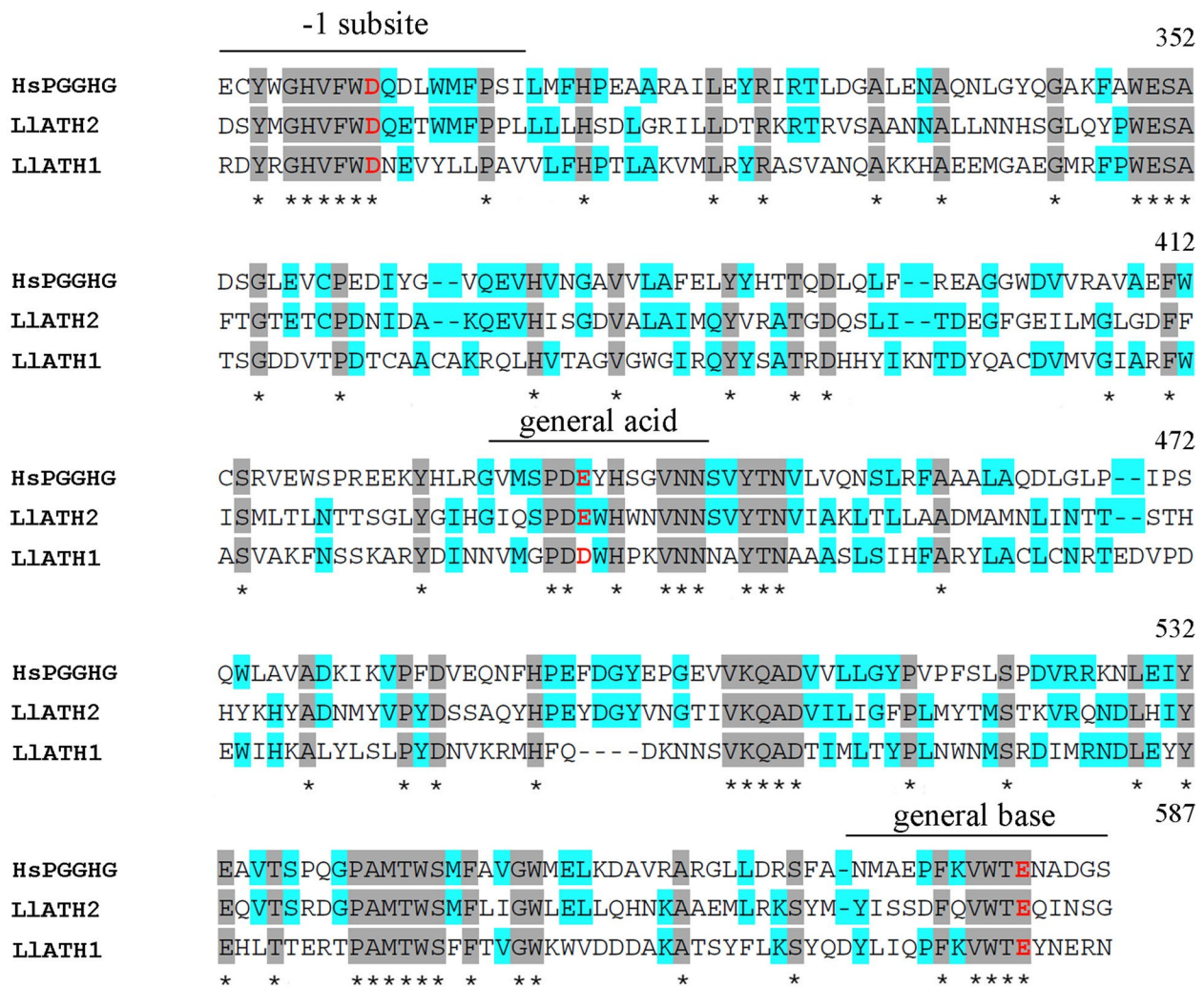


Fig. 10 Partial sequence alignment of *L. luymesii* ATHs and human PGGHG. HsPGGHG, *Homo sapiens* PGGHG (NCBI accession no. NP_079368.3); LlATH1/2, *L. luymesii* acid trehalase 1/2. Identical amino acids are marked by an asterisk and shaded in grey. More than 50% of the conserved residues are shaded in blue. The amino acids in red indicate that three carboxyl residues (corresponding to Asp301, Glu430 and Glu574 of human PGGHG) are essential for catalytic activity

the glycoside hydrolase family 65. Three carboxyl residues (corresponding to Asp301, Glu430, and Glu574 of human PGGHG) were essential for the catalytic activity in this family. The substitution of each of these three residues led to the complete elimination of catalytic activity [75]. Sequence analysis revealed that *L. luymesii* ATHs were members of GH 65, but ATH1 lacks glutamate residues (substitution of Asp) corresponding to human PGGHG Glu430 estimated to function as the catalytic acid, while three carboxyl residues are conserved in ATH2 (Fig. 10). Meanwhile, ATH2 has the highest expression among three trehalases (Fig. 8B). Thus, we speculated that *L. luymesii*

ATH2 might have PGGHG activity, and play more important role in cold-seep tubeworm.

Conclusions

We accomplished high-quality transcriptome assembly of the deep-sea hydrocarbon seep tubeworm *L. luymesii*. We further elucidated the molecular pathways of sulfate activation, cysteine and cholesterol synthesis, and trehalose metabolism. For the first time, two different routes for cysteine biosynthesis described in plants/bacteria and animals were found simultaneously in *L. luymesii*. Moreover, cold-seep tubeworm has the ability of de novo sterol biosynthesis and incorporation and transformation of cycloartenol (plant sterol) and

lanosterol (animal/fungi sterol) into unconventional sterols. At the same time, the critical enzymes involved in this process, such as lanosterol 14- α demethylase and cycloartenol-C-24-methyltransferase, might have properties similar to those of the enzymes from plants or fungi. Additionally, multiple trehalases might play different roles in cold-seep tubeworms. Contrary to the previous analysis, two pathways for cysteine synthesis and the cycloartenol-C-24-methyltransferase gene were identified in animals for the first time. The present study provides new insights into particular adaptations to chemosynthetic environments in *L. luymesii* and can serve as the basis for future molecular studies on host-symbiont interactions and biological evolution.

Materials and methods

Sample collection and RNA extraction

L. luymesii were collected in the northern Gulf of Mexico during an August–September 2006 cruise onboard Research Vessel *Seward Johnson* and submersible Johnson Sea-Link I.

Upon arrival at the sea surface, the trophosome tissue samples from three tubeworms were dissected and stored individually in RNAlater (QIAGEN) and used for RNA extraction. The sample was placed into a sterile mortar prechilled with liquid nitrogen. Additional liquid nitrogen was poured onto the tissue sample. A pestle prechilled in liquid nitrogen was used to crack the frozen tissues into a fine powder. TRI Reagent (Molecular Research Center) was then added, and the sample was mixed using a pestle. RNA extraction was performed according to the manufacturer's protocol. After digestion with RNase-free recombinant DNase I (Takara), RNA was electrophoresed on 1% agarose gels to examine the possibility of RNA degradation and contamination. RNA purity was checked using a NanoPhotometer[®] spectrophotometer (IMPLEN). RNA concentration was measured using a Qubit[®] RNA Assay Kit in a Qubit[®] 2.0 Fluorometer (ThermoFisher Scientific). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent).

Mitochondrial cytochrome c oxidase subunit I gene sequencing

The sample was genotyped by sequencing the mitochondrial gene cytochrome c oxidase subunit I (COI) using previously reported primers [76]. Briefly, genomic DNA was isolated by the phenol–chloroform DNA extraction method. The extracted DNA was used for PCR amplification. PCR was performed using LA Taq DNA polymerase (Takara). The PCR program was as follows: 95 °C for 3 min, followed by 35 cycles of 1 min at 95 °C,

1 min at 40 °C and 1 min at 72 °C and 5 min at 72 °C. The PCR products were purified and sequenced. The COI sequences were analysed using NCBI BLAST online (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Preparation for Illumina sequencing

Total RNA extracted from the sample was used for sequencing library, which was prepared using the NEBNext[®] Ultra[™] RNA Library Prep Kit for Illumina[®] (NEB) following the manufacturer's instructions. Index codes were added to attribute sequences to each sample. Briefly, mRNA was isolated from total RNA using oligo(dT) coupled to magnetic beads. Then, mRNA was fragmented into short fragments using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). Random hexamer primers were used for the first-strand cDNA synthesis, and the second-strand synthesis was subsequently performed using DNA Polymerase I and RNase H. After exchanging remaining overhangs into blunt ends, the cDNA fragments were adenylated at 3' ends and ligated with NEBNext Adaptor. Furthermore, the cDNA was purified with the AMPure XP system (Beckman Coulter) to separate fragments of sizes from 150 to 200 bp. Following PCR amplification, purification, and quality control, the cDNA library was constructed.

Illumina sequencing, transcriptome assembly, and functional annotation

Illumina sequencing, de novo assembly, and functional annotation were performed by Novogene Company Limited, China. Briefly, after the library was qualified, Illumina sequencing was performed using a HiSeq[™] 2000 (Illumina) in paired-end mode with a read length of 150 bp. The image data measured by the high-throughput sequencer were converted into sequence data (reads) by CASAVA base recognition. Raw data (raw reads) in fastq format were first processed using an in-house Perl script to acquire clean data (clean reads) by removing reads containing adapters, reads containing N bases, and low-quality reads. Meanwhile, the Q20, Q30 and GC content of the clean data was calculated. All downstream analyses were based on high-quality clean data. The assembly was accomplished using Trinity with `min_kmer_cov`, which can be set to 2, and all other parameters were set to default values.

Furthermore, the accuracy and completeness of the splicing results were evaluated according to the proportion and completeness of the comparison by BUSCO software. Gene family clustering analysis was performed using Corset (version 4.6) to aggregate transcripts into different clusters based on shared reads. Each cluster

was defined as “Gene”. The prefix “comp” was applied to these genes, such as comp10180_c0 and comp10180_c1. All the transcript sequences were searched against protein databases including NCBI Nr (non-redundant) database (e-value < 0.00001), Swiss-Prot (e-value < 0.00001), and KOG/COG (Clusters of Orthologous Groups of proteins) (e-value < 0.001) using BLASTx (Version: BLAST-2.2.28+). KEGG mapping was performed using KEGG automatic annotation server. Protein domain discovery was performed using the software HMMER 3.0 (hmmscan), with Pfam (Protein family) as the database and an e-value cutoff ≤ 0.01 . Gene Ontology terms information for the annotated transcripts was based on the Nr and Pfam annotation using the software blast2go (b2g4pipe_v2.5).

Sequence alignment and phylogenetic analyses

Characteristic domains or motifs of PAPSS and TPS were identified using a simple modular architecture research tool (SMART) (<http://smart.embl-heidelberg.de/>). Pair-wise and multiple sequence alignments of ATHs were analysed using MAFFT [77]. Multiple sequence alignment of CSs, CYP51s, and trehalases was performed using the ClustalW Multiple Alignment program (<http://www.ebi.ac.uk/clustalw/>). Phylogenetic tree construction was performed in MEGA 3.1 using the neighbour-joining method [78] with 1000 bootstraps.

Abbreviations

TPS	Trehalose-6-phosphate Synthase
TPP	Trehalose-6-phosphate Phosphatase
COI	Mitochondrial Cytochrome c Oxidase Subunit I
PAPSS	Bifunctional 3'-phosphoadenosine-5'-phosphosulfate Synthases
SUL/ATPS	ATP Sulfurylase
KIN/APSK	APS Kinase
FPKM	Fragments per kilobase million
PAP	Adenosine 3', 5'-bisphosphate
DIDS	4,4-Diisothiocyanostilbene-2, 2-disulfonic acid
CYS2/cysE	Serine-O-acetyltransferase
CYS3	Cystathionine gamma-lyase
CYS4	Cystathionine beta-synthase
CS	Cysteine synthase
OAS	O-acetylserine
CYP51A1	Lanosterol 14-alpha demethylase/Cytochrome P450 Family 51 Subfamily A Member 1
SRS	Substrate Recognition Sites
SMT	Sterol C24-methyltransferase
T6P	Trehalose 6-phosphate
TRE	Trehalase
NTH	Neutral Trehalase
ATH	Acid Trehalase
PGGHG	Protein-glucosylgalactosylhydroxylysine Glucosidase

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09267-8>.

Additional file 1: Figure S1a. The length distribution of all the transcripts (a) and unigenes (b).

Additional file 2: Table S1.

Additional file 3: Table S2.

Additional file 4: Table S3.

Additional file 5: Table S4.

Additional file 6.

Acknowledgements

We thank Prof. Fengping Wang for the sample collection. We thank Linfeng Gong for the help with bioinformatic analysis. We thank Betty Hung for the help with revising the manuscript.

Authors' contributions

HS designed the study, performed bioinformatics analysis and wrote the manuscript. LR collected the samples and performed RNA sequencing. ZC performed species identification. YL, WW and LL analyzed data. XX was involved in sample collection. The authors read and approved this manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2018YFC0310702), the Scientific Research Foundation of Third Institute of Oceanography, MNR (No. 2022002) and Major State Basic Research Development Program of China (973 Program, 2015CB755906). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The raw sequencing data have been submitted to the Short Read Archive (SRA) of NCBI under accession number SRR17999638 with Bio project number PRJNA804687 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA804687/>). All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory Breeding Base of Marine Genetic Resources, Key Laboratory of Marine Genetic Resources of Ministry of Natural Resources, Third Institute of Oceanography, Ministry of Natural Resources, Fujian Key Laboratory of Marine Genetic Resources, No. 178 Daxue Road, Xiamen, Fujian 361005, People's Republic of China. ²College of Marine Biology, Xiamen ocean vocational college, 361100 Xiamen, People's Republic of China. ³Co-Innovation Center of Jiangsu Marine Bio-Industry Technology, Jiangsu Ocean University, Lianyungang 222005, People's Republic of China. ⁴School of Advanced Manufacturing, Fuzhou University, Fuzhou 362200, People's Republic of China. ⁵Department of Biology and Guangdong Provincial Key Laboratory of Marine Biotechnology, Shantou University, Shantou 515063, People's Republic of China.

Received: 12 February 2022 Accepted: 20 March 2023

Published online: 05 April 2023

References

- Gardiner S, Hourdez S. On the occurrence of the vestimentiferan tube worm *Lamellibrachia luymesii* van de Land and Norrevang, 1975 (Annelida: Pogonophora) in hydrocarbon seep communities in the Gulf of Mexico. *Biological Society of Washington*. 2003;116:380–94.
- Freytag J. A paradox resolved: Sulfide acquisition by roots of seep tube-worms sustains net chemoautotrophy. *PNAS*. 2001;98:13408–13.
- Flores JF, Fisher CR, Carney SL, Green BN, Freytag JK, Schaeffer SW, Royer WE Jr. Sulfide binding is mediated by zinc ions discovered in the crystal structure of a hydrothermal vent tubeworm hemoglobin. *Proc Natl Acad Sci U S A*. 2005;102(8):2713–8.
- Bailly X, Jollivet D, Vanin S, Deutsch J, Zal F, Lallier F, Toulmond A. Evolution of the sulfide-binding function within the globin multigenic family of the deep-sea hydrothermal vent tubeworm *Riftia pachyptila*. *Mol Biol Evol*. 2002;19(9):1421–33.
- Goffredi SK, Childress JJ, Desaulniers NT, Lallier FJ. Sulfide acquisition by the vent worm *Riftia pachyptila* appears to be via uptake of HS⁻, rather than H₂S. *J Exp Biol*. 1997;200(Pt 20):2609–16.
- Dattagupta S, Miles LL, Barnabei MS, Fisher CR. The hydrocarbon seep tubeworm *Lamellibrachia luymesii* primarily eliminates sulfate and hydrogen ions across its roots to conserve energy and ensure sulfide supply. *J Exp Biol*. 2006;209(Pt 19):3795–805.
- Cordes EE, Arthur MA, Shea K, Arvidson RS, Fisher CR. Modeling the mutualistic interactions between tubeworms and microbial consortia. *PLoS Biol*. 2005;3(3):e77.
- Petersen JM, Zielinski FU, Pape T, Seifert R, Moraru C, Amann R, Hourdez S, Girguis PR, Wankel SD, Barbe V, Pelletier E, Fink D, Borowski C, Bach W, Dubilier N. Hydrogen is an energy source for hydrothermal vent symbioses. *Nature*. 2011;476(7359):176–80.
- Girguis PR, Lee RW, Desaulniers N, Childress JJ, Pospesel M, Felbeck H, Zal F. Fate of nitrate acquired by the tubeworm *Riftia pachyptila*. *Appl Environ Microbiol*. 2000;66(7):2783–90.
- Liao L, Wankel SD, Wu M, Cavanaugh CM, Girguis PR. Characterizing the plasticity of nitrogen metabolism by the host and symbionts of the hydrothermal vent chemoautotrophic symbioses *Ridgeia piscesae*. *Mol Ecol*. 2014;23(6):1544–57.
- De Cian M, Regnault M, Lallier FH. Nitrogen metabolites and related enzymatic activities in the body fluids and tissues of the hydrothermal vent tubeworm *Riftia pachyptila*. *J Exp Biol*. 2000;203(Pt 19):2907–20.
- Lee RW, Robinson JJ, Cavanaugh CM. Pathways of inorganic nitrogen assimilation in chemoautotrophic bacteria-marine invertebrate symbioses: expression of host and symbiont glutamine synthetase. *J Exp Biol*. 1999;202(Pt 3):289–300.
- Sanchez S, Andersen AC, Hourdez S, Lallier FH. Identification, sequencing, and localization of a new carbonic anhydrase transcript from the hydrothermal vent tubeworm *Riftia pachyptila*. *FEBS J*. 2007;274(20):5311–24. <https://doi.org/10.1111/j.1742-4658.2007.06050.x>. Epub 2007 Sep 24.
- Carney SL, Flores JF, Orobona KM, Butterfield DA, Fisher CR, Schaeffer SW. Environmental differences in hemoglobin gene expression in the hydrothermal vent tubeworm, *Ridgeia piscesae*. *Comp Biochem Physiol B Biochem Mol Biol*. 2007;146(3):326–37.
- Zal F, Suzuki T, Kawasaki Y, Childress JJ, Lallier FH, Toulmond A. Primary structure of the common polypeptide chain b from the multi-hemoglobin system of the hydrothermal vent tube worm *Riftia pachyptila*: an insight on the sulfide binding-site. *Proteins*. 1997;29(4):562–74.
- Zal F, Lallier FH, Green BN, Vinogradov SN, Toulmond A. The multi-hemoglobin system of the hydrothermal vent tube worm *Riftia pachyptila*. II. Complete polypeptide chain composition investigated by maximum entropy analysis of mass spectra. *J Biol Chem*. 1996;271(15):8875–81.
- Arp AJ, Childress JJ. Blood function in the hydrothermal vent vestimentiferan tube worm. *Science*. 1981;213(4505):342–4.
- Li Y, Tassia MG, Waits DS, Bogantes VE, David KT, Halanynch KM. Genomic adaptations to chemosymbiosis in the deep-sea seep-dwelling tubeworm *Lamellibrachia luymesii*. *BMC Biol*. 2019;17(1):91.
- Haines TH. Do sterols reduce proton and sodium leaks through lipid bilayers? *Prog Lipid Res*. 2001;40(4):299–324.
- Love GD, Grosjean E, Stalvies C, Fike DA, Grotzinger JP, Bradley AS, Kelly AE, Bhatia M, Meredith W, Snape CE, Bowring SA, Condon DJ, Summons RE. Fossil sterols record the appearance of Mesozoic forams during the Cryogenian period. *Nature*. 2009;457(7230):718–21.
- Bloch K. The biological synthesis of cholesterol. *Science*. 1965;150(3692):19–28. <https://doi.org/10.1126/science.150.3692.19>.
- Takishita K, Takaki Y, Chikaraishi Y, Ikuta T, Ozawa G, Yoshida T, Ohkouchi N, Fujikura K. Genomic evidence that methanotrophic endosymbionts likely provide deep-sea bathymodiolus mussels with a sterol intermediate in cholesterol biosynthesis. *Genome Biol Evol*. 2017;9(5):1148–60.
- Elbein AD, Pan YT, Pastuszak I, Carroll D. New insights on trehalose: a multifunctional molecule. *Glycobiology*. 2003;13(4):17R–27R.
- Elbein AD. The metabolism of alpha, alpha-trehalose. *Adv Carbohydr Chem Biochem*. 1974;30:227–56.
- Colaço CALS, Roser B. Trehalose—a multifunctional additive for food preservation. In: Mathlouthi M, editor. *Food Packaging and Preservation*. Boston, MA: Springer, US; 1994. p. 123–40.
- Asano N. Glycosidase inhibitors: update and perspectives on practical use. *Glycobiology*. 2003;13(10):93R–104R.
- Islam M, Schulze-Makuch D. Adaptations to environmental extremes by multicellular organisms. *Int J Astrobiol*. 2007;6(3):199–215.
- Behm CA. The role of trehalose in the physiology of nematodes. *Int J Parasitol*. 1997;27(2):215–29.
- Shukla E, Thorat LJ, Nath BB, Gaikwad SM. Insect trehalase: physiological significance and potential applications. *Glycobiology*. 2015;25(4):357–67.
- Kim CB, Moon SY, Gelder SR, Kim W. Phylogenetic relationships of annelids, molluscs, and arthropods evidenced from molecules and morphology. *J Mol Evol*. 1996;43(3):207–15.
- Patron NJ, Durnford DG, Kopriva S. Sulfate assimilation in eukaryotes: fusions, relocations and lateral transfers. *BMC Evol Biol*. 2008;4(8):39.
- Rosenthal E, Leustek T. A multifunctional *Urechis caupo* protein, PAPS synthetase, has both ATP sulfurylase and APS kinase activities. *Gene*. 1995;165(2):243–8.
- Bowman KG, Bertozzi CR. Carbohydrate sulfotransferases: mediators of extracellular communication. *Chem Biol*. 1999;6(1):R9–22.
- Hemmerich S. Carbohydrate sulfotransferases: novel therapeutic targets for inflammation, viral infection and cancer. *Drug Discov Today*. 2001;6(1):27–35.
- Hemmerich S, Rosen SD. Carbohydrate sulfotransferases in lymphocyte homing. *Glycobiology*. 2000;10(9):849–56.
- Mueller JW, Gilligan LC, Idkowiak J, Artl W, Foster PA. The regulation of steroid action by sulfation and desulfation. *Endocr Rev*. 2015;36(5):526–63.
- Vincourt J, Jullien D, Amalric F, Girard J. Molecular and functional characterization of SLC26A11, a sodium-independent sulfate transporter from high endothelial venules. *FASEB J*. 2003;17:1–21.
- Soleimani M. SLC26 Cl⁻/HCO₃⁻ exchangers in the kidney: roles in health and disease. *Kidney Int*. 2013;84(4):657–66.
- Sun J, Zhang Y, Xu T, Zhang Y, Mu H, Zhang Y, Lan Y, Fields CJ, Hui JHL, Zhang W, Li R, Nong W, Cheung FKM, Qiu JW, Qian PY. Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nat Ecol Evol*. 2017;1(5):121.
- Hell R. Molecular physiology of plant sulfur metabolism. *Planta*. 1997;202(2):138–48. <https://doi.org/10.1007/s004250050112>. PMID: 9202491.
- Wyboun N, Dermauw W, Tirry L, Stevens C, Grbić M, Feyereisen R, Van Leeuwen T. A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. *Elife*. 2014;3:e02365. <https://doi.org/10.7554/eLife.02365>. PMID: 24843024; PMCID: PMC4011162.
- Paszewski A. Sulfur amino acid metabolism and its regulation in fungi: studies with *Aspergillus nidulans*. *Acta Biochim Pol*. 1993;40(4):445–9. PMID: 8140816.
- Kerksick C, Willoughby D. The antioxidant role of glutathione and N-acetyl-cysteine supplements and exercise-induced oxidative stress. *J Int Soc Sports Nutr*. 2005;2(2):38–44. <https://doi.org/10.1186/1550-2783-2-2-38>. PMID: 18500954; PMCID: PMC2129149.
- Guo L, Ganguly A, Sun L, Suo F, Du LL, Russell P. Global Fitness Profiling Identifies Arsenic and Cadmium Tolerance Mechanisms in Fission Yeast G3 (Bethesda). 2016 Oct 13;6(10):3317–3333. <https://doi.org/10.1534/g3.116.033829>. PMID: 27558664; PMCID: PMC5068951.
- Waits DS, Santos SR, Thornhill DJ, Li Y, Halanynch KM. Evolution of Sulfur Binding by Hemoglobin in Siboglinidae (Annelida) with Special Reference to Bone-Eating Worms. *Osedax J Mol Evol*. 2016;82(4–5):219–29. <https://doi.org/10.1007/s00239-016-9739-7>. Epub 2016 Apr 21 PMID: 27100359.

46. Ma Y, Sugiura R, Saito M, Koike A, Sio SO, Fujita Y, Takegawa K, Kuno T. Six new amino acid-auxotrophic markers for targeted gene integration and disruption in fission yeast. *Curr Genet.* 2007;52(2):97–105. <https://doi.org/10.1007/s00294-007-0142-1>. Epub 2007 Jul 11 PMID: 17625533.
47. Nes WR, McKean ML. Biochemistry of steroids and other isopentenoids. 1977.
48. Schaller H. The role of sterols in plant growth and development. *Prog Lipid Res.* 2003;42(3):163–75. [https://doi.org/10.1016/s0163-7827\(02\)00047-4](https://doi.org/10.1016/s0163-7827(02)00047-4). PMID: 12689617.
49. Rauthan M, Pilon M. The mevalonate pathway in *C. elegans*. *Lipids Health Dis.* 2011;10:243. <https://doi.org/10.1186/1476-511X-10-243>. PMID: 22204706; PMCID: PMC3274489.
50. Kurzchalia TV, Ward S. Why do worms need cholesterol? *Nat Cell Biol.* 2003;5(8):684–8. <https://doi.org/10.1038/ncb0803-684>. PMID: 12894170.
51. Wootton JM, Wright LD. Biosynthesis of squalene by the annelid *Lumbricus terrestris*. *Nature.* 1960;17(187):1027–8. <https://doi.org/10.1038/1871027b0>. PMID: 13786660.
52. Voogt PA, van Rheeën JW, Zandee DI. What about squalene in the earthworm *Lumbricus terrestris*? *Comp Biochem Physiol B.* 1975;50(4):511–3. [https://doi.org/10.1016/0305-0491\(75\)90080-2](https://doi.org/10.1016/0305-0491(75)90080-2). PMID: 1122731.
53. Walton MJ, Pennock JF. Some studies on the biosynthesis of ubiquinone, isoprenoid alcohols, squalene and sterols by marine invertebrates. *Biochem J.* 1972;127(3):471–9. <https://doi.org/10.1042/bj1270471>. PMID: 4403925; PMCID: PMC1178687.
54. Goad LJ. Sterol biosynthesis and metabolism in marine invertebrates. *Pure Appl Chem.* 1981;53(4):837–52.
55. Knauer J, Kerr RG, Lindley D, Southgate PC. Sterol metabolism of pacific oyster (*Crassostrea gigas*) spat. *Comp Biochem Physiol B: Biochem Mol Biol.* 1998;119(1):81–4.
56. Giner JL, Zhao H, Dixon MS, Wikfors GH. Bioconversion of (13)C-labeled microalgal phytosterols to cholesterol by the Northern Bay scallop, *Argopecten irradians irradians*. *Comp Biochem Physiol B Biochem Mol Biol.* 2016;192:1–8. <https://doi.org/10.1016/j.cbpb.2015.11.003>. Epub 2015 Nov 11 PMID: 26577022.
57. Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PW, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A, Hedgecock D, Xu Z, Liu Y, Domazet-Lošo T, Du Y, Sun X, Zhang S, Liu B, Cheng P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li N, Wang J, Chen M, He Y, Tan F, Song X, Zheng Q, Huang R, Yang H, Du X, Chen L, Yang M, Gaffney PM, Wang S, Luo L, She Z, Ming Y, Huang W, Zhang S, Huang B, Zhang Y, Qu T, Ni P, Miao G, Wang J, Wang Q, Steinberg CE, Wang H, Li N, Qian L, Zhang G, Li Y, Yang H, Liu X, Wang J, Yin Y, Wang J. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature.* 2012;490(7418):49–54. <https://doi.org/10.1038/nature11413>. Epub 2012 Sep 19 PMID: 22992520.
58. Lepesheva GI, Waterman MR. Sterol 14 α -demethylase cytochrome P450 (CYP51), a P450 in all biological kingdoms. *Biochim Biophys Acta.* 2007;1770(3):467–77. <https://doi.org/10.1016/j.bbagen.2006.07.018>. Epub 2006 Aug 2. PMID: 16963187; PMCID: PMC2324071.
59. Daum G, Lees ND, Bard M, Dickson R. Biochemistry, cell biology and molecular biology of lipids of *Saccharomyces cerevisiae*. *Yeast.* 1998;14(16):1471–510.
60. Yoshida Y, Aoyama Y, Noshiro M, Gotoh O. Sterol 14-demethylase P450 (CYP51) provides a breakthrough for the discussion on the evolution of cytochrome P450 gene superfamily. *Biochem Biophys Res Commun.* 2000;273(3):799–804. <https://doi.org/10.1006/bbrc.2000.3030>. PMID: 10891326.
61. Gotoh O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J Biol Chem.* 1992;267(1):83–90. PMID: 1730627.
62. Lepesheva GI, Nes WD, Zhou W, Hill GC, Waterman MR. CYP51 from *Trypanosoma brucei* is obtusifoliol-specific. *Biochemistry.* 2004;43(33):10789–99. <https://doi.org/10.1021/bi048967t>. PMID: 15311940.
63. Pereira M, Song Z, Santos-Silva LK, Richards MH, Nguyen TT, Liu J, de Almeida Soares CM, da Silva Cruz AH, Ganapathy K, Nes WD. Cloning, mechanistic and functional analysis of a fungal sterol C24-methyltransferase implicated in brassicasterol biosynthesis. *Biochim Biophys Acta.* 2010;1801(10):1163–74. <https://doi.org/10.1016/j.bbali.2010.06.007>. Epub 2010 Jul 17 PMID: 20624480.
64. Carreón-Palau L, Özdemir NŞ, Parrish CC, Parzanini C. Sterol Composition of Sponges, Cnidarians, Arthropods, Mollusks, and Echinoderms from the Deep Northwest Atlantic: A Comparison with Shallow Coastal Gulf of Mexico. *Mar Drugs.* 2020;18(12):598. <https://doi.org/10.3390/md18120598>. PMID:33260983;PMCID:PMC7761341.
65. Kerr RG, Stoilov IL, Thompson JE, Djerassi C. Biosynthetic studies of marine lipids 16. De novo sterol biosynthesis in sponges. Incorporation and transformation of cycloartenol and lanosterol into unconventional sterols of marine and freshwater sponges. *Tetrahedron.* 1989;45(7):1893–904.
66. de Barros IB, Volkmer-Ribeiro C, da Veiga Junior VF. Sterols from sponges of Anavilhanas. *Biochem Syst Ecol.* 2013;49:167–71.
67. Argüelles JC. Physiological roles of trehalose in bacteria and yeasts: a comparative analysis. *Arch Microbiol.* 2000;174(4):217–24. <https://doi.org/10.1007/s002030000192>. Erratum in: *Arch Microbiol* 2000 Dec;174(6):456. PMID: 11081789.
68. Avonce N, Mendoza-Vargas A, Morett E, Iturriaga G. Insights on the evolution of trehalose biosynthesis. *BMC Evol Biol.* 2006;19(6):109. <https://doi.org/10.1186/1471-2148-6-109>. PMID:17178000;PMCID:PMC1769515.
69. Bell W, Klaassen P, Ohnacker M, Boller T, Herweijer M, Schoppink P, Van der Zee P, Wiemken A. Characterization of the 56-kDa subunit of yeast trehalose-6-phosphate synthase and cloning of its gene reveal its identity with the product of C1F1, a regulator of carbon catabolite inactivation. *Eur J Biochem.* 1992;209(3):951–9. <https://doi.org/10.1111/j.1432-1033.1992.tb17368.x>. PMID: 1425702.
70. Vuorio OE, Kalkkinen N, Londesborough J. Cloning of two related genes encoding the 56-kDa and 123-kDa subunits of trehalose synthase from the yeast *Saccharomyces cerevisiae*. *Eur J Biochem.* 1993;216(3):849–61. <https://doi.org/10.1111/j.1432-1033.1993.tb18207.x>. PMID: 8404905.
71. Song XS, Li HP, Zhang JB, Song B, Huang T, Du XM, Gong AD, Liu YK, Feng YN, Agboola RS, Liao YC. Trehalose 6-phosphate phosphatase is required for development, virulence and mycotoxin biosynthesis apart from trehalose biosynthesis in *Fusarium graminearum*. *Fungal Genet Biol.* 2014;63:24–41. <https://doi.org/10.1016/j.fgb.2013.11.005>. Epub 2013 Nov 27 PMID: 24291007.
72. Cui S-Y, Xia Y-X. Isolation and characterization of the trehalose-6-phosphate synthase gene from *Locusta migratoria manilensis*. *Insect Science.* 2009;16:287–95. <https://doi.org/10.1111/j.1744-7917.2009.01268.x>. PMID: 8404905.
73. Maicas S, Guirao-Abad JP, Argüelles JC. Yeast trehalases: two enzymes, one catalytic mission. *Biochim Biophys Acta.* 2016;1860(10):2249–54. <https://doi.org/10.1016/j.bbagen.2016.04.020>. Epub 2016 Apr 28 PMID: 27133444.
74. Lopes RG, Muñoz JE, Barros LM, Alves-Jr SL, Taborda CP, Stambuk BU. The secreted acid trehalase encoded by the *CgATH1* gene is involved in *Candida glabrata* virulence. *Mem Inst Oswaldo Cruz.* 2020;115:e200401. doi: <https://doi.org/10.1590/0074-02760200401>. PMID: 33146242; PMCID: PMC7607559.
75. Hamazaki H, Hamazaki MH. Catalytic site of human protein-glucosylglactosylhydroxylysine glucosidase: Three crucial carboxyl residues were determined by cloning and site-directed mutagenesis. *Biochem Biophys Res Commun.* 2016;469(3):357–62. <https://doi.org/10.1016/j.bbrc.2015.12.005>. Epub 2015 Dec 9 PMID: 26682924.
76. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol.* 1994;3(5):294–9.
77. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 2002;30(14):3059–66. <https://doi.org/10.1093/nar/gkf436>. PMID:12136088;PMCID:PMC135756.
78. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform.* 2004;5(2):150–63. <https://doi.org/10.1093/bib/5.2.150>. PMID: 15260895.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.