Lin et al. *Zool. Res.* 2023, 44(1): 106−125
https://doi.org/10.24272/j.issn.2095-8137.2022.279

ZR | Zoological Research

Article Open Access

Interactions among deep-sea mussels and their epibiotic and endosymbiotic chemoautotrophic bacteria: Insights from multi-omics analysis

Yi-Tao Lin[1,2,#], Ting Xu[2,3,#], Jack Chi-Ho Ip[1,2,#], Yanan Sun[1,2], Ling Fang[4], Tiangang Luan[5,6], Yu Zhang[7,*], Pei-Yuan Qian[2,3,*], Jian-Wen Qiu[1,2,*]

[1] *Department of Biology*, *Hong Kong Baptist University*, *Hong Kong SAR*, *China*
[2] *Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou)*, *Guangzhou*, *Guangdong* 511458, *China*
[3] *Department of Ocean Science*, *Hong Kong University of Science and Technology*, *Hong Kong SAR*, *China*
[4] *Instrumental Analysis & Research Center*, *Sun Yat-Sen University*, *Guangzhou*, *Guangdong* 510875, *China*
[5] *State Key Laboratory of Biocontrol*, *School of Life Sciences*, *Sun Yat-Sen University*, *Guangzhou*, *Guangdong* 510875, *China*
[6] *Institute of Environmental and Ecological Engineering*, *Guangdong University of Technology*, *Guangzhou*, *Guangdong* 510006, *China*
[7] *College of Life Sciences and Oceanography*, *Shenzhen University*, *Shenzhen*, *Guangdong* 518060, *China*

ABSTRACT

Endosymbiosis with Gammaproteobacteria is fundamental for the success of bathymodioline mussels in deep-sea chemosynthesis-based ecosystems. However, the recent discovery of Campylobacteria on the gill surfaces of these mussels suggests that these host-bacterial relationships may be more complex than previously thought. Using the cold-seep mussel (*Gigantidas haimaensis*) as a model, we explored this host-bacterial system by assembling the host transcriptome and genomes of its epibiotic Campylobacteria and endosymbiotic Gammaproteobacteria and quantifying their gene and protein expression levels. We found that the epibiont applies a sulfur oxidizing (SOX) multienzyme complex with the acquisition of *soxB* from Gammaproteobacteria for energy production and switched from a reductive tricarboxylic acid (rTCA) cycle to a Calvin-Benson-Bassham (CBB) cycle for carbon assimilation. The host provides metabolic intermediates, inorganic carbon, and thiosulfate to satisfy the materials and energy requirements of the epibiont, but whether the epibiont benefits the host is unclear. The endosymbiont adopts methane oxidation and the ribulose monophosphate pathway (RuMP) for energy production, providing the major source of energy for itself and the host. The host obtains most of its nutrients, such as lysine, glutamine, valine, isoleucine, leucine, histidine, and folate, from the endosymbiont. In addition, host pattern recognition receptors, including toll-like receptors, peptidoglycan recognition proteins, and C-type lectins, may

Received: 19 October 2022; Accepted: 21 November 2022; Online: 22 November 2022

Foundation items: This study was supported by the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0409, SMSEGL20SC02), Research Grants Council of Hong Kong (12101021), and Guangdong Natural Science Foundation (2020A1515011117)

[#]Authors contributed equally to this work

*Corresponding authors, E-mail: biozy@szu.edu.cn; boqianpy@ust.hk; qiujw@hkbu.edu.hk

participate in bacterial infection, maintenance, and population regulation. Overall, this study provides insights into the complex host-bacterial relationships that have enabled mussels and bacteria to thrive in deep-sea chemosynthetic ecosystems.

Keywords: Bathymodioline; Chemosynthesis; Cold seep; Deep-sea adaptation; Symbiosis

INTRODUCTION

Hydrothermal vents and cold seeps are considered extreme deep-sea habitats due to the dark conditions, high hydraulic pressure, variable temperatures, and high concentrations of toxic substances (Jing et al., 2020; Le Bris et al., 2016; Levin, 2005; Tunnicliffe, 1991). Nevertheless, the release of abundant reducing chemicals, such as hydrogen sulfide, methane, and hydrogen, from the seafloor of these habitats promotes bacterial chemosynthesis to support thriving communities of invertebrates such as giant tubeworms, clams, mussels, and snails (Sen et al., 2018; Van Dover, 2000). Symbiosis between deep-sea invertebrates and chemosynthetic bacteria has allowed the colonization of these deep-sea habitats despite very limited input of organic matter from photosynthesis in surface waters (Childress et al., 1986; Dubilier et al., 2008; Osman & Weinnig, 2022). The formation of symbiosis requires substantial structural changes in the host, such as the replacement of the digestive tract by trophosomes (bacterial-hosting organ in the trunk) in giant tubeworms and degeneration of the digestive tract in bathymodioline mussels and vesicomyid clams (Osman & Weinnig, 2022).

Research over the past four decades has revealed the diversity of chemosynthetic holobionts with respect to habitats, symbiont locations, and host and symbiont taxa (Ansorge et al., 2019; Dubilier et al., 2008; Lim et al., 2019; Osman & Weinnig, 2022). Among holobionts, bathymodioline mussels have been studied extensively (Duperron et al., 2008; Laming et al., 2018; Lin et al., 2021; Roeselers & Newton, 2012), with a focus on endosymbiotic Gammaproteobacteria that live inside host bacteriocytes – epithelial cells that contain symbionts. These bacteria can be further divided into two groups based on their energy production pathways: i.e., sulfur-oxidizing bacteria (SOB), which produce energy through oxidation of sulfur and sulfuric compounds, and methane-oxidizing bacteria (MOB), which produce energy through methane oxidation (Ansorge et al., 2019; Laming et al., 2018; Roeselers & Newton, 2012). While some species of deep-sea mussels host only one species of endosymbiotic Gammaproteobacteria, others can harbor up to six different species (Duperron et al., 2008). Previous genomic, transcriptomic, and proteomic studies have provided in-depth knowledge about the relationships between hosts and endosymbiotic SOB and MOB, especially with regard to symbiont energy production and nutrient complementarity between symbiotic partners (Ponnudurai et al., 2017a, 2017b, 2020; Sun et al., 2017).

Recent studies have discovered Campylobacteria (Campylobacterota, formerly Epsilonproteobacteria) (Waite et al., 2018) on the gill epithelium of *Gigantidas platifrons* (formerly *Bathymodiolus platifrons*), *G. childressi* (formerly *B. childressi*), and *B. azoricus* (Assié et al., 2016, 2020; Sun et al., 2022), suggesting that host-bacterial associations in deep-sea mussels may be more complex than previously thought. Campylobacteria, abundant in hydrothermal vents and cold seep environments (Campbell et al., 2006; Hou et al., 2020), have been identified as endosymbionts of deep-sea gastropods and epibionts of deep-sea shrimp and annelids (Dubilier et al., 2008; Goffredi, 2010; Lan et al., 2021). Furthermore, transintegumental transfer of inorganic carbon has been reported between the vent shrimp (*Rimicaris exoculata*) and its epibiotic bacteria (Ponsard et al., 2013). Assié et al. (2016) hypothesized that Campylobacteria in bathymodiolines occupy the same ecological niche as endosymbiotic SOB and if they perform sulfur-oxidizing chemosynthesis, and thus may have displaced the latter in some bathymodioline species. Assié et al. (2020) also suggested that horizontal gene transfer (HGT) may have played a role in the association between Campylobacteria and bathymodiolines. Free-living Campylobacteria rely on the reductive tricarboxylic acid (rTCA) cycle for carbon fixation, whereas bathymodioline epibiotic Campylobacteria rely on the Calvin-Benson-Bassham (CBB) cycle acquired through HGT from Betaproteobacteria and sulfur-oxidizing Gammaproteobacteria for carbon fixation (Assié et al., 2020). However, whether other metabolic pathways are affected by HGT, and the relationship between the host and these epibionts or endosymbiotic Gammaproteobacteria remain unclear.

The bathymodioline *G. haimaensis* (Bivalvia: Mytilidae) is a dominant epibenthic species found in the Haima cold seep on the northwestern continental slope of the South China Sea (Figure 1A, B). According to morphological and phylogenetic analysis, it is closely related to *G. platifrons*, *G. mauritanicus*, and *G. childressi* (Xu et al., 2019). Based on high-throughput sequencing of the V3–V4 region of the *16S rRNA* gene, the bacterial community of the *G. haimaensis* gill is dominated by methanotrophic Gammaproteobacteria (92.1% sequences) and Campylobacteria (5.0% sequences) (Xu et al., 2019). However, the spatial distribution of these bacteria on *G. haimaensis*, as well as their biological interactions, have not yet been reported. To address these issues, we conducted an integrated morphological and multi-omics study of *G. haimaensis*. Specifically, we used light and electron microscopy and fluorescence *in situ* hybridization (FISH) to identify and locate the epibiont and endosymbiont, metagenomics to assemble the genomes of these two bacteria, and metatranscriptomics and metaproteomics to determine the gene and protein expression levels of the three symbiotic partners. This integrative approach can be applied to other holobiont systems to enhance our understanding of the diversity and evolution of chemosynthesis-based symbiosis in the animal kingdom.

MATERIALS AND METHODS

Sample collection and fixation

Specimens of *G. haimaensis* were collected from the Haima

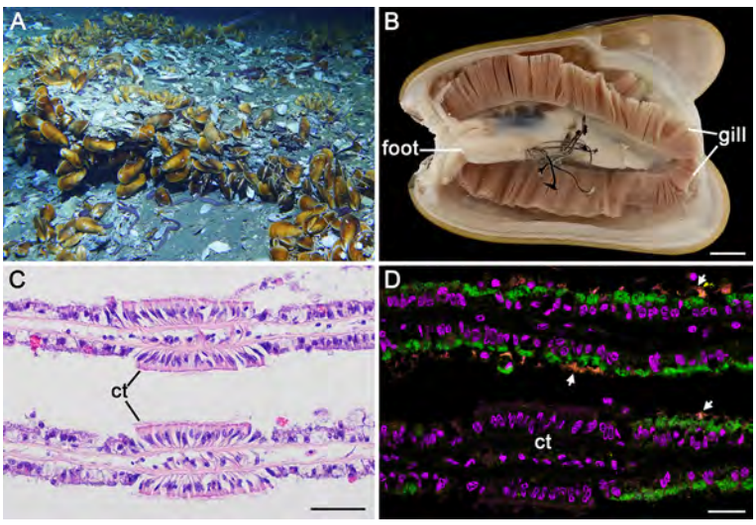

Figure 1 Habitat, morphology, and gill sections of *G. haimaensis*

A: Photograph showing a *G. haimaensis* population attached to authigenic carbonate rocks in Haima cold seep (Photo credit: Dr. Jun Tao from Guangzhou Marine Geological Survey (Guangzhou, China)). B: Dissected individual showing gills and foot. Scale bar: 1 cm. C: Gill tissue section colored by H&E stain. Scale bar: 50 μm. D: Fluorescence *in situ* hybridization (FISH) of a gill section showing distribution of epibiotic Campylobacteria (orange, probe EPSY549, Cy5, indicated by arrows) and endosymbiotic Gammaproteobacteria (green, probe IMedM-138, Cy3), as well as nucleus (purple, DAPI). Scale bar: 20 μm. ct: ciliary tuft.

cold seep (N16°43.80′, E110°28.50′) in the South China Sea at a depth of ~1 400 m for DNA and RNA extraction (Mussels 1–3, 29 April 2018), protein extraction (Mussels 4–6, 1 September 2020), and paraffin sectioning and electron microscopy (Mussel 7, 25 May 2021) using the remotely operated vehicle (ROV) *Haima 2* on board the research vessel (R/V) *Haiyang* 6 of the Guangzhou Marine Geological Survey, China. Typical habitat of *G. haimaensis* is shown in Figure 1A. Upon reaching the main deck, the specimens were immediately frozen at −80 °C for DNA extraction, dissected with the gill and foot then frozen at −80 °C for RNA extraction (Figure 1B), or dissected and fixed in 4% paraformaldehyde (PFA) for 12 h at 4 °C, then rinsed with Milli-Q water, dehydrated with an ethanol gradient, and stored at 4 °C for sectioning or in 2.5% glutaraldehyde at 4 °C for electron microscopy analysis.

Hematoxylin-eosin (H&E) staining and FISH analysis

H&E staining and FISH experiments were performed to visualize the gill morphology and distribution of the bacteria. Gill samples fixed in PFA and dehydrated with ethanol were soaked with xylene and embedded in Paraplast (Sigma, USA). Each paraffin block was cut into 6 μm sections using a RM 2126 microtone (Leica, Germany). Xylene was then used to remove the Paraplast, followed by rehydration. The sections were stained by H&E (Abcam, UK) and observed under a differential interference microscope (Olympus BX51, Japan). In addition, the rehydrated sections were treated with PBST (0.1% Tween 20 in phosphate-buffered saline (PBS)) to increase permeability for the FISH experiments. After pre-hybridization in formamide hybridization buffer without probes for 30 min at 46 °C, the sections were hybridized in formamide hybridization buffer (0.9 mol/L NaCl, 0.02 mol/L Tris-HCl, 0.01% sodium dodecyl-sulfate (SDS), 35% deionized formamide) with 0.5 μmol/L of each probe for 1 h at 46 °C. A washing buffer (0.1 mol/L NaCl, 0.02 mol/L Tris-HCl, 5 mmol/L ethylenediaminetetraacetic acid (EDTA), 0.01% SDS) was applied to resin the sections for 15 min at 48 °C. The sections were stained using 4',6-diamidino-2-phenylindole (DAPI) (Sigma, USA) for 3 min at room temperature, then washed with PBS buffer and covered with Antifade Mounting Medium (Beyotime, China) and a cover slip. The EPSY549 probe (Lin et al., 2006) and Cy5 dye were used to detect Campylobacteria. The IMedM-138 probe (Duperron et al., 2008) and Cy3 dye were used to identify Gammaproteobacteria. The sections were observed under a LSM 710 NLO laser scanning confocal microscope (ZEISS, Germany) and micrographs were analyzed using ZEN v3.3 (ZEISS, Germany).

Electron microscopy

Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) were used to confirm the distribution of the endosymbionts. The gill samples preserved in glutaraldehyde were washed in PBS. For SEM, the samples were dehydrated with a gradient of ethanol solutions. The samples were then transferred to 50% and 100% isoamyl acetate for the displacement of ethanol, then dried in a carbon dioxide critical point dryer XD-1 (EIKO, Japan). Conductive coating was performed using an ion coater IB-3 (EIKO, Japan) and micrographs were taken by a JSM-840 scanning electron microscope (JEOL, Japan). For TEM, the washed tissues were transferred to 1% osmic acid for further fixation at 4 °C for 2 h, rinsed in PBS, dehydrated in a gradient of ethanol solutions, and embedded in epoxy resin (EPON 812). After polymerization at 37 °C, 45 °C, and 65 °C for 24 h, respectively, ultrathin sections (70 nm) were cut with an EM UC7 ultramicrotome (Leica, Germany), stained with uranyl acetate then lead nitrate, and imaged with a JEM-1200EX transmission electron microscope (JEOL, Japan).

DNA extraction and metagenomic sequencing

Genomic DNA was extracted from the gill tissues of mussels 1–3 using the CTAB method (Stewart & Via, 1993). DNA quality was examined using 1.0% agarose gel electrophoresis, and DNA was quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, USA). A total of 1 μg of DNA per individual was used for library construction (insert size of 350 bp) using a NEBNext® Ultra™ DNA Library Prep Kit for Illumina (NEB, USA) following the manufacturer's protocols. The library was sequenced on an Illumina NovaSeq 6000 sequencer (Illumina, USA) to produce 150 bp paired-end reads at Novogene (Beijing, China).

Metagenomic assembly, genome binning and functional annotation

Raw Illumina reads were filtered using Trimmomatic v0.38 (Bolger et al., 2014) to remove adapter and low-quality reads under the following settings: LEADING=15, TRAILING=15, SLIDINGWINDOW=4:20, MINLEN=40. The clean reads of each sample were individually assembled using SPAdes v3.14.1 (Bankevich et al., 2012) in metagenomic mode with *k*-mer sizes of 21, 33, 55, 77, 99, and 127. Only scaffolds ≥1 kb were retained. To determine the symbiotic species in each sample, the *16S rRNA* gene sequences were recovered from the assembled scaffolds using Barrnap v0.9 (github.com/tseemann/barrnap), and BLAST searched against the National Center for Biotechnology Information Non-Redundant (NCBI NR) database. The *16S rRNA* sequences were further aligned with the sequences obtained from Xu et al. (2019) using MAFFT v7.475 (Katoh & Standley, 2013), confirming the presence of two bacteria in each gill sample. To obtain the bacterial genomes, genome binning was performed using MetaWRAP v1.2.1 (Uritskiy et al., 2018) with the "Binning", "Bin_refinement", and "Reassemble_bins" modules under default settings. Abundance of each bin was calculated using the "Quant_bins" module under default settings. Both the completeness and potential contamination of each assembled symbiont genome were estimated using CheckM v1.0.13 (Parks et al., 2015) under default settings. A manual approach was applied to further refine the genome assembly. Briefly, the best-assembled bins of the two bacteria from each individual were selected to serve as reference genomes. The sequences of the bins were searched reciprocally using BLASTn 2.11.0+ (Camacho et al., 2009) to identify missing genes in the references. The corresponding sequences were merged using Cap3 v10.2011 (Huang & Madan, 1999) under default settings to generate the best genome. The completeness of the final genome was assessed with CheckM v1.0.13 (Parks et al., 2015).

The symbiont genomes were annotated using Prokka v1.13.4 (Seemann, 2014) under default settings. Pfam and COG annotations of the predicted protein sequences were performed using WebMGA (Wu et al., 2011) with default settings, while Gene Ontology (GO) annotations were conducted using eggNOG-MAPPER v5.0 (Huerta-Cepas et al., 2019) against the eggNOG HMM database. Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation was conducted using KofamKOALA (Aramaki et al., 2020), KAAS (Moriya et al., 2007), and BlastKOALA (Kanehisa et al., 2016). KEGG Mapper (Moriya et al., 2007) was used to reconstruct the metabolic pathways.

Phylogenetic analysis

Maximum-likelihood (ML) phylogenetic analysis of the bacterial genomes was performed using IQ-TREE v2.1.4-beta (Kalyaanamoorthy et al., 2017) with the substitution model selected by ModelFinder in IQ-TREE and 1 000 ultrafast bootstraps. For each symbiont genome, 120 marker genes were generated using the "classify" workflow in GTDB-Tk v1.5.1 (Chaumeil et al., 2020) under default settings (data release r202). Distributions of the selected genes and complexes were determined using corresponding references and OrthoFinder v2.2.7 (Emms & Kelly, 2019) under default settings.

Identification of HGT events

To identify HGT events in the epibiont, the predicted protein sequences were searched against the NCBI NR database using BLASTp (Buchfink et al., 2015) implemented in Diamond v0.9.24 using the "more-sensitive" option and E-value threshold of 1e-10, with the associated taxonomy then extracted using a custom script. For each sequence, the horizontal gene transfer (*h*) index was calculated by subtracting the best hit score with the best campylobacterial hit score (Ip et al., 2021). The following stringent criteria were applied to select candidate horizontally transferred genes (HTGs) (Wang & Ruan, 2020): (1) ≥10 hit sequences, with the best hit of non-campylobacterial origin, and ≥70% of non-campylobacterial occupancy among all hits; (2) ≥35% sequence identity between the query protein across its entire length and its best hit; (3) Best hit score ≥100 and *h*≥30; (4) Similar Illumina read coverage of HTGs and their neighboring bona fide campylobacterial genes; (5) HTG candidates with monophyletic relationships with non-campylobacterial genes among the top 30 hits through phylogenetic analysis using the ML method implemented in IQ-TREE v2.1.4-beta (Kalyaanamoorthy et al., 2017), with the substitution model selected by ModelFinder and 1 000 ultrafast bootstraps.

RNA extraction and metatranscriptomic sequencing

Total RNA was extracted from gill and foot tissues of mussels 1–3 using Trizol reagent (Takara, Japan). RNA quality and quantity were examined as described for DNA samples and sent to Novogene (Beijing, China) for metatranscriptomic sequencing. Ribosomal RNA was removed from total RNA using a NEBNext® Ultra™ RNA Library Prep Kit for Illumina (NEB, USA), and RNA molecules were fragmented into 250–300 bp fragments and reverse transcribed into cDNA. The constructed libraries were paired-end sequenced on an Illumina NovaSeq 6000 sequencer (Illumina, USA) to produce 150 bp reads.

***De novo* metatranscriptome assembly, annotation, and analysis**

Adapters and low-quality reads were removed using Trimmomatic v0.38 (Bolger et al., 2014) under the following settings: LEADING=15, TRAILING=15, SLIDINGWINDOW=4:20, MINLEN=40. Clean reads from the gill and foot tissues were pooled and used for *de novo* assembly using Trinity

v2.8.5 (Grabherr et al., 2011) under default settings. TransDecoder v5.5.0 (Haas et al., 2013) was used to predict protein coding genes (PCGs) under default settings. Cd-Hit-Est v4.7 (Fu et al., 2012) was applied to remove redundant contigs using 95% as the sequence similarity cutoff. To remove sequences from bacteria and possible contamination, the PCGs were BLASTx 2.11.0+ (Camacho et al., 2009) searched against a custom database (Supplementary Table S1, including relevant published datasets and data from this study) with an E-value threshold of 1e-10 to generate host transcripts. The longest isoform was selected as the unigene for downstream analysis. To estimate gene expression, clean reads were mapped to the combined transcriptome, containing the host unigenes and two symbiont genes, and expression levels were quantified as transcripts per million (TPM) using Salmon v0.14.1 (Patro et al., 2017). Genes with TPM≥0.5 in at least two of the three samples were considered expressed. Unigenes were retained for downstream analyses. Transcriptome completeness was evaluated using Benchmarking Universal Single-Copy Orthologs (BUSCO) v4.1.1 (Simão et al., 2015) with the metazoa_odb10 database. Functional annotation was performed as applied to the bacterial genomes. In addition, read count data yielded by Salmon were used to determine differentially expressed genes (DEGs) between the gill and foot tissues using DESeq2 v1.30.0 (Love et al., 2014). TBTools v1.082 (Chen et al., 2020) was applied for KEGG enrichment analysis, with genes and KEGG pathways with a false discovery rate (FDR)<0.05 considered significantly enriched.

Protein extraction, liquid chromatography-tandem mass spectrometry (LC-MS/MS), and metaproteomic analysis

Gill tissues from mussels 4–6 were used for proteomic analysis. In brief, 100 mg of tissue was ruptured in Lysing Matrix Z (MPBIO, France) using a bench-top bead beating lysis system (MPBIO, France) in lysis buffer (8 mol/L urea, 40 mmol/L DTT, 0.5% SDS, and protease inhibitor cocktail, pH=8.0), and sonicated after 5 min at 400 W to break nucleic acids using an ultrasonic homogenizer (SCIENTZ, China). The protein solution was purified by centrifugation for 15 min at 12000 r/min and 4 °C, and quantified using the Bradford method (Bradford, 1976). Approximately 40 μg of protein was used in sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on a 4%–12% gradient gel and stained with Coomassie Brilliant Blue to visualize protein bands. Each gel was then cut into eight slices and destained with 100 mmol/L NH_4HCO_3, 50 mmol/L NH_4HCO_3, 50% acetonitrile (ACN), followed by 100% ACN. Each sample was reduced using 10 mmol/L DTT and alkylated using 55 mmol/L iodoacetamide. In-gel digestion was performed using 5 μg/mL MS-grade trypsin (Thermo Scientific, USA) for 14 h at 37 °C. Pierce C18 Tips (Thermo Scientific, USA) were applied for desalting. The resultant peptides were analyzed on a Dionex UltiMate 3000 RSLCnano system (Thermo Scientific, USA) coupled with an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific, USA). The following 120 min mobile phase mixing gradient was applied: initial 5 min from mobile phase A (0.1% formic acid in H_2O) to 5 min of 3% mobile phase B (0.1% formic acid in 80% ACN), 5 min from 3% to 8% B, 60 min from 8% to 25% B, 30 min from 25% to 40% B, 10 min from 40% to 98% B, and 5 min at 98% B. The flow rate was controlled at 300 nL/min. The following settings were applied for Orbitrap analysis: positive ion mode, MS resolution: 60 000, scan range: 300–1 500 m/z, AGC target: standard, dynamic exclusion duration: 60 s, high energy collision dissociation (HCD) in MS/MS, HCD collision energy: 30%, and isolation window: 1.6 m/z.

The LC-MS/MS data were converted into MGF files using ThermoRawFileParser v1.3.2 (Hulstaert et al., 2020) and searched using the MetaPro-IQ pipeline (Zhang et al., 2016) against our custom database (Supplementary Table S1). Search sensitivity and amount of protein identified were optimized by following an iterative pipeline (Zhang et al., 2016). In brief, the whole custom database was searched to generate a "pseudo-metaproteome" database for each sample. A typical target-decoy database search (FDR<0.01) was performed and the resulting proteins for all samples were combined and de-duplicated to generate a combined non-redundant database. The original raw files were searched for protein identification against the non-redundant database using MaxQuant v1.6.17.0 (Tyanova et al., 2016). Maximum missed digestion sites were set to 2 and minimum peptide length was set to 7. An FDR threshold of <0.01 was applied in each replicate for final protein identification. The intensity-based absolute protein quantification (iBAQ) algorithm was used for label-free quantification, including both razor and unique peptides for protein quantification, with a minimum ratio count of 2.

RESULTS

Localization of epibiotic and endosymbiotic bacteria

The H&E staining results showed the ciliary tufts of the gill filaments (Figure 1C). FISH analysis identified a low abundance of Campylobacteria (orange) on the surface of the gill filaments and outside the epithelial cells, but a high abundance of Gammaproteobacteria (green) inside the gill epithelial cells (Figure 1D). Bacterial fluorescence signals were not detected in the asymbiotic ciliary tufts (Figure 1D), consistent with previous results from *G. platifrons* and *B. septemdierum* (Yu et al., 2019). SEM analysis revealed epithelial bacteriocytes in the gill tissue, and spheroidal Gammaproteobacteria were found inside the host bacteriocytes for those cells with ruptured cell walls (Figure 2A, B). TEM analysis clearly revealed the bacteriocyte structures (Figure 2C), each containing many endosymbiotic MOB (Figure 2D). Endosymbionts possess stacked intracellular membranes indicative of type I methanotrophs (Cavanaugh et al., 1987). However, based on the SEM and TEM results, no typical twisted-shaped Campylobacteria were identified on the gill surface or in the bacteriocytes. Nevertheless, considering that the multi-omics data showed the presence of Campylobacteria, and FISH analysis revealed their location on the gill surface, we consider these bacteria to be epibionts of *G. haimaensis*. Similarly, Assié et al. (2016) found Campylobacteria on the gill surfaces of *G. childressi* and *B. azoricus* by FISH as well as filamentary Campylobacteria by SEM and TEM.

Epibiont and endosymbiont genomes

DNA sequencing of the three gill samples produced 41.79, 44.78, and 42.71 million paired-end reads, respectively (Supplementary Table S2). *De novo* genome assembly produced 199 k, 203 k, and 200 k contigs, ranging from 1 kb to 71 kb in length (Supplementary Table S3). Genome binning recovered a campylobacterial genome with 265 contigs measuring 2.12 Mb in total length, and a gammaproteobacterial genome with 714 contigs measuring 3.19 Mb in total length (Table 1). Based on CheckM analysis, the campylobacterial genome showed 94.72% completeness and 2.44% contamination, and the endosymbiont genome showed 97.89% completeness and 1.44% contamination. The two assembled *G. haimaensis* bacterial genomes were of high quality compared to other published bathymodioline gill-associated bacteria based on completeness (53.4% to 97.68%) and contamination (0 to 11.51%) (Table 1). The campylobacterial genome encoded 2 213 PCGs, 95.1% of which were annotated. The gammaproteobacterial genome encoded 3 312 PCGs, 91.8% of which were annotated

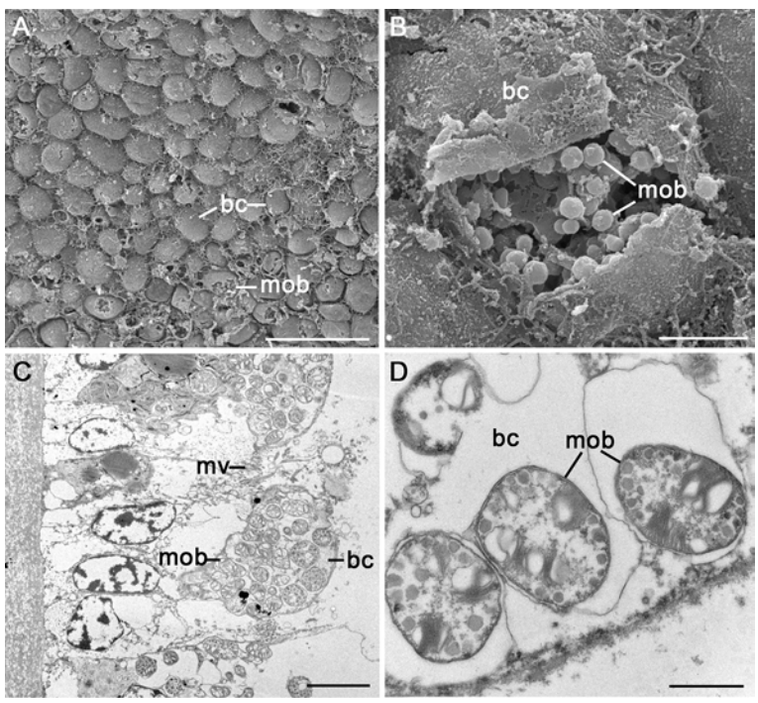

Figure 2 Distribution and morphology of *G. haimaensis* gill endosymbiont

A: SEM of gill filament surface, showing bacteriocytes. Scale bar: 50 μm. B: SEM of a bacteriocyte (close-up), showing endosymbiotic bacteria are coccus in shape and located inside a ruptured bacteriocyte. Scale bar: 5 μm. C: TEM of gill filament, showing distribution of endosymbiont. Scale bar: 5 μm. D: TEM of a bacteriocyte, showing morphological details of endosymbiotic bacteria. Scale bar: 1 μm. bc: bacteriocyte; mob: methane-oxidizing bacteria (gammaproteobacterial endosymbiont); mv: microvilli.

Table 1 General genomic features of bathymodioline mussel gill-associated bacteria

Genome	Size (Mb)	Contigs	Completeness (%)	Contamination (%)	N50 (bp)	GC (%)	CDS	RNA
Sulfur-oxidizing Campylobacteria								
Gigantidas haimaensis epibiont	2.12	265	94.72	2.44	15 458	28.70	2 213	34
Bathymodiolus azoricus epibiont	2.30	523	92.39	6.15	4 446	30.00	2 155	37
G. childressi epibiont	2.20	354	95.73	6.12	6 367	30.00	2 204	31
Sulfur-oxidizing Gammaproteobacteria								
B. thermophilus endosymbiont	2.83	1	97.19	1.99	2 832 685	38.60	2 171	44
B. azoricus endosymbiont	2.13	410	97.68	0	9 070	37.70	2 080	46
B. septemdierum endosymbiont	1.47	1	98.68	0	1 469 434	38.74	1 519	42
Methane-oxidizing Gammaproteobacteria								
G. haimaensis endosymbiont	3.19	714	97.89	1.44	6 829	39.50	3 312	40
G. platifrons endosymbiont	4.02	613	98.16	1.16	12 579	39.80	4 195	56
G. azoricus endosymbiont	1.73	1 094	53.44	0.79	2 380	39.00	2 207	33
B. sp. endosymbiont	1.75	5 123	92.00	11.51	1 457	42.80	5 295	N/A

N/A : Not available. Sources of genomes are shown in Supplementary Table S7.

(Supplementary Table S4). Alignment of the recovered *16S rRNA* gene V3–V4 region revealed 100% identity to the corresponding most abundant Campylobacteria and Gammaproteobacteria operational taxonomic units (OTUs) reported in a previous study that sequenced gill tissue of the same species (Xu et al., 2019). Mapping the clean reads against the genomic contigs of the three samples showed that the epibiont and endosymbiont accounted for 2.34%–8.17% and 91.83%–97.66% of the total reads of the two bacteria, respectively.

To determine their phylogenetic relationships, ML analysis of the bathymodioline gill-associated bacteria with their free-living relatives was conducted based on 120 conserved marker genes (Figure 3A; Supplementary Figure S1). The Campylobacteria and Gammaproteobacteria formed two large clades, with Gammaproteobacteria further divided into a MOB clade and SOB clade. All mussel-associated lineages within the three clades clustered together, distinct from their corresponding free-living relatives. The campylobacterial epibiont of *G. haimaensis* was closely related to that of *G. childressi* and *B. azoricus*, while the endosymbiont of *G. haimaensis* was in the gammaproteobacterial MOB clade,

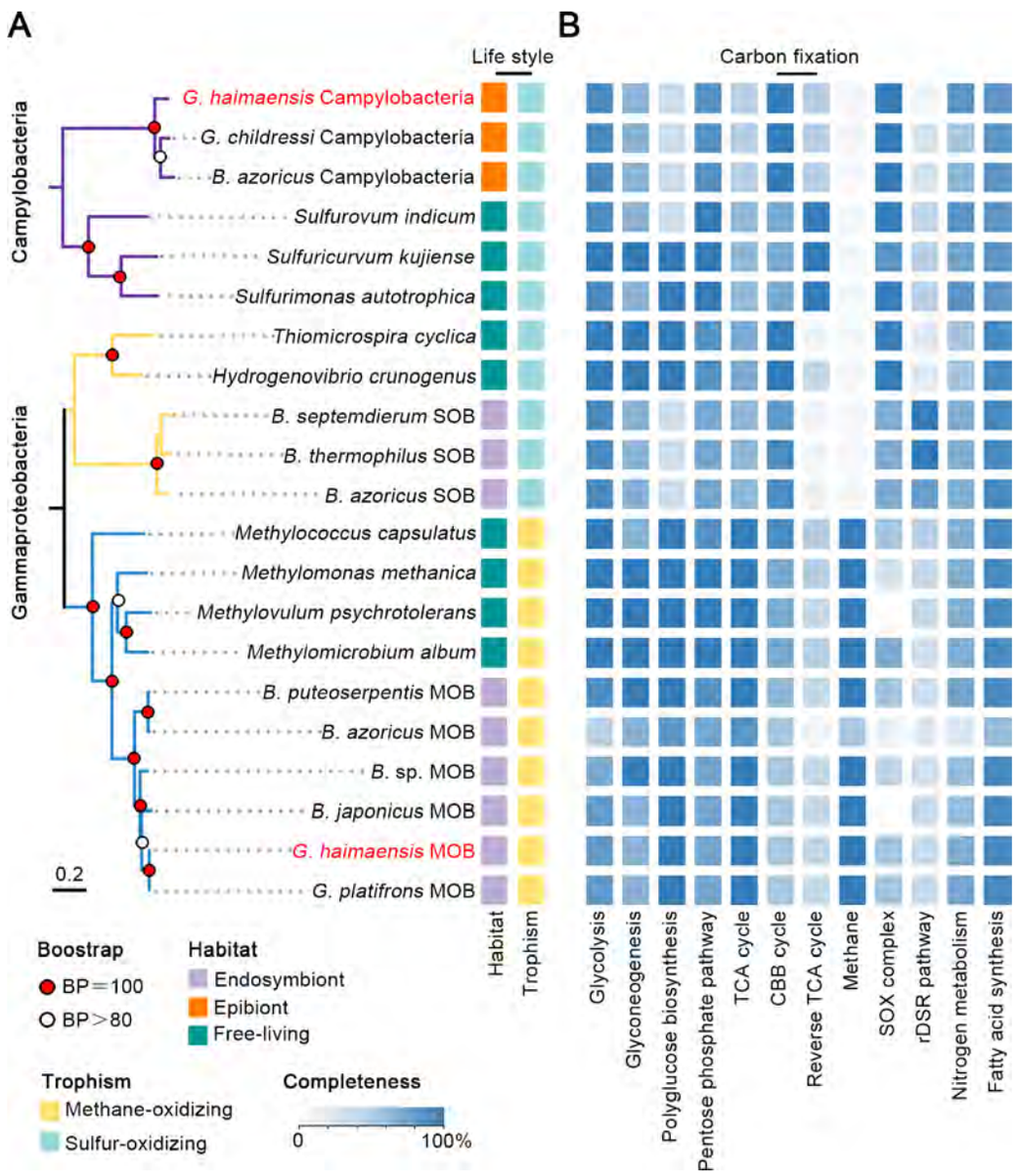

Figure 3 Phylogenetic relationship and central metabolic pathways of *G. haimaensis* gill-associated bacteria and their relatives

A: Simplified ML phylogenetic tree, with outgroup (*Neisseria meningitidis*) omitted and linkage between Campylobacterota and Gammaproteobacteria removed to save space. Full tree with additional genomes is shown in Supplementary Figure S1. Campylobacterial clade is indicated in purple, gammaproteobacterial SOB and MOB clades are labeled in orange and light blue, respectively, and genomes assembled in this study are highlighted in red letters. B: Completeness of central metabolic pathways. Details of metabolic pathways and gene abbreviations are included in Supplementary Table S7.

most closely related to the endosymbiotic MOB of *G. platifrons*.

Metatranscriptome and metaproteome of *G. haimaensis*

RNA sequencing (RNA-seq) of the gill and foot tissues of the three individuals generated 422.3 million reads in total (Supplementary Table S2), which were assembled into 2.14 million contigs (Supplementary Table S3). The *G. haimaensis* host contigs, determined by BLASTn analysis against a custom database (Supplementary Table S1), containing 61 968 unigenes with an N50 of 822 bp. Of these, 29 941 (48.32%) had at least one homolog in public protein databases, and BUSCO analysis showed that they contained 93.0% complete and 1.8% fragmented conserved metazoan genes (Supplementary Table S5). These assembly statistics are comparable to those of other bathymodioline transcriptomes produced using Illumina short reads (Wong et al., 2015; Zheng et al., 2017). Mapping of clean reads against the assembled contigs revealed 75.6% host, 23.5% endosymbiont, and 0.9% epibiont transcripts in the gill tissue, and 96.8% host, 3.2% endosymbiont, and 0.01% epibiont transcripts in the foot tissue (Supplementary Figure S2).

Proteomic analysis of gill tissues detected 26 234 peptides, corresponding to 5 713 proteins, including 4 479 host proteins, 1 171 endosymbiont proteins, and 63 epibiont proteins (Supplementary Table S6). The host, endosymbiont, and epibiont proteins accounted for 77.0%, 22.8%, and 0.18% of the total protein abundance, respectively (Supplementary Figure S2). The total number of proteins detected in this study is comparable to that reported in other studies using similar MS technology (Assié et al., 2020; Ponnudurai et al., 2017a, 2020), and the relative protein abundances of the three parties are consistent with the transcriptomic data. Without determining bacterial enrichment, we only discovered a small number of proteins from the epibiont, likely representing the most abundant proteins in this bacterium.

Different carbon assimilation strategies in two gill-associated bacteria

Carbon fixation by chemoautotrophic bacteria underlies deep-sea vent and seep ecosystems. We found that the epibiont and endosymbiont of *G. haimaensis* possessed different carbon assimilation pathways.

The epibiont genome encoded a complete gene set of enzymes in the CBB cycle, including the large and small subunits of RuBisCO form I (*rbcL* and *rbcS*), phosphoglycerate kinase (*pgk*), glyceraldehyde 3-phosphate dehydrogenase (*gapdh*), transketolase (*tkt*), ribose 5 phosphate isomerase (*rpi*), phosphoribulokinase (*prk*), triosephosphate isomerase (*tpi*), fructose bisphosphate aldolase (*fba*), fructose 1,6-bisphosphatase (*fbp*), and ribulose phosphate 3-epimerase (*rpe*) (Figures 3B, 4; Supplementary Figure S3 and Table S7). RuBisCO form I was transcriptionally active, with *rcbS* and *rbcL* being the 10th and 19th most transcribed genes, respectively. Furthermore, TKT in the CBB cycle was found in the epibiont proteome (Figure 4; Supplementary Figure S4). The epibiont encoded pyruvate: ferredoxin oxidoreductase (*porABCD*) and adenosine triphosphate (ATP)-citrate lyase subunit A (*aclA*) but lacked the five key genes of the rTCA cycle, including the 2-oxoglutarate oxidoreductase (*oorABDG*) and ATP-citrate lyase subunit B genes (*aclB*) (Figure 4; Supplementary Table S7). Consistent with recent research on *G. childressi* and *B. azoricus* (Assié et al., 2020), we found that HGT events were responsible for the acquisition of several key genes in the CBB cycle of the epibiont of *G. haimaensis*. Notably, the *rbcL*, *rbcS* and *gapdh* genes of the *G. haimaensis* epibiont were horizontally transferred from sulfur-oxidizing Gammaproteobacteria (Supplementary Figures S5–S7), while *fba*, *fbp*, *tkt*, and *rpe* were from Betaproteobacteria (Supplementary Figures S8–S11). However, the origin of the *prk* gene was not clear as the bathymodioline epibiont clade was sister to a clade of mixed proteobacteria (Supplementary Figure S12). In contrast, free-living Campylobacteria, including *Sulfurovum indicum*, *Sulfuricurvum kujiense*, and *Sulfurimonas autotrophica*, encoded a complete rTCA cycle (Figure 3), while the CBB cycle was incomplete, missing key genes *rbcL*, *rbcS*, and *prk* (Supplementary Table S7).

Like the closely related bathymodioline methanotrophic symbionts (Figure 3; Supplementary Table S7), the MOB of *G. haimaensis* contained a complete set of genes encoding enzymes for methane oxidation, including particulate methane monooxygenase (*pmoABC*), lanthanide-dependent methanol dehydrogenase (*xoxF*), and formate dehydrogenase (*fdh*), which convert CH_4 to CH_3OH, HCHO, and CO_2 sequentially (Supplementary Figure S3) (DeChaine & Cavanaugh, 2006). Our transcriptomic and proteomic data showed that methane oxidation was highly active in the gill tissue of *G. haimaensis*. The *pmoB*, *pmoC*, *pmoA*, and 5,6,7,8-tetrahydromethanopterin hydro-lyase (*fae*) genes were the 12th, 14th, 16th, and 19th most highly expressed gammaproteobacterial genes, respectively, while PmoB, PmoA, Fae, XoxF, and PmoC ranked 1st, 6th, 7th, 12th, and 15th in terms of gammaproteobacterial protein abundance, respectively (Supplementary Figure S13). Both the CBB and rTCA pathways were incomplete in the MOB of *G. haimaensis*, lacking the *rbcL*, *rbcS*, *prk*, *fba*, and *fbp* genes of the CBB pathway, and the *porBC* and *oorABDG* genes of the rTCA pathway, consistent with the MOB of other bathymodioline mussels (Figures 3, 4; Supplementary Table S7). In contrast, carbon assimilation in the *G. haimaensis* endosymbiont proceeded via the ribulose monophosphate (RuMP) pathway and serine cycle with formaldehyde in the first two steps of methane oxidation (Supplementary Figure S3), as reported in the MOB of *B. azoricus* (Ponnudurai et al., 2017a). Several key genes involved in the RuMP pathway, including hexulose-6-phosphate formaldehyde lyase and 3-hexulose-6-phosphate isomerase, were highly expressed (Figure 4; Supplementary Figure S13A), with the former ranking 2nd in abundance among endosymbiont proteins (Supplementary Figure S13B).

TCA cycle is incomplete in the epibiont but complete in the endosymbiont

The TCA cycle is used to release stored energy and produce precursors of amino acids and reducing agent NADH (Ponnudurai et al., 2017a). Like the epibiont genomes of *G. childressi* and *B. azoricus*, that of *G. haimaensis* encoded an

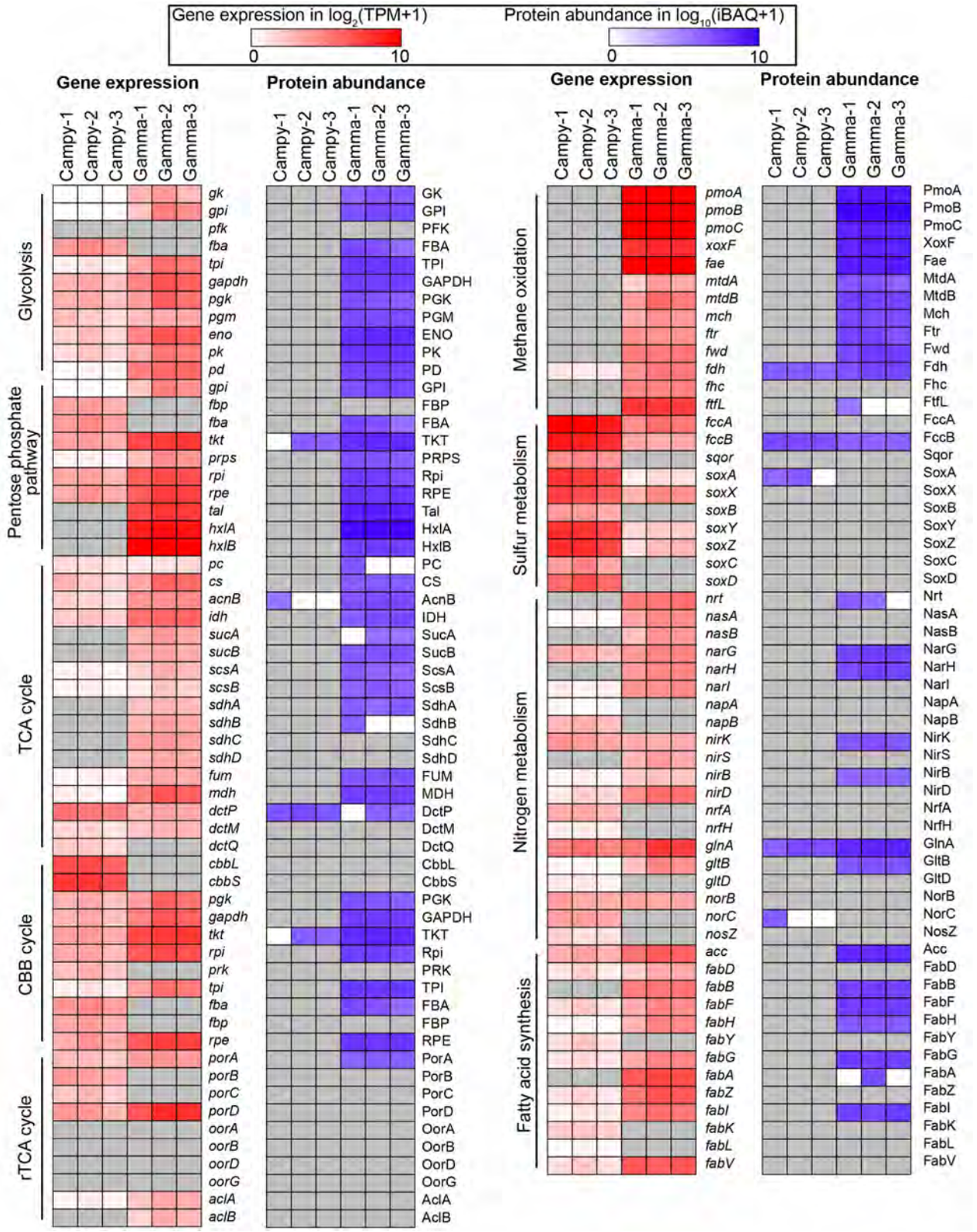

Figure 4 Overview of several central metabolic pathways in *G. haimaensis* campylobacterial epibiont and gammaproteobacterial endosymbiont with gene expression levels and protein abundances

Campy-1 to 3 and Gamma-1 to 3 represent three replicates of campylobacterial epibiont and gammaproteobacterial endosymbiont samples, respectively. Missing expression levels of genes or proteins are shown in gray, and full gene names are included in Supplementary Table S7.

incomplete TCA cycle, with the absence of oxoglutarate dehydrogenase (*sucAB*) and succinate dehydrogenase (*sdhABCD*) genes (Figure 3; Supplementary Table S7). Nevertheless, we found that the *G. haimaensis* epibiont may be capable of importing succinate and fumarate extracellularly from the host to complete its TCA cycle through the TRAP-type C4 dicarboxylate transporter complex (DctQMP) (Supplementary Figure S3), as suggested for the thiotrophic endosymbiont of *B. azoricus* (Ponnudurai et al., 2017a). We identified four copies of each of the three components of DctQMP in the *G. haimaensis* epibiont genome, including the small permease component (*dctQ*), large permease protein (*dctM*), and periplasmic component (*dctP*) (Supplementary Table S8). Among them, *dctP* showed a high transcription level (Figure 4), and its corresponding protein was among the most abundant epibiont proteins (Figure 3; Supplementary Figure S4).

Like free-living MOB, the *G. haimaensis* MOB endosymbiont contained a complete TCA cycle (Figure 3; Supplementary Figure S3 and Table S7). RNA-seq of gill tissues revealed active transcription of all enzyme-encoding genes and most proteins were detected by LC-MS/MS, with malate dehydrogenase being the most highly expressed gene and the corresponding product being the most abundant protein in the TCA cycle (Figure 4).

Active sulfur metabolism and origin of SoxB in the epibiont

Sulfur oxidation is the central energy production pathway for sulfur-oxidizing Campylobacteria and Gammaproteobacteria. The oxidation of thiosulfate to sulfate is catalyzed by the SOX multienzyme complex, which includes L-cysteine S-thiosulfotransferase (*soxAX*), S-sulfosulfanyl-L-cysteine sulfohydrolase (*soxB*), sulfur-oxidizing protein (*soxYZ*), sulfane dehydrogenase (*soxC*), and S-disulfanyl-L-cysteine oxidoreductase (*soxD*) (Figure 3; Supplementary Table S7) (Yamamoto & Takai, 2011). We found hyperactive thiosulfate oxidation in the *G. haimaensis* epibiont (Figure 4), where *soxZ*, *soxY*, *soxA*, *soxX*, and *soxD* were the 6th, 7th, 9th, 16th, and 23rd most highly expressed genes (Supplementary Figure S4A), and SoxA was one of the few proteins from this bacterium detected by LC-MS/MS (Figure 4 ; Supplementary Figure S4B).

Given the importance of the SOX pathway for energy metabolism, we examined the SOX multienzyme complex in the epibiont for evidence of HGT. Among the SOX enzymes, only SoxB, a sulfate thioesterase for sulfone group release and sulfonate moiety hydrolysis (Friedrich et al., 2001), was likely horizontally transferred from thiotrophic Gammaproteobacteria. This was supported by the discrepancy between the phylogenetic tree of the SoxB protein sequences and the species tree (Figure 5; Supplementary Figure S15). In the SoxB phylogenetic tree, the *G. haimaensis*, *G. childressi*, and *B. azoricus* epibiont clade was nested within the free-living and endosymbiotic thiotrophic Gammaproteobacteria (Figure 5; Supplementary Figure S14). Furthermore, this epibiotic campylobacterial SoxB clade was distant from the two other campylobacterial SoxB clades (Figure 5; Supplementary Figure S14).

The reverse dissimilatory sulfite reduction (rDSR) pathway oxidizes sulfide to sulfate anaerobically. Our examination of the epibiont genomes of *G. haimaensis*, *G. childressi*, and *B. azoricus* revealed that they could not oxidize sulfur through the rDSR pathway due to the lack of dissimilatory sulfite reductase (*dsrAB*), which reduces sulfide to sulfite, adenylylsulfate reductase (*aprAB*), which catalyzes sulfite to adenosine-5'-phosphosulfate, and sulfate adenylyltransferase (*sat*), which catalyzes the conversion of adenosine-5'-phosphosulfate to sulfide (Figure 3; Supplementary Table S7).

In addition, we found the genes encoding sulfide dehydrogenase cytochrome subunit (*fccA*) and flavocytochrome c sulfide dehydrogenase (*fccB*) in both the epibiont and endosymbiont of *G. haimaensis*, as well as sulfide: quinone oxidoreductase (*sqor*) in the epibiont (Supplementary Table S7). These enzymes mediate the oxidation of sulfide to elemental sulfur in *G. haimaensis* gill-associated bacteria, which detoxifies hydrogen sulfide and allows the mussels to store sulfur as globules (Sun et al., 2022; Xia et al., 2017). Here, both *fccA* and *fccB* were highly expressed (Figure 4), ranking 2nd and 4th in the epibiont transcriptome, while FccB was the 8th most abundant protein (Supplementary Figure S4), indicating active involvement in sulfur storage, sulfide detoxification, and provision of H^+.

Nutrient biosynthesis in the two gill-associated bacteria

Because nutritional adaptation is the basis of many symbioses, we examined multi-omics data to determine the capacity of the host and two bacteria to synthesize amino acids, vitamins, and cofactors (Table 2; Supplementary Figure S16 and Table S9). The epibiont genome encoded 143 genes for the synthesis of 21 amino acids and five vitamins and cofactors, while the endosymbiont genome encoded 158 genes for the synthesis of 21 amino acids and 13 vitamins and cofactors. In contrast, the host gill transcriptome and proteome involved 51 transcripts and proteins for the synthesis of only nine amino acids and five vitamins and cofactors (Table 2). Remarkably, the GTP cyclohydrolase II gene involved in the initial step of riboflavin synthesis was highly expressed in the epibiont (Supplementary Table S9), ranking 26th in the proteome (Supplementary Figure S4). However, genes encoding enzymes involved in riboflavin synthesis were absent in the host transcriptome and lowly expressed in the endosymbiont genome. In contrast, the synthetic pathways for lysine, glutamine, valine, isoleucine, leucine, histidine, and folate were all active in the endosymbiont (Supplementary Table S9).

Potential transfer of metabolic intermediates from the host to the epibiont

Although the epibiont genome does not encode genes for the synthesis of succinate and fumarate, *DctQMP* genes responsible for C4 dicarboxylate transportation were highly expressed, suggesting that the epibiont may complete the TCA cycle by taking up succinate and fumarate from the host. We found transcripts of *sdhA*, *sdhCD*, and *scsAB*, which are involved in the synthesis of succinate and fumarate, in the host transcriptome (Supplementary Table S10), as well as SdhABC and ScsAB in the gill proteome (Supplementary

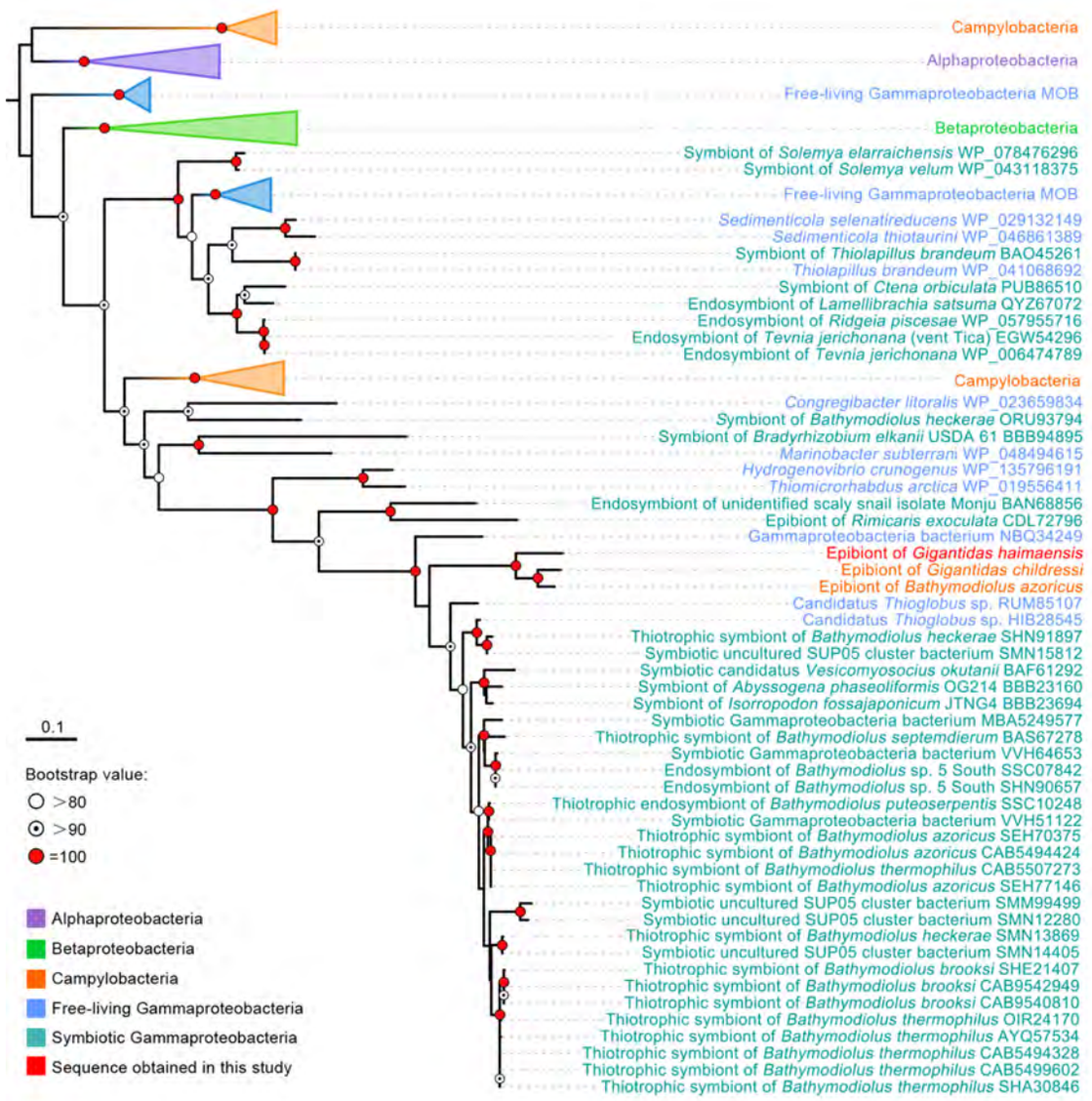

Figure 5 ML phylogenetic tree of several groups of bacteria showing sources of S-sulfosulfanyl-L-cysteine sulfohydrolase (SoxB) amino acid sequences

Tree was constructed under LG+R7 model with 1 000 bootstrap replicates. Several branches with the same class of bacteria are collapsed. Original tree is included in Supplementary Figure S14.

Table S11). Expression of these genes was higher in the gill than in the foot (up to 4.1 times). In the gill transcriptome of *G. haimaensis*, we found active transcription of solute carrier (SLC) family 13 sodium-dependent dicarboxylate transporter (SLC13A2/3/5) (Supplementary Table S10), which is known to transport dicarboxylates (Markovich & Murer, 2004). We also detected its corresponding protein in the gill proteome (Supplementary Table S11), thus supporting succinate and fumarate uptake from the host gill by the epibiont.

Epibiont carbon acquisition from the endosymbiont via host transportation

Carbon fixation in the epibiont CBB cycle requires the use of inorganic carbon (CO_2, CO_3^{2-}, or HCO_3^-), which may be supplied from methane oxidation by the endosymbiont and transportation by the host. Carbonic anhydrases (CAs) interconvert between diffusible CO_2 and ionic HCO_3^- (Supuran, 2018). We identified eight cytoplasmic CA (CCA) and seven membrane-bound CA (MCA) genes in the inorganic carbon transport system of *G. haimaensis* (Supplementary Table S10). Among them, two CCAs were highly expressed (Supplementary Figure S17 and Table S10) and their corresponding proteins were among the most abundant in the gill (Figure 6). We also found 23 and 12 CAs in the deep-sea mussel *G. platifrons* and deep-sea clam *A. marissinica*, respectively (Supplementary Figure S17), and most CAs (13 out of 15) in *G. haimaensis* were closely related to those in *G. platifrons* (Supplementary Figure S17). SLC family 4 bicarbonate transporters (SLC4COs) transport HCO_3^-

Table 2 Amino acid, vitamin, and cofactor biosynthesis capabilities of *G. haimaensis* holobiont inferred from multi-omics analyses

Nutrient	Host	Epibiont	Endosymbiont
Biosynthesis of amino acids			
A, D, G, N, P, Q, R	++	++	++
C, H, HS, I, L, T, V	+	++	++
Cho, E, F, K, Orn, W, Y	−	++	++
S	++	+	++
M	+	+	+
Biosynthesis of vitamins and cofactors			
FAD, CoA, NAD, isopentenyl diphosphate	++	++	++
Glutathione	++	−	++
Protoheme	+	++	++
Pyridoxine, folate, lipoic acid, ubiquinone	+	+	++
Riboflavin	−	++	++
Biotin, thiamine	−	+	++
Siroheme	−	−	++
Pantothenate, pyridoxine phosphate	−	+	+

++: complete gene set; +: incomplete gene set; -: missing gene set. Full gene list and completeness of biosynthesis gene sets are shown in Supplementary Figure S16 and Table S9. A: alanine; C: cysteine; Cho: chorismate; D: aspartate; E: glutamate; F: phenylalanine; G: glycine; H: histidine; HS: homoserine; I: isoleucine; K: lysine; L: leucine; M: methionine; N: asparagine; Orn: ornithine; P: proline; Q: glutamine; R: arginine; S: serine; T: threonine; V: valine; W: tryptophan; Y: tyrosine.

generated by CAs from the inner cellular membrane to the hemolymph in the clam *A. marissinica* (Ip et al., 2021). Here, we identified eight SCL4COs and seven SLC family 26 chloride/bicarbonate transporter (SLC26A6/10) genes in *G. haimaensis* (Supplementary Table S10). Among them, *SCL4COs* and *SLC26A10* were highly expressed (Supplementary Table S10), and their protein products were also identified in the gill proteome (Supplementary Table S11). Furthermore, we found a bicarbonate intake transporter (BicA) and two CCAs in the genome of the epibiont (Supplementary Figure S3), which participate in bicarbonate import and carbon transformation (Kamennaya et al., 2015; Supuran, 2018).

Epibiont thiosulfate acquisition via sulfide detoxification through the host

As stated above, thiosulfate is a major energy source of epibionts and can be obtained from sulfide detoxification through the host. In bivalves, sulfide: quinone reductase (Sqr) and sulfurtransferase (Tst) are proposed to catalyze the oxidation of sulfide to thiosulfate and provide a thiosulfate reservoir for their endosymbionts (Ponnudurai et al., 2020; Sun et al., 2022). We identified three sulfide genes, quinone reductase (*sqr*) and two sulfurtransferase (*tst*), in the host transcriptome of *G. haimaensis*, which were highly expressed in the gill tissue with transcription levels 2.5–5.9- and 2.4–8.5-fold higher than in the foot, respectively (Supplementary Table S10). Corresponding proteins were also identified in the host gill proteome (Supplementary Table S11), highlighting the sulfide detoxification and thiosulfate producing capabilities of the mussel host.

Active metabolism in host gill

We examined the highly expressed transcripts (i.e., top 10%) to gain biological insight into the symbiotic gill transcriptome (Supplementary Figures S18, S19). KEGG pathway annotations indicated that the *G. haimaensis* gill was active in energy, carbohydrate, lipid, amino acid, cofactor, and vitamin metabolism (Supplementary Figure S18). GO enrichment also indicated active metabolism in the gill (Supplementary Figure S19). Based on analysis of DEGs, we identified 1 169 and 856 highly expressed genes in the gill and foot, respectively (Supplementary Figure S20). Most DEGs involved in energy, central carbon, amino acid, cofactor, and vitamin metabolism were more highly expressed in the gill, consistent with the role of this organ in active energy and material metabolism, as well as transportation between mussel gill cells and endosymbiotic bacteria (Figure 7).

Host immune system related to bacteria

The host immune system plays multiple roles associated with bacteria, including invasion, maintenance, and population regulation (Sun et al., 2017, 2021). In the *G. haimaensis* transcriptome, we identified 60 toll-like receptors (TLRs), eight peptidoglycan recognition proteins (PGRPs), 88 C-type lectins (CTLs), and 167 C1q-domain-containing proteins (C1qs) potentially involved in symbiosis (Supplementary Tables S10, S11). We found substantially higher expression of several TLRs (i.e., TLR2, TLR4, TLR6, and TLR13) and PGRPs in the gill compared to the foot (Supplementary Figure S21 and Table S10). We also found highly expressed transcripts of C-type lectins in the gill (Supplementary Figure S18), along with 19 C-type lectins in the gill proteome (Supplementary Table S11). Among the C1qs, several transcripts were extremely highly expressed in the gill compared to the foot (Supplementary Figure S21).

Lysosomal digestion can control bacterial populations and provide nutrients to the host (Sun et al., 2021). In the *G. haimaensis* transcriptome, we identified 132 proteinase genes, of which 26 lysosomal cathepsins (Supplementary Figure S22), showed higher expression in the gill than in the foot (Supplementary Table S10). Among the 14 copies of cathepsin L (*ctsL*), 12 were highly transcribed in the gill, and

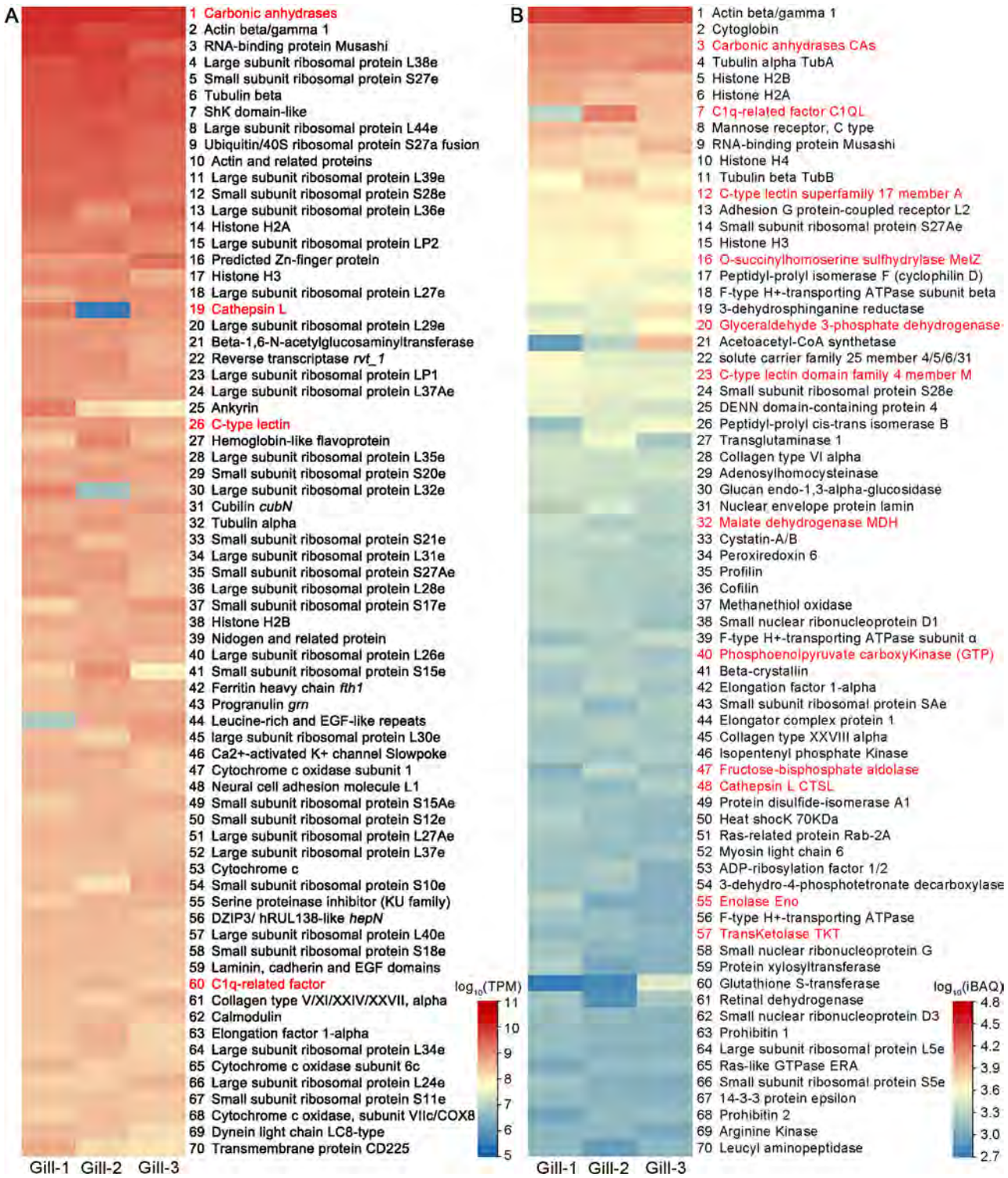

Figure 6 Top 70 highly expressed host genes (A) and abundant proteins (B) of the host in three gill tissue samples

Genes and proteins with no annotation or unknown function are not included. Color gradient represents gene expression levels and protein abundances based on $\log_{10}$-transformed transcripts per million ($\log_{10}$(TPM)) and intensity-based absolute protein quantification ($\log_{10}$(iBAQ)), respectively.

their corresponding proteins were also abundant in the gill proteome, indicating their potential immune function in symbiosis (Figure 6; Supplementary Table S11). Nine other cathepsins, including *ctsABDFOZ*, showed significantly higher expression in the gill than in the foot (Supplementary Table S10).

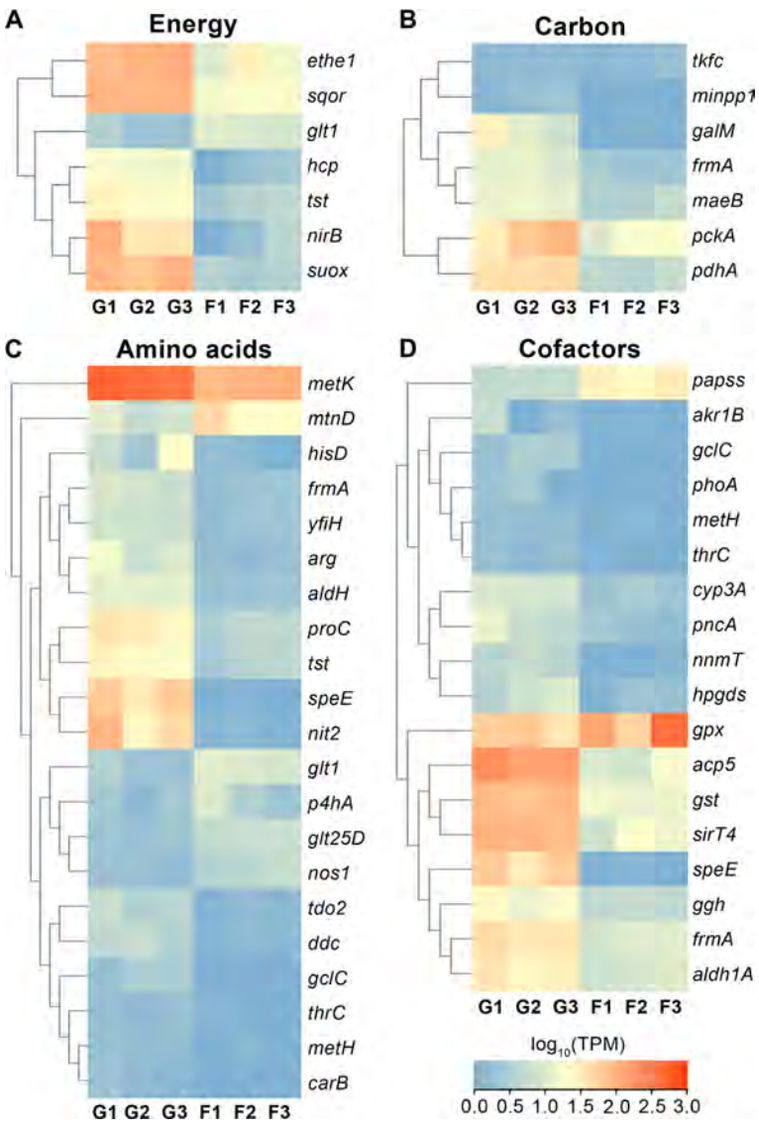

Figure 7 Heatmaps of DEGs between *G. haimaensis* gill and foot involved in metabolism based on KEGG annotation

A: DEGs of energy metabolism. B: DEGs of central carbon metabolism. C: DEGs of amino acid metabolism. D: DEGs of cofactor and vitamin metabolism. Full gene names are available in Supplementary Table S12.

DISCUSSION

Central metabolic pathways of Campylobacteria

Various vent Campylobacteria have been shown to use the rTCA cycle for carbon fixation, including endosymbionts of deep-sea snail *Gigantopelta aegis* (Lan et al., 2021) and epibionts of the yeti crab *Kiwa* sp. (Zwirglmaier et al., 2015) and shrimp *R. exoculata* (Jan et al., 2014). Previous analysis of the epibiont genome of the cold-seep mussel *Gigantidas childressi* only identified the *porAB* and *aclA* genes of the rTCA cycle (Assié et al., 2020). Although we identified orthologs of the *porCD* gene in the *B. azoricus* and *G. haimaensis* epibionts, our results indicated that campylobacterial epibionts rely on the CBB cycle rather than the rTCA cycle for carbon assimilation. Microaerophilic and anaerobic organisms typically employ the rTCA cycle. In contrast, the CBB cycle is more commonly found in aerobic organisms (Hügler & Sievert, 2011). Our phylogenetic analyses corroborate previous studies suggesting that the shift in the carbon assimilation pathway from the rTCA to CBB cycle may have been a key adaptation allowing Campylobacteria to become epibionts in the bathymodioline gills, presumably located inside the well-ventilated mantle cavity (Assié et al., 2020; Cambon-Bonavita et al., 2021).

Both *sucAB* and *sdhCD* are missing in the TCA cycle of free-living Campylobacteria *Sulfurovum* and *Sulfurimonas*, although their function in reversible α-ketoglutarate oxidation to produce succinyl-CoA can be replaced by heterotetrameric *oorABDG*, with ferredoxin or flavodoxin as the electron acceptor (Kendall et al., 2014). These missing genes in the *G. haimaensis* epibiont indicate that it is unable to produce succinyl-CoA. Deletion of the *sucAB* gene is considered an indicator of obligate autotrophy in gammaproteobacterial SOB endosymbionts (Dmytrenko et al., 2014; Newton et al., 2007),

as this can prevent the oxidation of autotrophically fixed organic carbon (Ponnudurai et al., 2017a). Therefore, the absence of *sucAB*, *sdhABCD*, and *oorABDG* in the TCA cycle indicates that the *G. haimaensis* campylobacterial epibiont may be an obligate autotroph.

In the SOX multienzyme complex, SoxAX catalyzes the initial formation linkage of thiosulfate and SoxYZ, then the sulfur substrate, which is covalently linked to the sulfhydryl moiety of SoxYZ, is oxidized by SoxCD, with SoxB responsible for the final stage of thiosulfate oxidation to sulfate (Meyer et al., 2007; Tourova et al., 2013). Previous phylogenetic research has shown that a complete SOX multienzyme complex was present in the common ancestor of Campylobacteria, but this complex has undergone extensive HGT and loss during its evolutionary history (Ghosh et al., 2009). Certain free-living campylobacterial species, such as *Campylobacter* and *Sulfurospirillum*, do not possess *soxB*, and thus may have lost the ability to oxidize sulfur using the SOX multienzyme complex (Supplementary Figure S15). Both *Arcobacter* and bathymodioline epibionts possess *soxB*, but their origins differ, with the former retained from ancestral Campylobacteria and the latter from Gammaproteobacteria. Due to their ability to produce energy via sulfur oxidation, epibiotic Campylobacteria are speculated to overlap with endosymbiotic SOB in ecological niches, which may help explain the absence of the former in bathymodioline mussels that only harbor endosymbiotic SOB (Assié et al., 2016). Our discovery of the acquisition of *soxB* by HGT may support this hypothesis; however, whether the newly acquired epibiont *soxB* performs better on the gill surface than inside the host bacteriocyte warrants further study. The rDSR pathway, which oxidizes sulfide to sulfate anaerobically, is also widely used by endosymbionts of chemosynthetic invertebrates, such as bathymodioline mussels *B. azoricus* and *B. thermophilus* (Ponnudurai et al., 2020), clams *Solemya velum* (Stewart et al., 2011) and *Archivesica marissinica* (Ip et al., 2021), and giant tubeworm *Riftia pachyptila* (Gardebrecht et al., 2012). However, the relative importance of the two sulfur oxidation systems differs among these organisms, with those capable of obtaining sulfide from the sediment more dependent on the rDSR system (Gardebrecht et al., 2012; Stewart et al., 2011), while those mainly utilizing thiosulfate from ambient water more dependent on the SOX system (Ip et al., 2021; Ponnudurai et al., 2020). Bathymodioline mussels usually attach to the surface of authigenic carbonate rocks on the seafloor. The reliance on the SOX multienzyme complex rather than the rDSR system for sulfur oxidation in deep-sea mussel epibionts may indicate that they primarily utilize thiosulfate in the seawater instead of hydrogen sulfide, which is generally more abundant in sediment (Ip et al., 2021).

Host-Campylobacteria interactions

Epibionts may provide mussels with the specific nutrient riboflavin, the precursor of coenzymes flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), which participate in many oxidative and reductive reactions crucial for energy metabolism, cellular respiration, antibody production, growth, and development (Pinto & Zempleni, 2016). Based on the high expression of the riboflavin synthesis enzyme in the epibiont, and the absence of this pathway in the host transcriptome, we speculated that the host may satisfy its demand for riboflavin via importation from the epibiont rather than from the endosymbiont, which shows very low expression of the corresponding genes. However, whether this process is carried out by digestion or transportation remains unclear. Nevertheless, the role of riboflavin in the epibiont warrants further study, as the highly expressed riboflavin genes in the epibiont of *G. haimaensis* (Supplementary Table S9) may be a stress response to the ambient environment outside the gill cells.

Assié et al. (2016) suggested that epibiotic Campylobacteria may have taken over the ecological niche of endosymbiotic sulfur-oxidizing Gammaproteobacteria, forming a reticular bacterial layer to protect mussel gill cells from pathogen affection, thus showing a symbiotic relationship with the host. However, if these low-abundance Campylobacteria are unable to support their hosts with energy and material, they may be considered parasitic epibionts on the *G. haimaensis* gill based on the following evidence.

(1) The epibiont requires metabolic intermediates from the host to complete its TCA cycle. The absence of key genes in the TCA cycle means that the epibiont lacks the ability to synthesize succinate and fumarate. Our finding that succinate and fumarate are actively synthesized by the host gill indicates that these chemicals may be transported from the cytosol to extracellular space and subsequently transported to the epibiont.

(2) The epibiont may acquire additional inorganic carbon from methane oxidation of the endosymbiont via host transportation. Our metatranscriptomic analysis showed that the endosymbiont likely produces considerable CO_2 during methane oxidation, and the mussel gill cells may transport CO_2 inside the bacteriocytes and export it to the extracellular space via CAs and SLC4COs. High expression of *CA* genes has also been reported in the gills of deep-sea bivalves (Hongo et al., 2013; Ponnudurai et al., 2020). Furthermore, in the deep-sea clam *Archivesica marissinica*, MCAs have been proposed to catalyze the conversion of membrane-impermeable HCO_3^- from hemolymph and seawater to permeable CO_2 in both asymbiotic and symbiotic cells (Ip et al., 2021). Inside bacteriocytes, the conversion of HCO_3^- to CO_2 is catalyzed by CCAs (Hongo et al., 2016). In *G. platifrons* and *B. azoricus*, CAs have been suggested to remove CO_2 released by methanotrophic symbionts as an end-product of methane oxidation (Ponnudurai et al., 2020; Wang et al., 2021). Our discovery of highly expressed CCAs in the *G. haimaensis* gill tissue supports their role in concentrating endosymbiont-generated CO_2 during methane oxidation (Supplementary Figure S3). Fixed carbon is subsequently exported by SLC4COs (Ip et al., 2021), and may be assimilated by the epibiont. The bicarbonate outside the gill cells can be further taken up by the epibiont via BicA, and the bicarbonate in the epibiont cytoplasm can be converted to CO_2 again by the CCAs as a source of carbon fixation in the CBB cycle.

(3) The epibiont may utilize thiosulfate generated by the host through sulfide detoxification. Bivalves living in reducing environments need to detoxify sulfide and deliver sulfuric

compounds to their SOB symbionts (Sun et al., 2022). The high transcript levels of *sqr* and *tst*, and the high abundance of their corresponding proteins (Supplementary Table S10), indicated that the gill cells of *G. haimaensis* can efficiently oxidize sulfide for detoxification, and produce thiosulfate as a byproduct. For deep-sea mussels hosting both Campylobacteria and MOB, the positioning of the epibionts on the gill epithelial cells inside the well-ventilated mantle cavity makes it easier for them to gain access to thiosulfate in seawater (Sun et al., 2022). Therefore, the epibionts may be more advantageous than the endosymbiotic bathymodioline SOB, which solely rely on the host to supply thiosulfate. Our discovery of high levels of sulfur-detoxifying transcripts and enzymes in the gills of *G. haimaensis* suggests that the epibiotic Campylobacteria may obtain extra thiosulfate from the mussel for its energy production through the SOX pathway.

Host-Gammaproteobacteria relationship

Consistent with other mussel symbiotic Gammaproteobacteria (Ponnudurai et al., 2017a; Sun et al., 2017), the endosymbiont of *G. haimaensis* adopted highly active methane-oxidation and RuMP as its primary energy production and carbon assimilation pathways. The high endosymbiont abundance suggested that they may provide most of the energy and carbon elements to support host survival and maintenance (Figure 2). In addition, the active lysine, glutamine, valine, isoleucine, leucine, histidine, and folate synthesis pathways in the endosymbionts indicated that the mussel may obtain these amino acids from Gammaproteobacteria, which are unable to be produced by the host itself. These results suggest that the host obtains most of its nutrients from the endosymbiont, and perhaps some essential and less abundant nutrients from the epibiont. These findings are consistent with previous proteomic analysis of the deep-sea mussel *B. azoricus* (Ponnudurai et al., 2017a).

Limited interactions between the epibiont and endosymbiont

In the vent snail *Gigantopelta aegis*, MOB are suggested to supply four-carbon compounds to SOB (Lan et al., 2021), suggesting co-adaptation and evolution of symbiotic partners living in a single cell. A similar situation occurs in *B. azoricus*, where MOB provide SOB with CO_2 derived from methane oxidation (Ponnudurai et al., 2017a). In the *Gigantidas haimaensis* holobiont, however, direct exchange of metabolites between the epibiont and endosymbiont is not possible due to the lack of physical contact. Nevertheless, our analysis of the carbonate transport system in the mussel holobiont suggests that the epibiont likely assimilates some CO_2 generated from methane oxidation of the endosymbiont and transported to the epibiont through the host.

Immune responses to bacteria by the host

Pattern recognition receptors (PRRs) are important for immune recognition and response (Byeon et al., 2015; Graça, 2015). Previous studies have reported high levels of PRRs in the symbiotic gill tissue of bathymodiolines, such as *G. platifrons*, *B. septemdierum*, and *B. azoricus* (Bettencourt et al., 2014; Chen et al., 2021; Ikuta et al., 2019; Li et al., 2020; Wang et al., 2019), suggesting their importance in symbiosis. Both TLRs and PGRPs have been suggested to participate in endosymbiont establishment and symbiosis regulation (Li et al., 2021; Sun et al., 2017). TLRs and PGRPs also function in pathogenic infection, with the latter able to induce phagocytosis (Bunet et al., 2019). Thus, they may also play a role in the recognition and response of epibiotic Campylobacteria. As another PRR, C-type lectins are also reported to be involved in immune recognition and phagocytosis in deep-sea invertebrates (Graça, 2015). For example, Wang et al. (2019) found high expression of C-type lectins in the gills of *G. platifrons*. Liu et al. (2019) considered that the high expression of gill C-type lectins in the vent shrimp *R. exoculata* hosting epibiotic bacteria was related to symbiotic recognition and epibiont attachment. Furthermore, Gourdine & Smith-Ravin (2007) considered that lectins played major roles in symbiont acquisition and maintenance in the clam *Codakia orbicularis* harboring endosymbiotic SOB. Thus, the high expression levels of TLRs, PGRPs, and C-type lectins in the gill of *G. haimaensis* indicate that they may be involved in bacterial recognition, acquisition, and maintenance. However, further studies are required to clarify their functions. C1qs may bind to the surface of pathogens and kill them directly or trigger downstream immune activities (Jiang et al., 2015; Cheng et al., 2019). Consequently, C1qs have been suggested to function as scaffolds of PRRs and in symbiont recognition (Chen et al., 2021). Wang et al. (2021) reported high expression of C1qs in the gill tissue of *G. platifrons*. However, their exact functions in bathymodiolines and whether Campylobacteria and Gammaproteobacteria can be directly killed by C1qs remain unclear. In a previous proteomic study, Ponnudurai et al. (2017a) found high abundance of cathepsin in the gill of *B. azoricus*, suggesting involvement in endosymbiont digestion. High expression of *ctsBL* in the symbiotic trophosome of the tubeworm *Paraescarpia echinospica* has also been implicated in the control of symbiont populations by digestion (Sun et al., 2021). In *G. haimaensis*, CTSB and CTSL may also control endosymbiont populations through direct digestion, as only a few substantive transporters were identified in the endosymbiont genome. These results suggest that the host immune system of *G. haimaensis* may play important roles in bacterial infection, maintenance, and population regulation, which warrant further exploration.

CONCLUSIONS

In this study, we revealed the spatial distribution patterns and relative abundance of the epibiont and endosymbiont of *G. haimaensis* and provided insights into their metabolic functions. We assembled two genomes of these deep-sea mussel gill-associated bacteria. The epibiont, being less abundant than the endosymbiont, adopted the SOX multienzyme complex for energy production, the CBB cycle for carbon assimilation, and an incomplete TCA cycle with input from the host to release stored energy. For the first time, we detected an HGT event in *soxB* in the bathymodioline epibiont. The host may benefit the epibiotic Campylobacteria by supplying metabolic intermediates, CO_2, and thiosulfate.

However, whether the epibiont benefits the host, and is therefore a symbiont, remains unclear. The endosymbiont, which relied on methane oxidation for energy production and the RuMP pathway for carbon assimilation, was the major source of energy and nutrients for itself and the host. The host obtained most of its nutrients from the endosymbiont, and its immune system may play important roles in bacterial infection, maintenance, and population regulation. Overall, our integrative multi-omics study sheds light on the mechanisms of the host-bacterial system that has enabled a group of bathymodioline mussels to successfully colonize extreme deep-sea habitats.

DATA AVAILABILITY

All raw sequencing data for the metagenomes, metatranscriptomes, and symbiont genomes were deposited in the National Center for Biotechnology Information Sequence Read Archive (NCBI: PRJNA785362), National Genomics Data Center (GSA: PRJCA012895), and Science Data Bank (CSTR: 31253.11.sciencedb.06106). The mass spectrometry proteomics data were deposited in the ProteomeXchange (PRIDE identifier: PXD03067). The metagenome assemblies and annotations were deposited in Figshare DOI: 10.6084/m9.figshare.17693405.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

J.W.Q., P.Y.Q., and Y.Z. designed the project. Y.T.L. and T.X. collected the samples. T.X. performed the DNA and RNA extraction. Y.T.L., T.X., J.C.H.I., and Y.S. conducted sequencing, assembly, and annotation of the metagenome and metatranscriptome. Y.T.L. extracted the proteins. L.F. and T.L. conducted the LC-MS/MS. Y.T.L., T.X., and J.C.-H.I. analyzed the proteome. Y.T.L., T.X., J.C.H.I., Y.Z., and J.W.Q. drafted the manuscript. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank the captain and crew of the R/V *Haiyang 6* and the operation team of the *Haima 2* ROV for their professional assistance during the cruise.

REFERENCES

Ansorge R, Romano S, Sayavedra L, Porras MÁG, Kupczok A, Tegetmeyer HE, et al. 2019. Functional diversity enables multiple symbiont strains to coexist in deep-sea mussels. *Nature Microbiology*, **4**(12): 2487−2497.

Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*, **36**(7): 2251−2252.

Assié A, Borowski C, Van Der Heijden K, Raggi L, Geier B, Leisch N, et al. 2016. A specific and widespread association between deep-sea *Bathymodiolus* mussels and a novel family of Epsilonproteobacteria. *Environmental Microbiology Reports*, **8**(5): 805−813.

Assié A, Leisch N, Meier DV, Gruber-Vodicka H, Tegetmeyer HE, Meyerdierks A, et al. 2020. Horizontal acquisition of a patchwork Calvin cycle by symbiotic and free-living Campylobacterota (formerly Epsilonproteobacteria). *The ISME Journal*, **14**(1): 104−122.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, **19**(5): 455−477.

Bettencourt R, Rodrigues M, Barros I, Cerqueira T, Freitas C, Costa V, et al. 2014. Site-related differences in gene expression and bacterial densities in the mussel *Bathymodiolus azoricus* from the Menez Gwen and Lucky Strike deep-sea hydrothermal vent sites. *Fish & Shellfish Immunology*, **39**(2): 343−353.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**(15): 2114−2120.

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, **12**(1): 59−60.

Bunet R, Prévot JM, Vicente N, García-March JR, Martinović R, Tena-Medialdea J, et al. 2019. Genome description and inventory of immune-related genes of the endangered pen shell *Pinna nobilis*: a giant bivalve experiencing a mass mortality event. *Research Square*: 1−40, doi: 10.21203/rs.2.15332/v1.

Byeon JH, Seo ES, Lee JB, Lee MJ, Kim JK, Yoo JW, et al. 2015. A specific cathepsin-L-like protease purified from an insect midgut shows antibacterial activity against gut symbiotic bacteria. *Developmental & Comparative Immunology*, **53**(1): 79−84.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72**: 248−254.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, **10**: 421.

Cambon-Bonavita MA, Aubé J, Cueff-Gauchard V, Reveillaud J. 2021. Niche partitioning in the *Rimicaris exoculata* holobiont: the case of the first symbiotic *Zetaproteobacteria*. *Microbiome*, **9**(1): 87.

Campbell BJ, Engel AS, Porter ML, Takai K. 2006. The versatile ε-proteobacteria: key players in sulphidic habitats. *Nature Reviews Microbiology*, **4**(6): 458−468.

Cavanaugh CM, Levering PR, Maki JS, Mitchell R, Lidstrom ME. 1987. Symbiosis of methylotrophic bacteria and deep-sea mussels. *Nature*, **325**(6102): 346−348.

Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB–Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*, **36**(6): 1925−1927.

Chen CJ, Chen H, Zhang Y, Thomas HR, Frank MH, He YH, et al. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, **13**(8): 1194−1202.

Chen H, Wang MX, Zhang H, Wang H, Zhou L, Zhong ZS, et al. 2021. MicroRNAs facilitate comprehensive responses of Bathymodiolinae mussel against symbiotic and nonsymbiotic bacteria stimulation. *Fish & Shellfish Immunology*, **119**: 420−431.

Cheng J, Hui M, Sha ZL. 2019. Transcriptomic analysis reveals insights into deep-sea adaptations of the dominant species, *Shinkaia crosnieri* (Crustacea: Decapoda: Anomura), inhabiting both hydrothermal vents and cold seeps. *BMC Genomics*, **20**(1): 388.

Childress JJ, Fisher CR, Brooks JM, Kennicutt MC, Bidigare R, Anderson AE. 1986. A methanotrophic marine molluscan (Bivalvia, Mytilidae) symbiosis: mussels fueled by gas. *Science*, **233**(4770): 1306−1308.

DeChaine EG, Cavanaugh CM. 2006. Symbioses of methanotrophs and deep-sea mussels (Mytilidae: Bathymodiolinae). *In*: Overmann J. Molecular Basis of Symbiosis. Berlin: Springer, 227–249.

Dmytrenko O, Russell SL, Loo WT, Fontanez KM, Liao L, Roeselers G, et al. 2014. The genome of the intracellular bacterium of the coastal bivalve, *Solemya velum*: a blueprint for thriving in and out of symbiosis. *BMC Genomics*, **15**(1): 924.

Dubilier N, Bergin C, Lott C. 2008. Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nature Reviews Microbiology*, **6**(10): 725−740.

Duperron S, Halary S, Lorion J, Sibuet M, Gaill F. 2008. Unexpected co-occurrence of six bacterial symbionts in the gills of the cold seep mussel *Idas* sp. (Bivalvia: Mytilidae). *Environmental Microbiology*, **10**(2): 433−445.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, **20**(1): 238.

Friedrich CG, Rother D, Bardischewsky F, Quentmeier A, Fischer J. 2001. Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism?. *Applied and Environmental Microbiology*, **67**(7): 2873−2882.

Fu LM, Niu BF, Zhu ZW, Wu ST, Li WZ. 2012. CD-HIT: accelerated for clustering the next–generation sequencing data. *Bioinformatics*, **28**(23): 3150−3152.

Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, et al. 2012. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. *The ISME Journal*, **6**(4): 766−776.

Ghosh W, Mallick S, DasGupta SK. 2009. Origin of the Sox multienzyme complex system in ancient thermophilic bacteria and coevolution of its constituent proteins. *Research in Microbiology*, **160**(6): 409−420.

Goffredi SK. 2010. Indigenous ectosymbiotic bacteria associated with diverse hydrothermal vent invertebrates. *Environmental Microbiology Reports*, **2**(4): 479−488.

Gourdine JP, Smith-Ravin EJ. 2007. Analysis of a cDNA-derived sequence of a novel mannose-binding lectin, codakine, from the tropical clam *Codakia orbicularis*. *Fish & Shellfish Immunology*, **22**(5): 498–509.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**(7): 644−652.

Graça ELM. 2015. Immune responses in *Bathymodiolus azoricus* upon *Vibrio* challenges: approaches to characterize mussel survival strategies and physiological adaptations in deep sea extreme environments. PhD thesis, University of the Azores.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8**(8): 1494–1512.

Hongo Y, Ikuta T, Takaki Y, Shimamura S, Shigenobu S, Maruyama T, et al. 2016. Expression of genes involved in the uptake of inorganic carbon in the gill of a deep-sea vesicomyid clam harboring intracellular thioautotrophic bacteria. *Gene*, **585**(2): 228−240.

Hongo Y, Nakamura Y, Shimamura S, Takaki Y, Uematsu K, Toyofuku T, et al. 2013. Exclusive localization of carbonic anhydrase in bacteriocytes of the deep-sea clam *Calyptogena okutanii* with thioautotrophic symbiotic bacteria. *Journal of Experimental Biology*, **216**(Pt 23): 4403–4414.

Hou JL, Sievert SM, Wang YZ, Seewald JS, Natarajan VP, Wang FP, et al. 2020. Microbial succession during the transition from active to inactive stages of deep-sea hydrothermal vent sulfide chimneys. *Microbiome*, **8**(1): 102.

Huang XQ, Madan A. 1999. CAP3: A DNA sequence assembly program. *Genome Research*, **9**(9): 868−877.

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. 2019. EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, **47**(D1): D309−D314.

Hügler M, Sievert SM. 2011. Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Annual Review of Marine Science*, **3**: 261−289.

Hulstaert N, Shofstahl J, Sachsenberg T, Walzer M, Barsnes H, Martens L, et al. 2020. ThermoRawFileParser: modular, scalable, and cross-platform RAW file conversion. *Journal of Proteome Research*, **19**(1): 537−542.

Ikuta T, Tame A, Saito M, Aoki Y, Nagai Y, Sugimura M, et al. 2019. Identification of cells expressing two peptidoglycan recognition proteins in the gill of the vent mussel. *Bathymodiolus septemdierum. Fish & Shellfish Immunology*, **93**: 815−822.

Ip JCH, Xu T, Sun J, Li RS, Chen C, Lan Y, et al. 2021. Host-endosymbiont genome integration in a deep-sea chemosymbiotic clam. *Molecular Biology and Evolution*, **38**(2): 502−518.

Jan C, Petersen JM, Werner J, Teeling H, Huang SX, Glöckner FO, et al. 2014. The gill chamber epibiosis of deep-sea shrimp *Rimicaris exoculata*: an in-depth metagenomic investigation and discovery of Zetaproteobacteria. *Environmental Microbiology*, **16**(9): 2723−2738.

Jiang S, Li H, Zhang DX, Zhang H, Wang LL, Sun JS, et al. 2015. A C1q domain containing protein from *Crassostrea gigas* serves as pattern recognition receptor and opsonin with high binding affinity to LPS. *Fish & Shellfish Immunology*, **45**(2): 583−591.

Jing HM, Wang RN, Jiang QY, Zhang Y, Peng XT. 2020. Anaerobic methane oxidation coupled to denitrification is an important potential methane sink in deep-sea cold seeps. *Science of The Total Environment*, **748**: 142459.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**(6): 587−589.

Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, **428**(4): 726−731.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**(4): 772−780.

Kamennaya NA, Ahn S, Park H, Bartal R, Sasaki KA, Holman HY, et al. 2015. Installing extra bicarbonate transporters in the cyanobacterium *Synechocystis* sp. PCC6803 enhances biomass production. *Metabolic Engineering*, **29**: 76−85.

Kendall JJ, Barrero-Tobon AM, Hendrixson DR, Kelly DJ. 2014. Hemerythrins in the microaerophilic bacterium *Campylobacter jejuni* help protect key iron–sulphur cluster enzymes from oxidative damage. *Environmental Microbiology*, **16**(4): 1105−1121.

Laming SR, Gaudron SM, Duperron S. 2018. Lifecycle ecology of deep-sea chemosymbiotic mussels: a review. *Frontiers in Marine Science*, **5**: 282.

Lan Y, Sun J, Chen C, Sun YN, Zhou YD, Yang Y, et al. 2021. Hologenome analysis reveals dual symbiosis in the deep-sea hydrothermal vent snail

Gigantopelta aegis. *Nature Communications*, **12**(1): 1165.
Le Bris N, Arnaud-Haond S, Beaulieu S, Cordes E, Hilario A, Rogers A, et al. 2016. Hydrothermal vents and cold seeps. *In*: United Nations. The First Global Integrated Marine Assessment: World Ocean Assessment I. Cambridge: Cambridge University Press, 853–862.
Levin LA. 2005. Ecology of cold seep sediments: interactions of fauna with flow, chemistry and microbes. *In*: Gibson RN, Atkinson RJA, Gordon JDM. Oceanography and Marine Biology. Boca Raton: CRC Press, 1–46.
Li MN, Chen H, Wang MX, Zhong ZS, Wang H, Zhou L, et al. 2021. A Toll-like receptor identified in *Gigantidas platifrons* and its potential role in the immune recognition of endosymbiotic methane oxidation bacteria. *PeerJ*, **9**: e11282.
Li MN, Chen H, Wang MX, Zhong ZS, Zhou L, Li CL. 2020. Identification and characterization of endosymbiosis-related immune genes in deep-sea mussels *Gigantidas platifrons*. *Journal of Oceanology and Limnology*, **38**(4): 1292–1303.
Lim SJ, Davis BG, Gill DE, Walton J, Nachman E, Engel AS, et al. 2019. Taxonomic and functional heterogeneity of the gill microbiome in a symbiotic coastal mangrove lucinid species. *The ISME Journal*, **13**(4): 902−920.
Lin GM, Lu JG, Sun ZL, Xie JG, Huang JR, Su M, et al. 2021. Characterization of tissue-associated bacterial community of two Bathymodiolus species from the adjacent cold seep and hydrothermal vent environments. *Science of the Total Environment*, **796**: 149046.
Lin XJ, Wakeham SG, Putnam IF, Astor YM, Scranton MI, Chistoserdov AY, et al. 2006. Comparison of vertical distributions of prokaryotic assemblages in the anoxic Cariaco Basin and Black Sea by use of fluorescence in situ hybridization. *Applied and Environmental Microbiology*, **72**(4): 2679−2690.
Liu XL, Ye S, Cheng CY, Li HW, Lu B, Yang WJ, et al. 2019. Identification and characterization of a symbiotic agglutination-related C-type lectin from the hydrothermal vent shrimp *Rimicaris exoculata*. *Fish & Shellfish Immunology*, **92**: 1–10.
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**(12): 550.
Markovich D, Murer H. 2004. The SLC13 gene family of sodium sulphate/carboxylate cotransporters. *Pflügers Archiv*, **447**(5): 594−602.
Meyer B, Imhoff JF, Kuever J. 2007. Molecular analysis of the distribution and phylogeny of the *soxB* gene among sulfur-oxidizing bacteria - evolution of the Sox sulfur oxidation enzyme system. *Environmental Microbiology*, **9**(12): 2957−2977.
Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, **35**(S2): W182−W185.
Newton ILG, Woyke T, Auchtung TA, Dilly GF, Dutton RJ, Fisher MC, et al. 2007. The *Calyptogena magnifica* chemoautotrophic symbiont genome. *Science*, **315**(5814): 998−1000.
Osman EO, Weinnig AM. 2022. Microbiomes and obligate symbiosis of deep-sea animals. *Annual Review of Animal Biosciences*, **10**: 151−176.
Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, **25**(7): 1043−1055.
Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, **14**(4): 417−419.
Pinto JT, Zempleni J. 2016. Riboflavin. *Advances in Nutrition*, **7**(5): 973−975.
Ponnudurai R, Heiden SE, Sayavedra L, Hinzke T, Kleiner M, Hentschker C, et al. 2020. Comparative proteomics of related symbiotic mussel species reveals high variability of host-symbiont interactions. *The ISME Journal*, **14**(2): 649−656.
Ponnudurai R, Kleiner M, Sayavedra L, Petersen JM, Moche M, Otto A, et al. 2017a. Metabolic and physiological interdependencies in the *Bathymodiolus azoricus* symbiosis. *The ISME Journal*, **11**(2): 463−477.
Ponnudurai R, Sayavedra L, Kleiner M, Heiden SE, Thürmer A, Felbeck H, et al. 2017b. Genome sequence of the sulfur-oxidizing *Bathymodiolus thermophilus* gill endosymbiont. *Standards in Genomic Sciences*, **12**(1): 50.
Ponsard J, Cambon-Bonavita MA, Zbinden M, Lepoint G, Joassin A, Corbari L, et al. 2013. Inorganic carbon fixation by chemosynthetic ectosymbionts and nutritional transfers to the hydrothermal vent host-shrimp *Rimicaris exoculata*. *The ISME Journal*, **7**(1): 96–109.
Roeselers G, Newton ILG. 2012. On the evolutionary ecology of symbioses between chemosynthetic bacteria and bivalves. *Applied Microbiology and Biotechnology*, **94**(1): 1−10.
Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, **30**(14): 2068−2069.
Sen A, Åström EKL, Hong WL, Portnov A, Waage M, Serov P, et al. 2018. Geophysical and geochemical controls on the megafaunal community of a high Arctic cold seep. *Biogeosciences*, **15**(14): 4533−4559.
Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, **31**(19): 3210−3212.
Stewart CN Jr, Via LE. 1993. A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. *Biotechniques*, **14**(5): 748−750.
Stewart FJ, Dmytrenko O, Delong EF, Cavanaugh CM. 2011. Metatranscriptomic analysis of sulfur oxidation genes in the endosymbiont of *Solemya velum*. *Frontiers in Microbiology*, **2**: 134.
Sun J, Zhang Y, Xu T, Zhang Y, Mu HW, Zhang YJ, et al. 2017. Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nature Ecology & Evolution*, **1**(5): 0121.
Sun Y, Wang MX, Zhong ZS, Chen H, Wang H, Zhou L, et al. 2022. Adaption to hydrogen sulfide-rich environments: Strategies for active detoxification in deep-sea symbiotic mussels. *Gigantidas platifrons. Science of the Total Environment*, **804**: 150054.
Sun YN, Sun J, Yang Y, Lan Y, Ip JCH, Wong WC, et al. 2021. Genomic signatures supporting the symbiosis and formation of chitinous tube in the deep-sea tubeworm *Paraescarpia echinospica*. *Molecular Biology and Evolution*, **38**(10): 4116–4134.
Supuran CT. 2018. Carbonic anhydrase activators. *Future Medicinal Chemistry*, **10**(5): 561−573.
Tourova TP, Slobodova NV, Bumazhkin BK, Kolganova TV, Muyzer G, Sorokin DY. 2013. Analysis of community composition of sulfur-oxidizing bacteria in hypersaline and soda lakes using *sox*B as a functional molecular marker. *FEMS Microbiology Ecology*, **84**(2): 280−289.
Tunnicliffe V. 1991. The biology of hydrothermal vents: ecology and evolution. *Oceanography and Marine Biology*, **29**: 319−407.
Tyanova S, Temu T, Cox J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols*, **11**(12): 2301−2319.
Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP-a flexible pipeline for

genome–resolved metagenomic data analysis. *Microbiome*, **6**(1): 158.

Van Dover CL. 2000. The Ecology of Deep-Sea Hydrothermal Vents. Princeton: Princeton University Press, 424.

Waite DW, Vanwonterghem I, Rinke C, Parks DH, Zhang Y, Takai K, et al. 2018. Addendum: comparative genomic analysis of the class *Epsilonproteobacteria* and proposed reclassification to Epsilonbacteraeota (phyl. nov.). *Frontiers in Microbiology*, **9**: 772.

Wang H, Zhang H, Wang MX, Chen H, Lian C, Li CL. 2019. Comparative transcriptomic analysis illuminates the host-symbiont interactions in the deep-sea mussel *Bathymodiolus platifrons*. *Deep Sea Research Part I: Oceanographic Research Papers*, **151**: 103082.

Wang H, Zhang H, Zhong ZS, Sun Y, Wang MX, Chen H, et al. 2021. Molecular analyses of the gill symbiosis of the bathymodiolin mussel *Gigantidas platifrons*. *iScience*, **24**(1): 101894.

Wang MS, Ruan RX. 2020. Genome-wide identification and functional analysis of the horizontally transferred genes in *Penicillium*. *Genomics*, **112**(6): 5037–5043.

Wong YH, Sun J, He LS, Chen LG, Qiu JW, Qian PY. 2015. High-throughput transcriptome sequencing of the cold seep mussel *Bathymodiolus platifrons*. *Scientific Reports*, **5**: 16597.

Wu ST, Zhu ZW, Fu LM, Niu BF, Li WZ. 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics*, **12**: 444.

Xia YZ, Lü CJ, Hou NK, Xin YF, Liu JH, Liu HL, et al. 2017. Sulfide production and oxidation by heterotrophic bacteria under aerobic conditions. *The ISME Journal*, **11**(12): 2754−2766.

Xu T, Feng D, Tao J, Qiu JW. 2019. A new species of deep-sea mussel (Bivalvia: Mytilidae: *Gigantidas*) from the South China Sea: Morphology, phylogenetic position, and gill-associated microbes. *Deep Sea Research Part I:Oceanographic Research Papers*, **146**: 79−90.

Yamamoto M, Takai K. 2011. Sulfur metabolisms in epsilon- and gamma-*Proteobacteria* in deep-sea hydrothermal fields. *Frontiers in Microbiology*, **2**: 192.

Yu JJ, Wang MX, Liu BZ, Yue X, Li CL. 2019. Gill symbionts of the cold-seep mussel *Bathymodiolus platifrons*: composition, environmental dependency and immune control. *Fish & Shellfish Immunology*, **86**: 246−252.

Zhang X, Ning ZB, Mayne J, Moore JI, Li J, Butcher J, et al. 2016. MetaPro-IQ: a universal metaproteomic approach to studying human and mouse gut microbiota. *Microbiome*, **4**(1): 31.

Zheng P, Wang MX, Li CL, Sun XQ, Wang XC, Sun Y, et al. 2017. Insights into deep-sea adaptations and host-symbiont interactions: A comparative transcriptome study on *Bathymodiolus* mussels and their coastal relatives. *Molecular Ecology*, **26**(19): 5133−5148.

Zwirglmaier K, Reid WDK, Heywood J, Sweeting CJ, Wigham BD, Polunin NVC, et al. 2015. Linking regional variation of epibiotic bacterial diversity and trophic ecology in a new species of Kiwaidae (Decapoda, Anomura) from East Scotia Ridge (Antarctica) hydrothermal vents. *MicrobiologyOpen*, **4**(1): 136−150.