

Supplementary Materials for

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 Fang⁴, Tiangang Luan^{5,6}, Yu Zhang^{7,*}, Pei-Yuan Qian^{2,3,*}, Jian-Wen Qiu^{1,2,*}

¹Department of Biology, Hong Kong Baptist University, Hong Kong SAR, China

²Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, Guangdong 511458, China

³Department of Ocean Science, Hong Kong University of Science and Technology, Hong Kong SAR, China

⁴Instrumental Analysis & Research Center, Sun Yat-Sen University, Guangzhou, Guangdong 510875, China

⁵State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou, Guangdong 510875, China

⁶Institute of Environmental and Ecological Engineering, Guangdong University of Technology, Guangzhou, Guangdong 510006, China

⁷College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, Guangdong 518060, China

#Authors contributed equally to this work

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

DOI: 10.24272/j.issn.2095-8137.2022.279

Supplementary Results and Discussion

Carbon metabolism

The genomes of the epibiont and endosymbiont encode genes for glycolysis, glyconeogenesis and non-oxidative branch of the pentose phosphate pathway (Fig. 3B & Table S4). Like the epibionts of *Bathymodiolus azoricus* and *Gigantidas childressi*, the 6-phosphofructokinase (*pfk*) and transaldolase (*tal*) are missing in the epibiont of *G. haimaensis*, indicating its potential deficiency in using sucrose and fructose (Assié et al., 2020), and the conversion between D-erythrose-4P and D-sedo-heptulose-7P. However, the epibiont encodes *pfk*, fructose-1,6-bisphosphatase (*fbp*) and fructose-bisphosphate aldolase (*fba*), which allows the production of 5-phospho-alpha-D-ribose 1-diphosphate (PRPP). The missing of *fbp* and *fba* in the endosymbiont indicates its deficiency of PRPP synthesis. The missing of phosphoenolpyruvate synthase genes (*ppsA*) and phosphoenolpyruvate carboxykinase (*pck*) in the genomes of the epibiont and endosymbiont indicates that they lack the ability to convert pyruvate and oxaloacetate to phosphoenolpyruvate. Besides, most genes involved in polyglucose biosynthesis are missing in the epibionts, but the complete pathway is present in the endosymbiont (Fig. 3B). The difference of carbon metabolism may imply the tripartite interaction among host and two symbionts to confirm the needs in carbon metabolism.

Hydrogen oxidation

Some deep-sea chemosynthetic symbionts, such as the thiotroph endosymbiont and campylobacterial epibiont of vent mussel *B. azoricus* (Petersen et al., 2011; Assié et al., 2020) are known to encode genes as potential auxiliary electron donors from hydrogen. However, the epibiont and endosymbiont of *G. haimaensis* do not encode hydrogenase genes, indicating their lack of the ability to use hydrogen as an energy source. This result is consistent with the absent of hydrogenase gene in the epibiont of *G. childressi* (Assié et al., 2020). Since both *G. childressi* and *G. haimaensis* inhabit cold seeps, our results appear to support the hypothesis that hydrogenases are unique to bacterial symbionts of hydrothermal vent mussels (Petersen et al., 2011).

Nitrogen assimilation in the two symbionts

Exogenous nitrogen, typically available in cold seeps in the form of nitrate, is crucial for deep-sea mussels (Petersen et al., 2011). A previous study found nitrate reductase activity for ammonium assimilation in the gill tissue of symbiotic mussels (Lee & Childress, 1996). Assimilatory nitrate reduction (ANR) generally uses cytoplasmic nitrate reductase (NasAB), while dissimilatory nitrate reduction (DNR) usually applies membrane-bound respiratory nitrate reductase (NarGHI) or the periplasmic nitrate reductase enzyme complex (NapAB) for the reducing nitrate to nitrite (Potter et al., 2001). We found both ANR and DNR genes in the genome of the *G. haimaensis* epibiont, among them *narG* and *napB* were actively expressed (Fig. 4, Table S7). Following nitrate reduction to nitrite, the product may be further reduced to ammonia via DNR by a NADH form nitrite reductase (NirBD) and a cytochrome c nitrite reductase (NrfAH) (Yuan & Wang, 2021), which were also identified in the *G. haimaensis* epibiont (Fig. 4). The resultant ammonium can then be used for the glutamine and glutamate synthesis (Reitzer, 2003). In addition, we found evidence of respiratory denitrification, through which nitrite is converted to nitrogen by nitric oxide reductase (NorBC) and nitrous oxide reductase (NosZ) (Fig. 4, Table S7). Among these enzymes, glutamine synthetase and NorC were highly abundant in the epibiont proteome (Fig. 4 & S4). Unexpectedly, nitrate/nitrite transporter (*nrt*) is missing in the campylobacterial epibiont, indicating that either there is another source of nitrate, or their nitrogen assimilatory primary depends on ammonia transportation from the fluid via ammonia permease (Table S8). The epibionts of *G. childressi* and *B. azoricus* have been suggested to incorporate ammonia via high-affinity ammonium uptake transporters (Assié et al., 2020). Consistent with this, we found a transcriptionally active ammonia permease in the epibiont of *G. haimaensis*, while nitrate reduction genes were all lowly expressed (Fig. 4), indicating the potential role of the ammonium transporter in its nitrogen assimilation. In contrast, the endosymbiont of *G. haimaensis* encodes a highly expressed ammonia permease and Nrt for nitrate intake (Fig. 4, Table S7), and the *narGHI* was expressed at a high level than *nasAB*, and NarGH proteins were abundant (Fig. 4). These results imply that the endosymbiotic MOB of *G. haimaensis* primary relies on DNR for nitrate reduction. Besides, respiratory denitrification is incomplete in the MOB genome of *G. haimaensis*, *B. azoricus* and *B. japonicus*, with the missing of *norC* and *nosZ* (Table S7). Taken together, our results indicate that both the epibiont and endosymbiont have the abilities of nitrogen

assimilation, which is critical for these bacteria living in the deep ocean with very little nitrogen supply.

Complementary fatty acid metabolism in the holobiont

Both the epibiont and endosymbiont genomes encode a complete type II fatty acid (FA) synthesis pathway but lack most enzymes for FA degradation, such as acyl-CoA dehydrogenase and enoyl-CoA hydratase (Fig. 3 & Table S7). This implies that the two symbionts lack the ability to degrade FAs by beta-oxidation. In contrast, our searching of the host transcripts recovered all these key genes in the FA degradation pathway, whereas some enzymes are missing in the FA biosynthesis pathway, such as 3-oxoacyl-[acyl-carrier protein] reductase and 3-hydroxyacyl-thioester dehydratase. Indeed, the enoyl-[acyl-carrier protein] reductase I (FabI) was among the most abundant 70 proteins of Gammaproteobacteria (Fig. 7). Therefore, the host's requirement for long-chain FAs may be met by the epibiont and endosymbiont. These results highlight the complementary FAs metabolism in the tripartite holobiont. A previous study of the cold-seep vesicomyid clam *Calyptogena phaseoliformis* showed that its endosymbiotic sulfur-oxidizing Gammaproteobacteria could supply n-4 family non-methylene interrupted polyunsaturated fatty acids (NMI-PUFA) to the host (Saito, 2007). Similar to *C. phaseoliformis*, the primary unsaturated FAs are n-4 and n-7 NMI-PUFA in the deep-sea mussels *B. japonicus* and *B. platifrons* which host endosymbiotic methane-oxidizing Gammaproteobacteria, in comparison the shallow-water mussel *Mytilus galloprovincialis* whose FAs are characterized by n-3 unsaturated FAs derived from their algal food (Saito, 2008). Besides, the differences in unsaturated FAs between *Calyptogena* (n-4 NMI-PUFA) and *Bathymodiolus* (n-4 and n-7 NMI-PUFA) illustrate a taxonomic specificity of symbiotic bacteria and differences in FA physiology between thiotrophic and methanotrophic symbionts (Saito, 2008). Our result shows that *G. haimaensis* host thiotrophic epibiont and methanotrophic endosymbiont, which might provide different FAs to their host.

Virulence of symbiotic bacteria

Several genes encoding cold shock proteins (*csp*) were discovered in the endosymbiotic Gammaproteobacteria, with the total transcriptional abundance of all the *csp* transcripts being the highest (Fig. S13). Present in all organisms, Csp regulates the cold adaption of cell by mediating the expression to plasma membrane component and downstream genes (Wu et al., 2006), which improves the cellular tolerance of low temperature, high hydraulic pressure and salinity (Wemekamp-Kamphuis et al., 2002; Schmid et al., 2009). Csp had been demonstrated to play a key role in the stress response of membrane, movement, and biofilm formation in bacteria (Michaux et al., 2017). Besides, Csp has been shown to affect the pathogenic bacterial virulence, which could help the bacteria to resist host cell and adapt to the severe intercellular environment (Wang et al., 2014). Therefore, high expression of *csp* detected in this study may indicate its involvement in adaptation of the endosymbiont to the host's intracellular environment.

Several genes related to the two-component system were found in both two symbiont genomes, including the OmpR family that functions in bacterial adhesion and invasion and chemotaxis protein that functions in signal transduction (Groisman, 2001; Zhao et al., 2020). Notably, the KDP operon response regulator (KdpE, belonging to OmpR family), highly expressed among the gammaproteobacterial sequences (Fig. S13), is an

adaptive regulatory protein that has been shown to affect the virulence and intercellular survival of pathogens (Freeman et al., 2013). The chemotaxis protein CheY is a primary two-component system protein in the campylobacterial epibiont of *G. haimaensis*, which may be responsible for the transmission of external stimulus (Yang et al., 2020). Therefore, the different dominant two-component system proteins between the epibiont and endosymbiont may reflect their responses to the different intracellular and extracellular associations with the host, respectively. In addition, proteases Clp, which is responsible for bacterial virulence, was highly expressed in the endosymbiont but lowly expressed in the epibiont (Frees et al., 2014), which may be related to their different capabilities in host invasion and resistances to host digestion.

Secretion systems and transporters in the symbionts

Protein secretion is crucial for the interaction between symbiotic bacteria and the host (Tseng et al., 2009). The *G. haimaensis* epibiont and endosymbiont encode complete general secretion (Sec) dependent and twin arginine targeting (Tat) dependent translocation pathways (Table S8). These two secretion systems have been shown to translocate folded and unfolded proteins in the periplasmic space, including adhesins, peptidase and toxins (Natale et al., 2008). After transported by the Sec or Tat pathway, these proteins are secreted into the extracellular environment by type II secretion system (T2SS) (Costa et al., 2015). The endosymbiont of *G. haimaensis* encodes a complete T2SS, while the epibiont encodes nearly half of the 11 general secretion pathway protein genes (*gspDGHJE*) with relatively low expression levels comparing to the endosymbiont (Table S8). T2SS, identified in the genomes of SOB and MOB endosymbionts of *B. azoricus*, has been suggested to transport folded exoproteins to the host (Ponnudurai et al., 2017). It is important for the mediation of symbiotic relationship in the symbionts of deep-sea mussel *B. manusensis* and tubeworm *Arcovestia ivanovi*, such as the providing of nutrients to the host, adjusting symbiotic environment, and facilitating the symbiosis establishment (Li et al., 2020). It has also been reported in the campylobacterial symbiont of the vent snail *Alviniconcha marisindica* and *A. boucheti* (Li et al., 2020; Yang et al., 2020). Besides, T2SS has been suggested to be an essential virulence protein in the endosymbiont of the vent tubeworm *Riftia pachyptila* (Li et al., 2018), and mediate cellular interactions between a cold seep sponge and its sulfur-oxidizing symbiont (Tian et al., 2017). The presence of a complete T2SS in the endosymbiont and an incomplete T2SS in the epibiont indicates a potentially strong interactions between the endosymbiont and the host, whereas a weak interaction between the epibiont and the host gill.

Transporters are important for material transfer between eukaryotic hosts and bacterial symbionts (Zheng et al., 2017). We found three and four kinds of amino acids, vitamins and cofactors transporters in the epibiont and endosymbiont genomes, respectively (Table S8). In the epibiont, there were transporters are highly expressed with high protein abundance. For instance, polar amino acid transporter (ranks 1st in the epibiont transcriptome) and general L-amino acid transporter (ranks 22nd in the epibiont transcriptome and 27th in proteome), while these transporters have not been detected in the top 70 highly expressed genes or the top 70 abundant proteins in the endosymbiont (Fig. S4 & S13). These results unveil that the epibiont may be active in the transportation of required amino acids (e.g., tryptophan and tyrosine) to the host (Table 2), in comparison, the endosymbiont may provide such nutrients primarily depend on host intercellular digestion instead of transportation (Yang et al., 2020). Meanwhile, other unique transporters in the symbionts, such as the glucose/mannose and zinc

transporters in the epibiont, and the heme and phospholipid transporters in the endosymbiont, perhaps reflecting their different substrates for transportation. Besides, expression levels of the unique transporters are lower than those of the amino acid transporters (Table S8). For instance, the relatively high level of heme exporter proteins in the endosymbiont allows the exportation of heme (e.g., siroheme) into host cell (Newton & Rice, 2020), which may not be able to synthesize in the host (Table 2). In addition, the epibiont and endosymbiont both encode ammonia, molybdate, phosphate, branched-chain amino acid, and lipopolysaccharide transporters (Table S8). Among them, ammonia transporter is highly expressed in both symbionts, which is responsible for the intake of $\text{NH}_4^+/\text{NH}_3$ to assimilate nitrogen (Yang et al., 2007). In addition, phosphate transporter (efficient phosphate uptake system PstSCAB) is highly expressed in the endosymbiont, which is also identified in the symbiont of *B. puteoserpentis* (Ansorge et al., 2019), indicating the endosymbiont of *G. haimaensis* may obtain phosphate from host cell.

Highly expressed genes in host gill

Gill is the organ that harbors the symbionts. Therefore, to gain insight into the host's functions in symbiosis, we compared the gene expression levels between gill and foot tissues. DEGSeq2 analysis resulting in 1 169 highly expressed genes in the gill tissue and 856 in the foot tissue (Fig. S20). Among the highly expressed genes in the gill tissue, 73 KEGG pathways were enriched, including innate immunity (lysosome, Toll and Imd signaling pathway), small molecules metabolisms (sulfur, amino acid, sugar, and lipid), and transporters and membrane trafficking that are likely related to symbiosis (Fig. S23). Besides, the enrichment of sulfur metabolism genes is consistent with the need of the host to detoxify sulfide, which is in general toxic to eukaryotic cells. It has been suggested that deep-sea mussel *G. platifrons* oxidizes sulfide to less toxic sulfur rapidly using duplicated SQR proteins and provide electrons (Sun et al., 2022). The epibiont of *G. haimaensis* requires sulfur-rich environment to support its energy demand, which may bring more sulfide to the holobiont. Therefore, the active sulfur metabolism in the host gill is significant for its survival. Besides, the active transporters and membrane trafficking proteins prove the nutrients interdependency between the symbionts and host. As we mentioned above, the epibiont encodes a high-level general L-amino acid transporter (Fig. S4 & S13). Particularly, we found four and one genes of L-type amino acid transporter SLC7A9_15 and SLC7A5, respectively, which were highly expressed in the gill of *G. haimaensis* with about 2.2 to 8.4-fold higher than in foot, indicating the active L-amino acid intake in the gill (Palacín & Kanai, 2004; Elolimy et al., 2020). The highly expressed metabolic pathways and transporters in symbiotic gill of *G. haimaensis* indicate its intimate symbiotic relationship with the symbionts (Yang et al., 2020). These results are consistent with previous transcriptomic studies of *B. azoricus*, which revealed the amino acid metabolism and membrane transport genes were enriched in the gill bacteriocytes that harbor endosymbiotic bacteria (Egas et al., 2012).

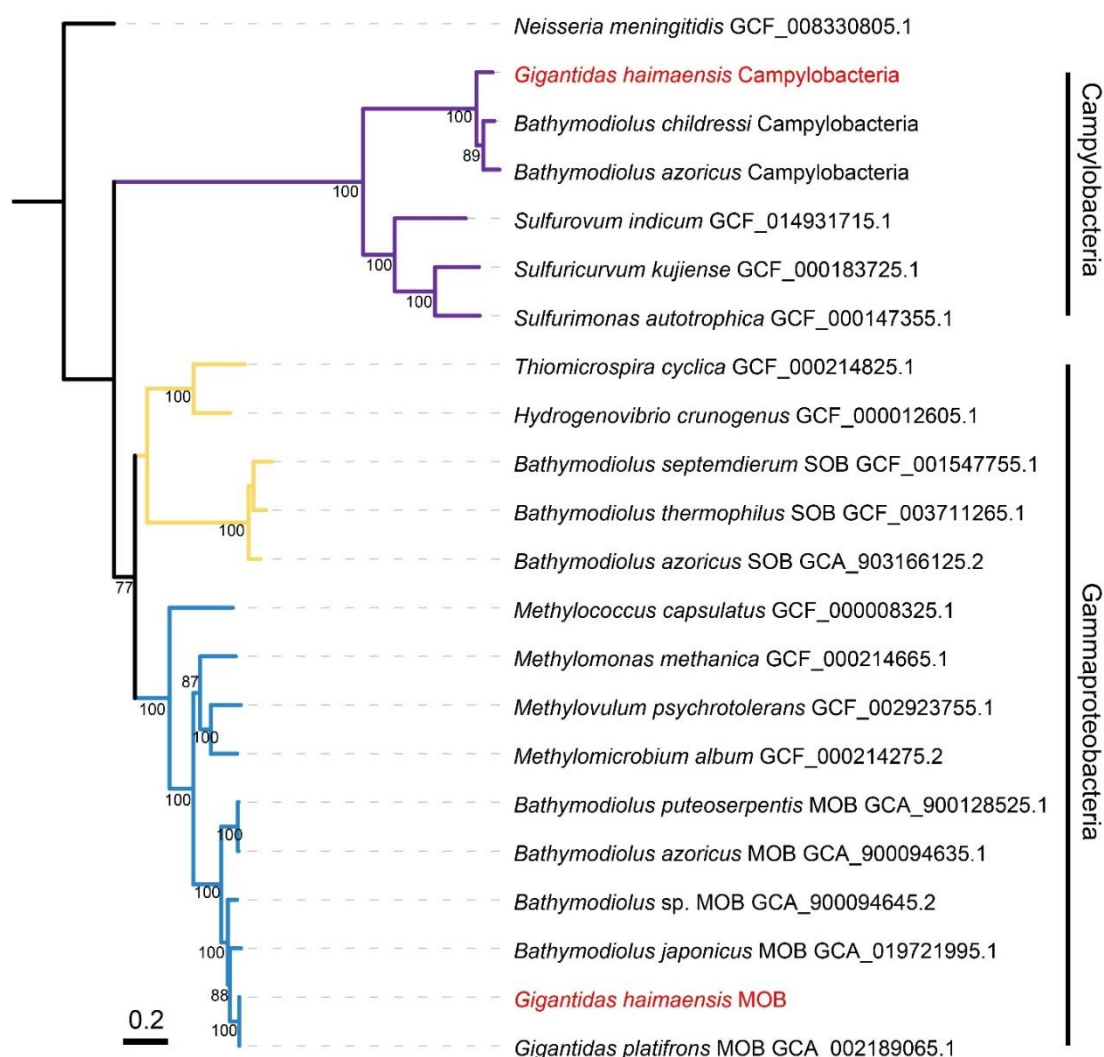
References

- Ansorge R, Romano S, Sayavedra L, Porras MÁ G, 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.
- 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.
- Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, et al. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology*, **13**(6): 343–359.
- Cheng J, Hui M, Sha Z. 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.
- Egas C, Pinheiro M, Gomes P, Barroso C, Bettencourt R. 2012. The transcriptome of *Bathymodiolus azoricus* gill reveals expression of genes from endosymbionts and free-living deep-sea bacteria. *Marine Drugs*, **10**(8): 1765–1783.
- Elolimy A, Alharthi A, Zeineldin M, Parys C, Loor JJ. 2020. Residual feed intake divergence during the preweaning period is associated with unique hindgut microbiome and metabolome profiles in neonatal Holstein heifer calves. *Journal of Animal Science and Biotechnology*, **11**: 13.
- Freeman ZN, Dorus S, Waterfield NR. 2013. The KdpD/KdpE two-component system: integrating K⁺ homeostasis and virulence. *PLOS Pathogens*, **9**(3): e1003201.
- Frees D, Gerth U, Ingmer H. 2014. Clp chaperones and proteases are central in stress survival, virulence and antibiotic resistance of *Staphylococcus aureus*. *International Journal of Medical Microbiology*, **304**(2): 142–149.
- Groisman EA. 2001. The pleiotropic two-component regulatory system PhoP-PhoQ. *Journal of Bacteriology*, **183**(6): 1835–1842.
- Lee RW, Childress JJ. 1996. Inorganic N assimilation and ammonium pools in a deep-sea mussel containing methanotrophic endosymbionts. *Biology Bulletin*, **190**(3): 373–384.
- Li Y, Liles MR, Halanych KM. 2018. Endosymbiont genomes yield clues of tubeworm success. *The ISME Journal*, **12**(11): 2785–2795.
- Li L, Wang M, Li L, Du Z, Sun Y, Wang X, et al. 2020. Endosymbionts of metazoans dwelling in the PACManus hydrothermal vent: diversity and potential adaptive features revealed by genome analysis. *Applied and Environmental Microbiology*, **86**(21): e00815-20.
- Michaux C, Holmqvist E, Vasicek E, Sharan M, Barquist L, Westermann AJ, et al. 2017. RNA target profiles direct the discovery of virulence functions for the cold-shock proteins CspC and CspE. *Proceedings of the National Academy of Sciences*, **114**(26): 6824–6829.
- Natale P, Brüser T, Driessen AJ. 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-distinct translocases and mechanisms. *Biochimica et Biophysica Acta*, **1778**(9): 1735–1756.
- Newton ILG, Rice DW. 2020. The jekyll and hyde symbiont: could wolbachia be a nutritional mutualist?. *Journal of Bacteriology*, **202**(4): e00589-19.
- Palacín M, Kanai Y. 2004. The ancillary proteins of HATs: SLC3 family of amino acid transporters. *Pflügers Archiv European Journal of Physiology*, **447**(5): 490–494.

- Potter L, Angove H, Richardson D, Cole J. 2001. Nitrate reduction in the periplasm of gram-negative bacteria. *Advances in Microbial Physiology*, **45**: 51–112.
- Petersen JM, Zielinski FU, Pape T, Seifert R, Moraru C, Amann R, et al. 2011. Hydrogen is an energy source for hydrothermal vent symbioses. *Nature*, **476**(7359): 176–180.
- Ponnudurai R, Kleiner M, Sayavedra L, Petersen JM, Moche M, Otto A, et al. 2017. Metabolic and physiological interdependencies in the *Bathymodiolus azoricus* symbiosis. *The ISME Journal*, **11**(2): 463–477.
- Reitzer L. 2003. Nitrogen assimilation and global regulation in *Escherichia coli*. *Annual Review of Microbiology*, **57**: 155–176.
- Saito H. 2007. Identification of novel n-4 series polyunsaturated fatty acids in a deep-sea clam, *Calymene phaseoliformis*. *Journal of Chromatography A*, **1163**(1–2): 247–259.
- Saito H. 2008. Unusual novel n-4 polyunsaturated fatty acids in cold-seep mussels (*Bathymodiolus japonicus* and *Bathymodiolus platifrons*), originating from symbiotic methanotrophic bacteria. *Journal of Chromatography A*, **1200**(2): 242–254.
- Schmid B, Klumpp J, Raimann E, Loessner MJ, Stephan R, Tasara T. 2009. Role of cold shock proteins in growth of *Listeria monocytogenes* under cold and osmotic stress conditions. *Applied and Environmental Microbiology*, **75**(6): 1621–1627.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, **30**(14): 2068–2069.
- Sun Y, Wang M, Zhong Z, 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.
- Tseng TT, Tyler BM, Setubal JC. 2009. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiology*, **9**(1): 1–9.
- Tian RM, Zhang W, Cai L, Wong YH, Ding W, Qian PY. 2017. Genome reduction and microbe-host interactions drive adaptation of a sulfur-oxidizing bacterium associated with a cold seep sponge. *mSystems*, **2**(2): e00184–16.
- Wemekamp-Kamphuis HH, Karatzas AK, Wouters JA, Abee T. 2002. Enhanced levels of cold shock proteins in *Listeria monocytogenes* LO28 upon exposure to low temperature and high hydrostatic pressure. *Applied and Environmental Microbiology*, **68**(2): 456–463.
- Wu Q, Pei J, Turse C, Ficht TA. 2006. Mariner mutagenesis of *Brucella melitensis* reveals genes with previously uncharacterized roles in virulence and survival. *BMC Microbiology*, **6**: 102.
- Wang Z, Wang S, Wu Q. 2014. Cold shock protein A plays an important role in the stress adaptation and virulence of *Brucella melitensis*. *FEMS Microbiology Letters*, **354**(1): 27–36.
- Yang H, Xu Y, Zhu W, Chen K, Jiang H. 2007. Detailed mechanism for AmtB conducting NH₄⁺/NH₃: molecular dynamics simulations. *Biophysical Journal*, **92**(3): 877–885.
- Yang Y, Sun J, Sun Y, Kwan YH, Wong WC, Zhang Y, et al. 2020. Genomic, transcriptomic, and proteomic insights into the symbiosis of deep-sea tubeworm holobionts. *The ISME Journal*, **14**(1): 135–150.
- Yang Y, Sun J, Chen C, Zhou YD, Lan Y, Dover CLV, et al. 2020. Tripartite holobiont system in a vent snail broadens the concept of chemosymbiosis. *BioRxiv*. <https://doi.org/10.1101/2020.09.13.295170>.
- Yuan H, Li Y, Wang K. 2021. Effect of influent ammonia nitrogen concentration on microbial community in MBBR reactor. *Water Science & Technology*, **83**(1): 162–172.

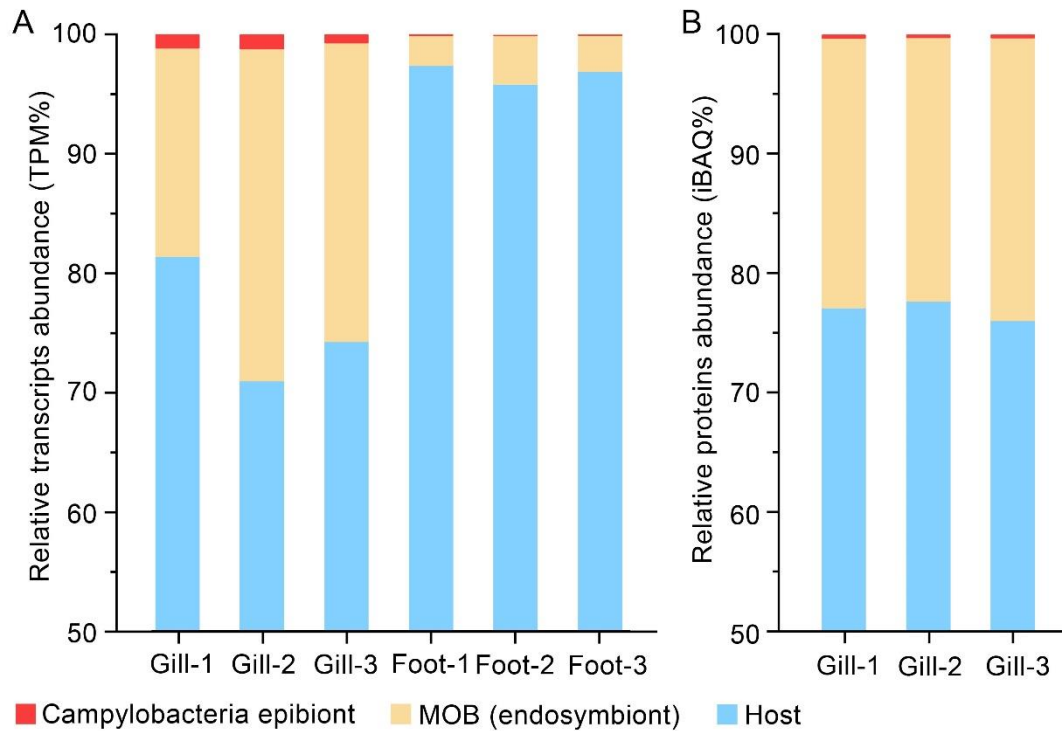
Zheng P, Wang M, Li C, Sun X, Wang X, Sun Y, et al. 2017. Insights into deep-sea adaptations and host-symbiont interactions: A comparative transcriptome study on *Bathymodiolus* mussels and their coastal relatives. *Molecular Ecology*, **26**(19): 5133–5148.

Zhao W, Ma X, Liu X, Jian H, Zhang Y, Xiao X. 2020. Cross-stress adaptation in a piezophilic and hyperthermophilic archaeon from deep sea hydrothermal vent. *Frontiers in Microbiology*, **11**: 2081.



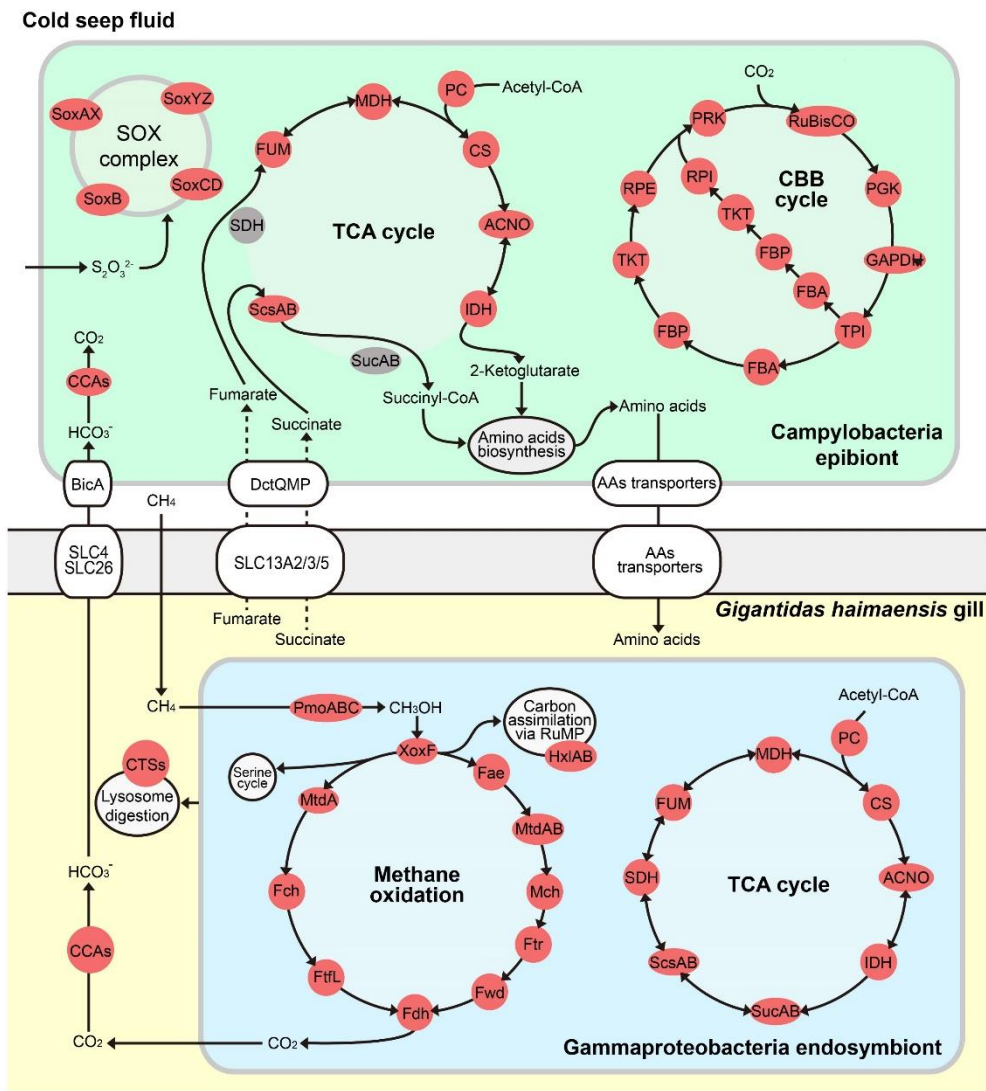
Supplementary Figure S1 Original maximum-likelihood (ML) phylogenetic tree of Figure 3

Tree was constructed using 120 marker genes under the LG+F+R4 model with 1 000 bootstrap replicates. Campylobacterial branch is indicated by purple lines, gammaproteobacterial SOB and MOB branch is indicated by orange and light blue lines, respectively. Genomes assembled in this study are indicated by red letters.



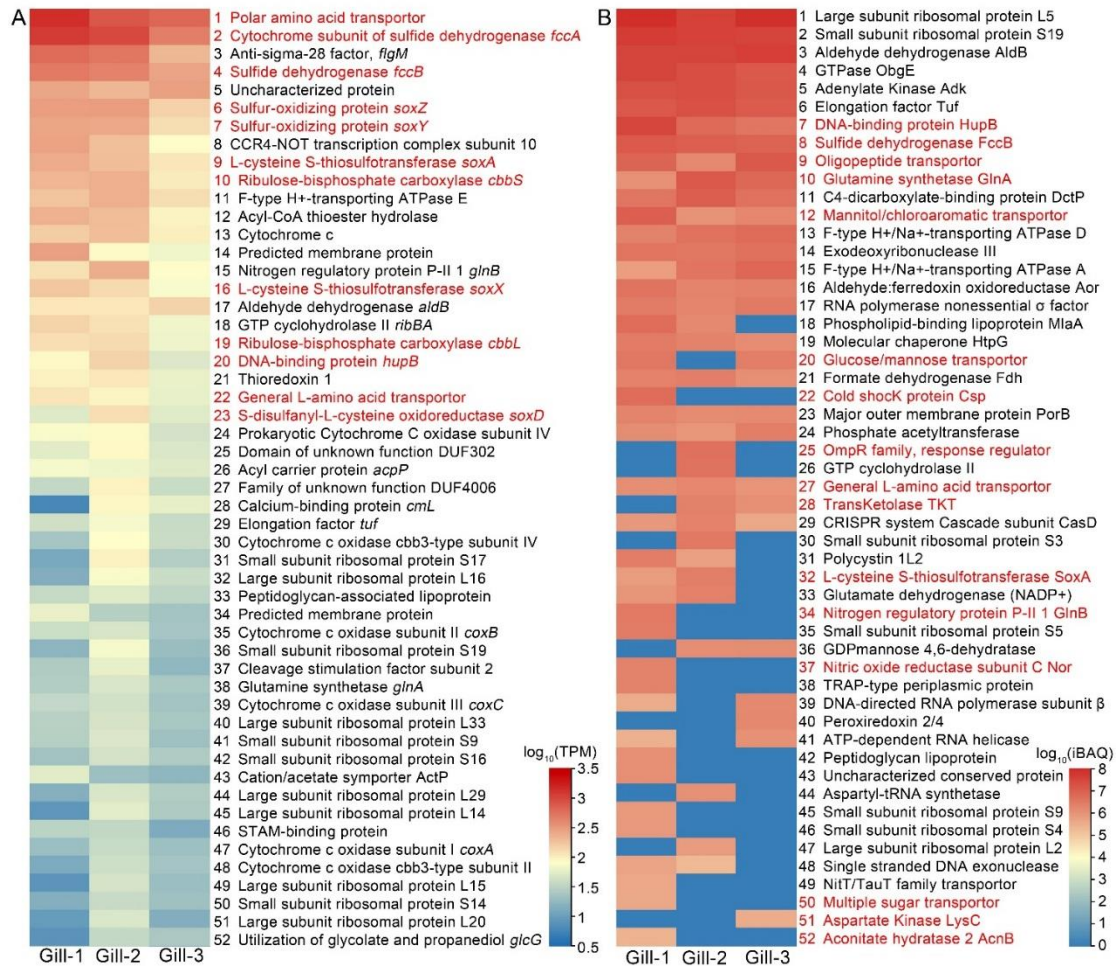
Supplementary Figure S2 Gene expression levels and protein abundances of *G. haimaensis* and its symbionts in the gill and foot tissues

A: Relative gene expression levels of campylobacterial epibiont, gammaproteobacterial endosymbiont (MOB), and host in the gill and foot tissues. B: Relative protein abundances of campylobacterial epibiont, gammaproteobacterial endosymbiont (MOB), and host in the gill tissue.



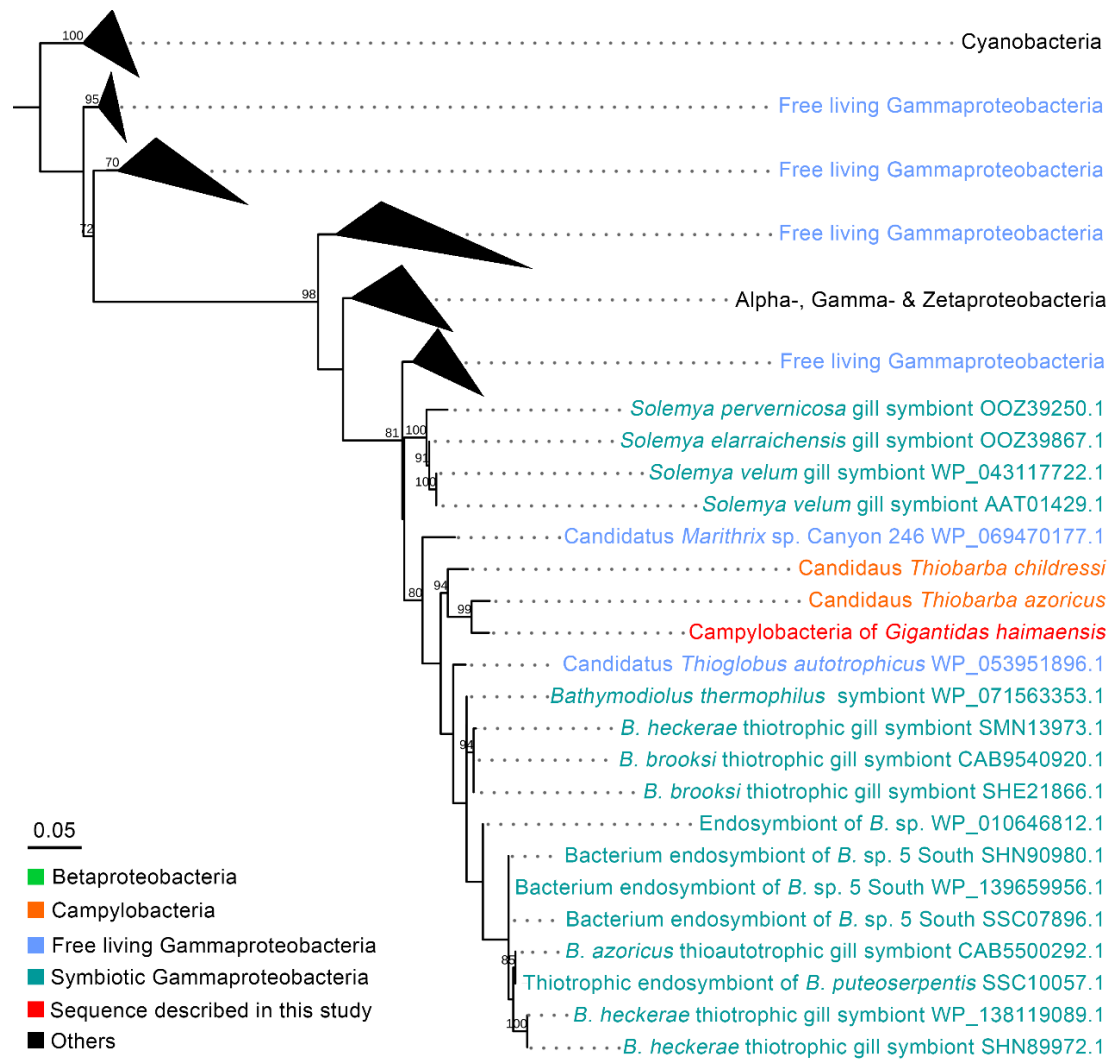
Supplementary Figure S3 Diagram of energy sources, putative carbon assimilation, and TCA cycle in tripartite holobiont of *G. haimaensis*

Host gill cytosol is indicated in yellow, Campylobacteria and Gammaproteobacteria are indicated in green and blue, respectively. Red bubbles indicate enzymes present, gray bubbles indicate missing enzymes. Full names of enzymes are included in Supplementary Table S7.



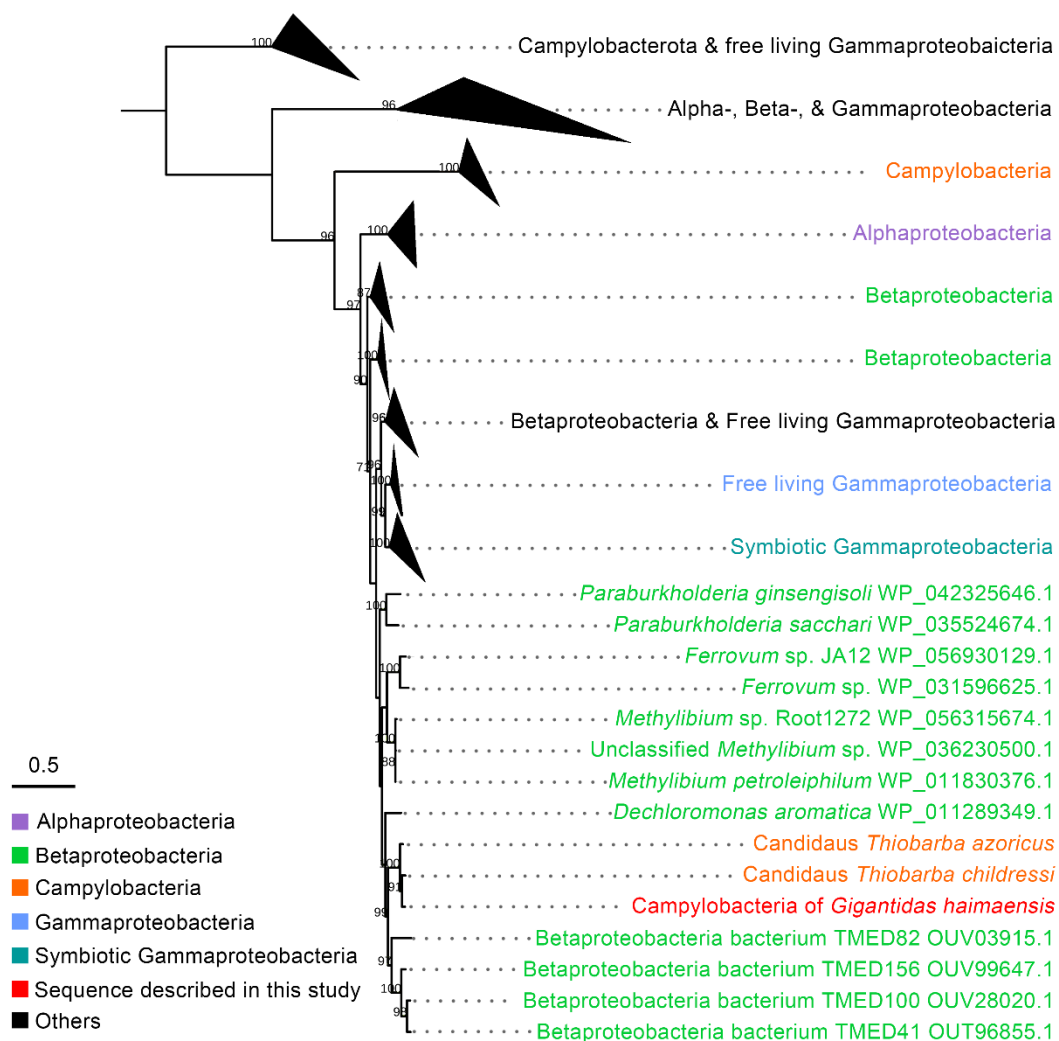
Supplementary Figure S4 Top 52 highly expressed genes (A) and detected proteins (B) of *G. haimaensis* campylobacterial epibiont

Genes and proteins with no annotation or unknown function are excluded. In (A), color gradient represents gene expression level based on log₁₀-transformed transcripts per million (log₁₀(TPM)), and in (B) color gradient indicates protein abundance evaluated according to absolute protein quantification (log₁₀(iBAQ)).



Supplementary Figure S5 ML tree of amino acid sequences of ribulose 1,5-bisphosphate carboxylase large subunit (CbbL) involved in CBB cycle

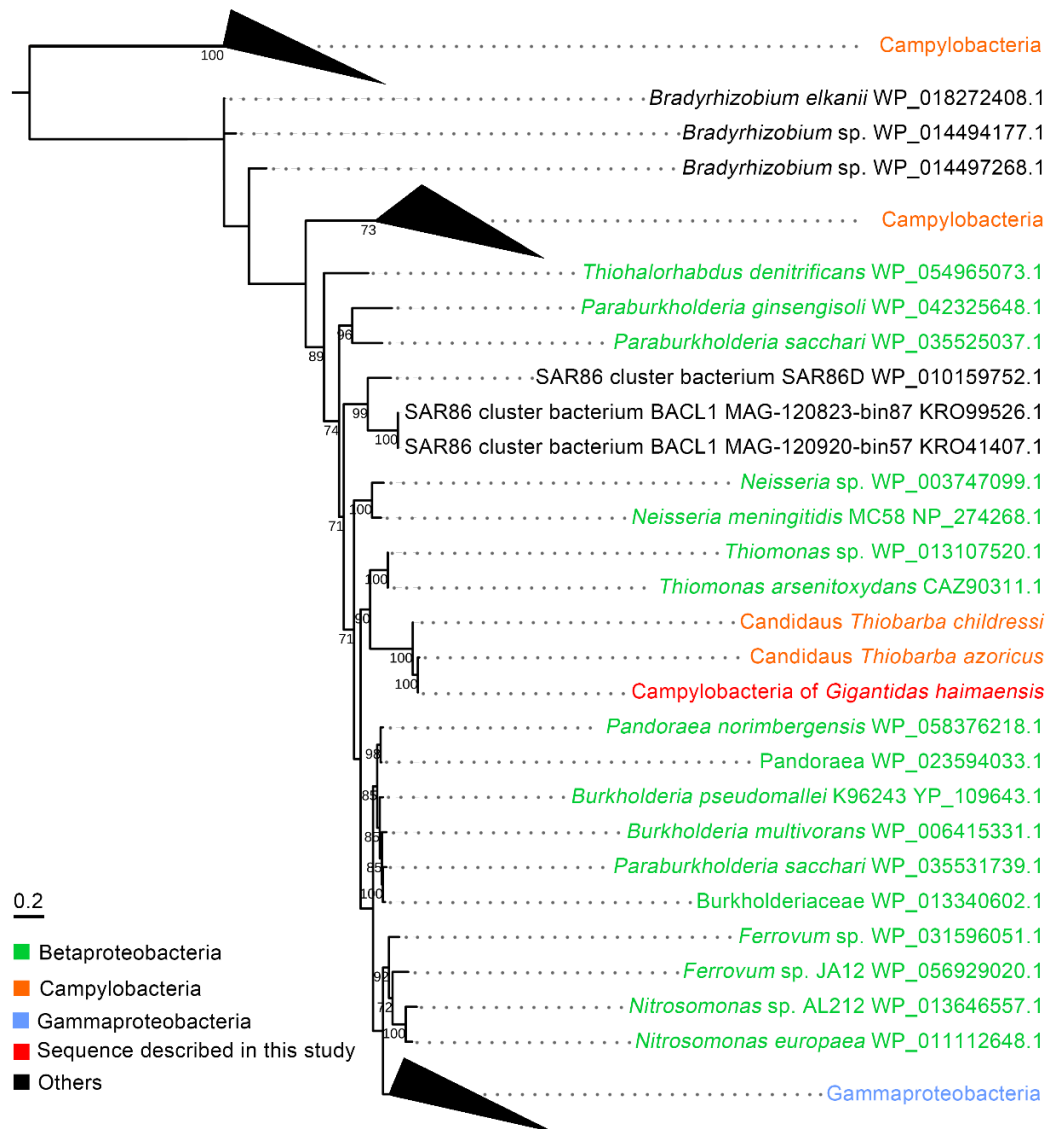
Tree was constructed based on LG+R4 model with 1 000 bootstrap replicates. Each branch comprising sequences from the same taxon was collapsed into one leaf. Numbers associated with each terminal are identical to those in Supplementary Figure S6 of Assié et al. (2020).



Supplementary Figure S8 ML tree of amino acid sequences of fructose biphosphate aldolase (FBA) involved in the CBB cycle

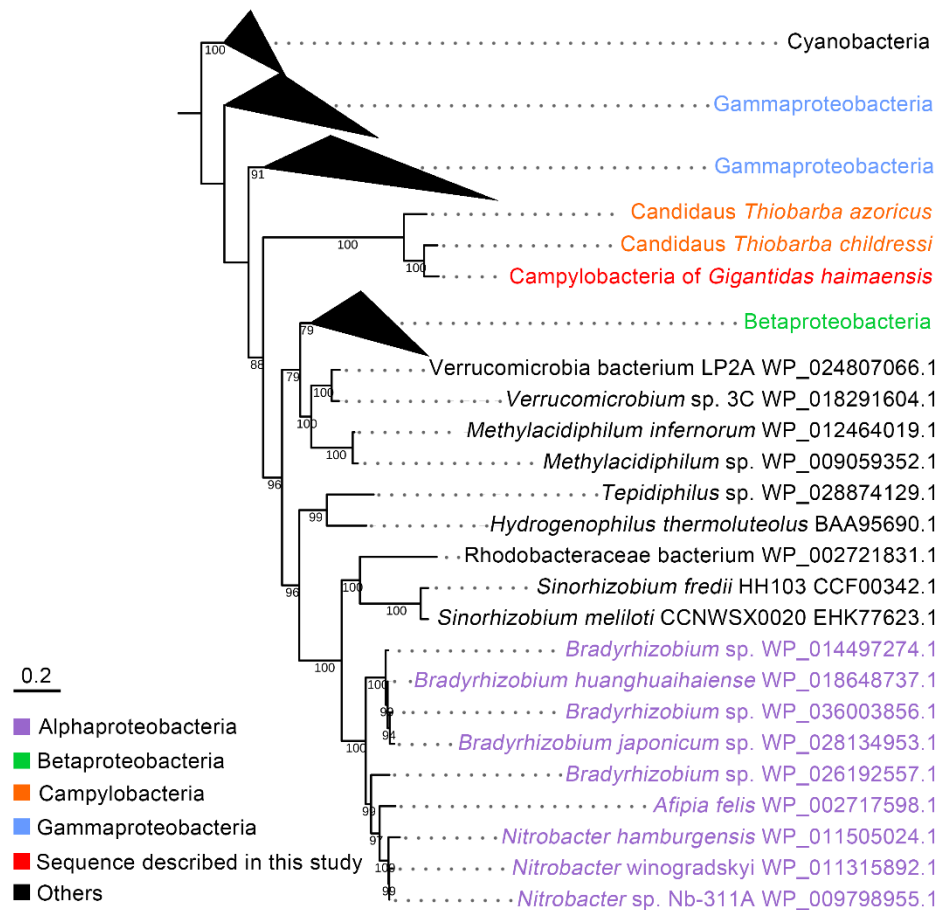
Tree was constructed based on LG+R5 model with 1 000 bootstrap replicates. Each branch comprising sequences from the same taxon was collapsed into one leaf. Numbers associated with each terminal are identical to those in Supplementary Figure S11 of Assié et al. (2020).

Tree was constructed based on LG+R5 model with 1 000 bootstrap replicates. Each branch comprising sequences from the same taxon was collapsed into one leaf. Numbers associated with each terminal are identical to those in Supplementary Figure S13 of Assié et al. (2020).



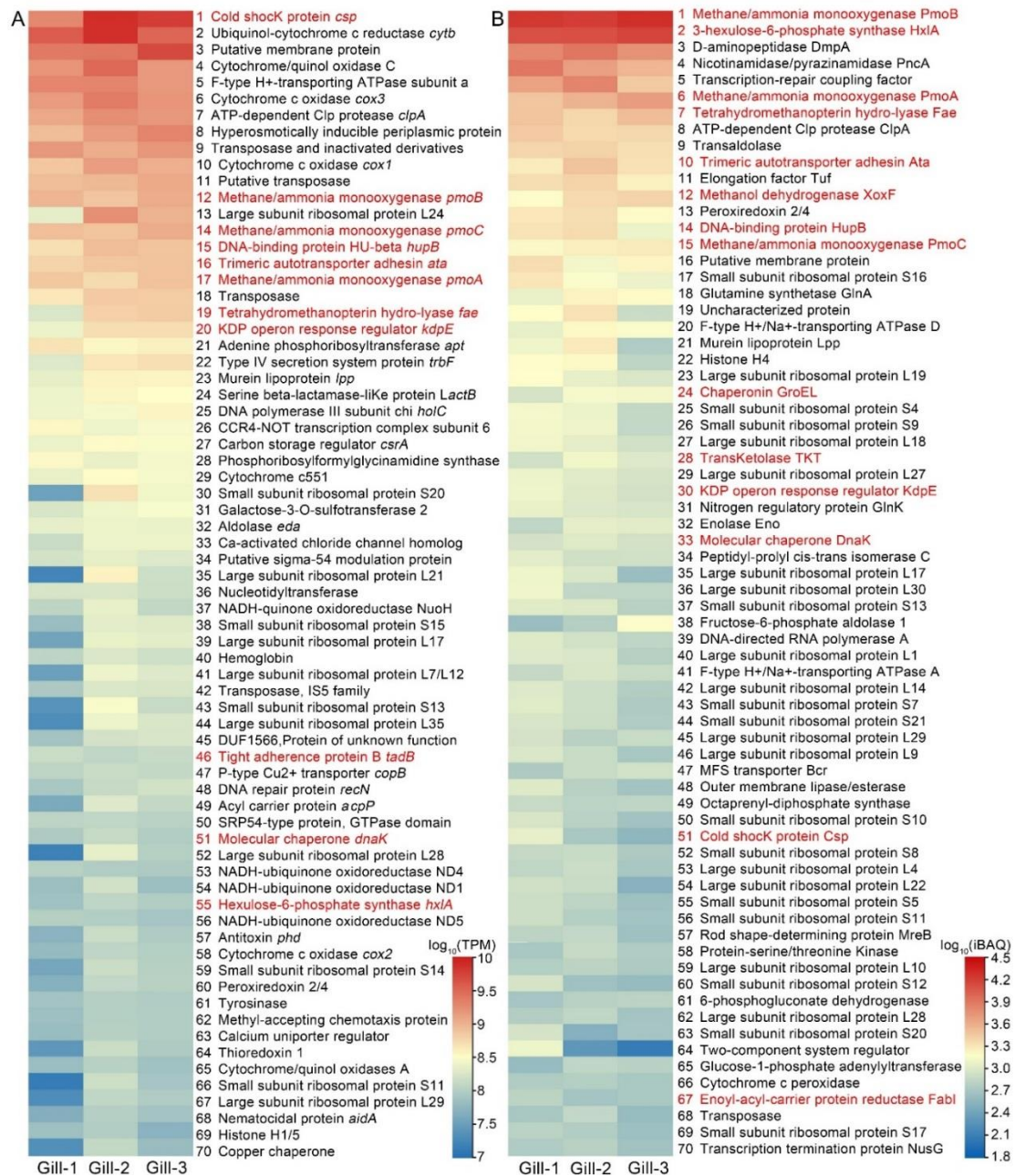
Supplementary Figure S11 ML tree of amino acid sequences of ribulose-phosphate-3-epimerase (RPE) involved in the CBB cycle

Tree was constructed based on LG+G4 model with 1 000 bootstrap replicates. Each branch comprising sequences from the same taxon was collapsed into one leaf. Numbers associated with each terminal are identical to those in Supplementary Figure S14 of Assié et al. (2020).



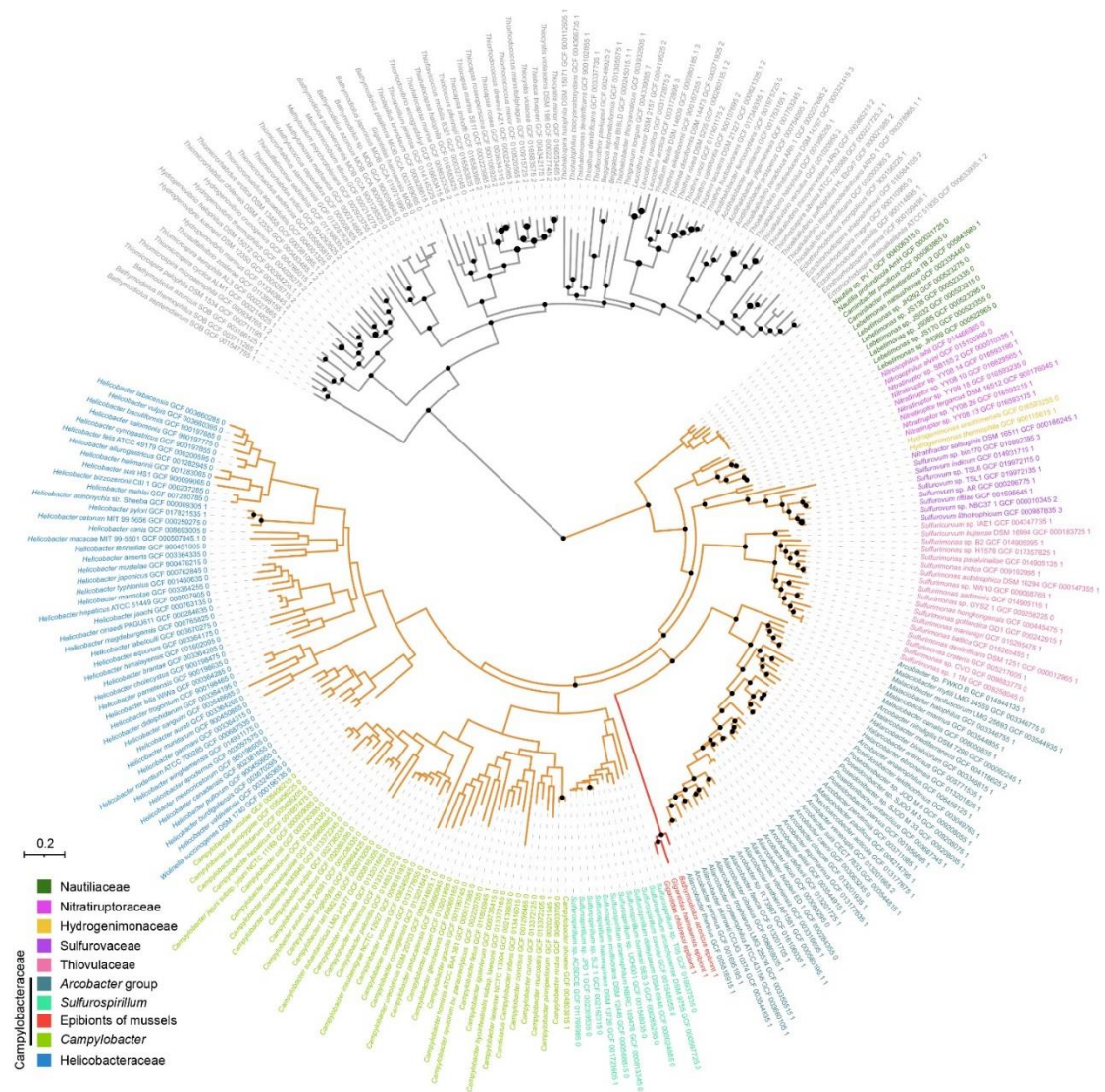
Supplementary Figure S12 ML tree of amino acid sequences of phosphoribulokinase (PRK) involved in the CBB cycle

Tree was constructed based on LG+G4 model with 1 000 bootstrap replicates. Each branch comprising sequences from the same taxon was collapsed into one leaf. Numbers associated with each terminal are identical to those in Supplementary Figure S8 of Assié et al. (2020).



Supplementary Figure S13 Top 70 expressed genes (A) and abundant proteins (B) of the *G. haimaensis* gammaproteobacterial endosymbiont

Unknown genes and proteins were excluded. Color gradient represents transcriptional and protein levels based on \log_{10} -transformed transcripts per million ($\log_{10}(\text{TPM})$) and intensity-based absolute protein quantification ($\log_{10}(\text{iBAQ})$), respectively.



Supplementary Figure S15 Phylogenetic relationships among Campylobacteria and distribution of SoxB orthologs

Black dots at branch points indicate distribution of SoxB orthologs identified by OrthoFinder. Number after a genome association number represents number of SoxB orthologs. Campylobacteria are in orange, Gammaproteobacteria are in gray and considered to be the outgroup. Tree was constructed using 120 marker genes and the Q.yeast+R10 model with 1 000 bootstrap replicates.

Ornithine and Arg										Val, Ile (from Thr), and Leu										Chorismate and Phe										Trp (from chorismate)										Homoserine and Thr									
Host	argJ	argB	argC	argD	argF	argG	argH	ilvBH	ilvG	ilvD	ilvE	ilvA	leuA	leuCD	leuB	aroG	aroB	aroQ	aroE	aroK	aroA	aroC	pheA	hisC	trpEG	trpD	trpF	trpC	trpAB	lysC	asd	thrA	thrB	thrC															
Camp																																																	
Gamm																																																	

Existing gene

Missing gene

Met (from homoserine)				Lys				His				Tyr		Pro		Cys		Glu	Gln	Gly	Ser	Ala	Asp	Asn									
Host	metA	metB	metH	dapA	dapB	dapD	dapC	dapE	dapF	lysA	hisG	hisE	hisI	hisA	hisH	hisB	hisD	pheA	tyrA	proB	proA	proC	cysE	cysK	gluB	glnA	glyA	serA	serC	serB	aspT	aspB	asnB
Camp																																	
Gamm																																	

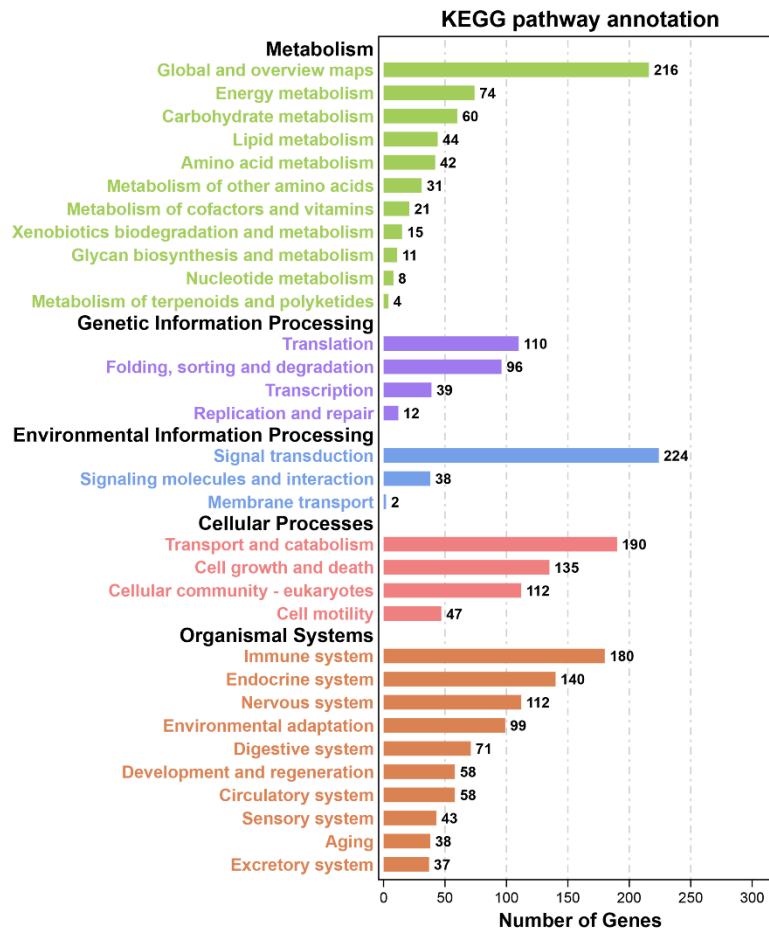
Biotin				Lipoic acid		Riboflavin and FAD				Folate				Pantothenate			CoA		Thiamine		Pyridoxine and pyridoxine phosphate												
Host	bioC	bioH	bioF	bioA	bioD	bioB	lipB	lipA	ribA	ribD	ribEH	ribF	folE	folB	folK	folP	folC	folA	ilvE	panB	apbA	panC	dip	coaD	coaE	thiC	thiD	thiE	pdxH	pdxB	serC	pdxA	pdxJ
Camp																																	
Gamm																																	

Protoheme and siroheme										Ubiquinone						NAD		Glutathione		Isopentenyl diphosphate (from pyruvate)						Isopentenyl diphosphate (from Acetyl-CoA)										
Host	gltX	hemA	hemI	hemB	hemC	hemD	hemE	hemN	hemK	hemH	cysG	ubiC	ubiA	ubiDX	ubiI	ubiG	ubiH	ubiE	ubiF	nadC	nadD	andE	gshA	gshB	dxs	dxr	ispD	ispE	ispF	ispG	lytB	hcs	mvaA	mvK	pmvK	mvD
Camp																																				
Gamm																																				

Supplementary Figure S16 Genes related to biosynthesis of amino acids, vitamins, and cofactors found in *G. haimaensis* gill and its epibiont and endosymbiont

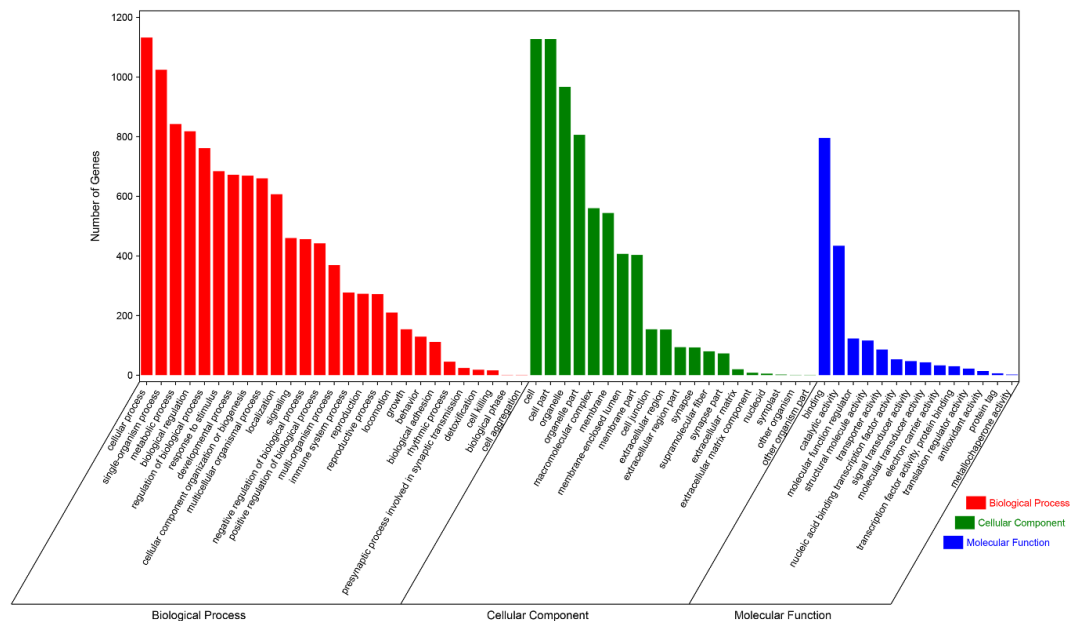
Green box indicates gene is present, and gray indicates gene is missing.

Abbreviations: Camp: Campylobacteria; Gamm: Gammaproteobacteria; Genes and proteins required for each pathway were obtained from Ponnudurai et al. 2017a. Full gene names and their expression levels are given in Supplementary Table S9.



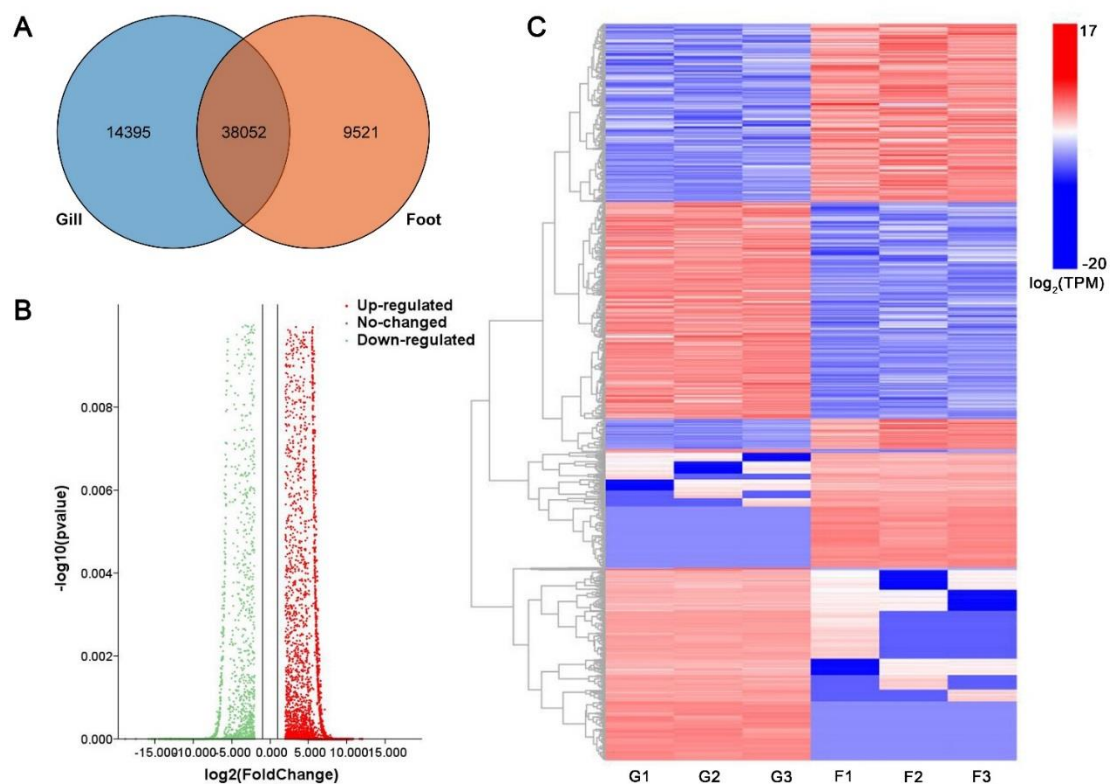
Supplementary Figure S18 Top 30 KEGG enrichment results of top 10% of expressed genes in gill transcriptome

Different groups are labeled in different colors. Unknown functions or those related to human disease were excluded.



Supplementary Figure S19 Top 30 GO enrichment results of top 10% of expressed genes in gill transcriptome

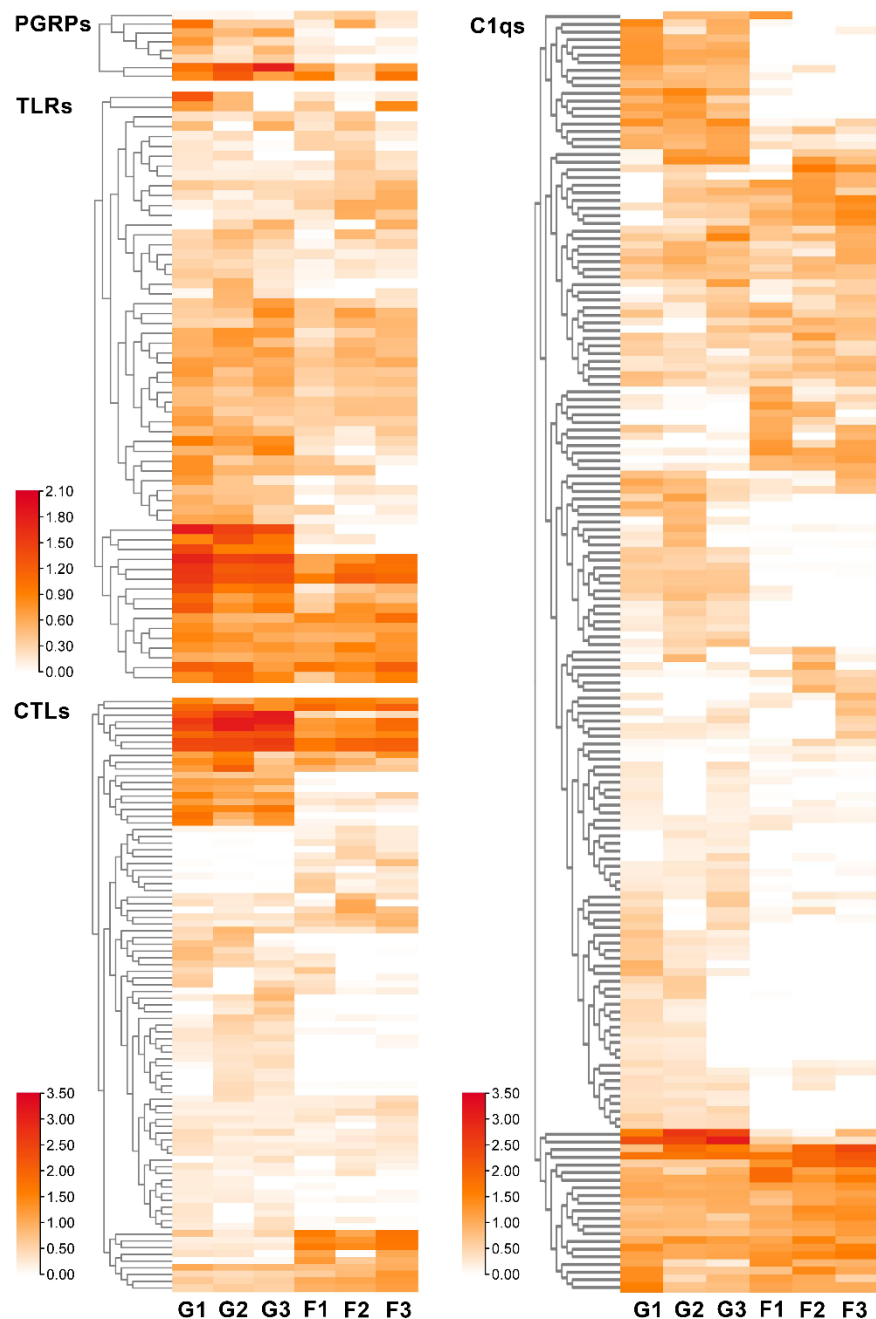
Different groups are labeled in different colors. Unknown functions or those related to human disease were excluded.



Supplementary Figure S20 Differentially expressed genes (DEGs) between *G.*

***haimaensis* gill and foot**

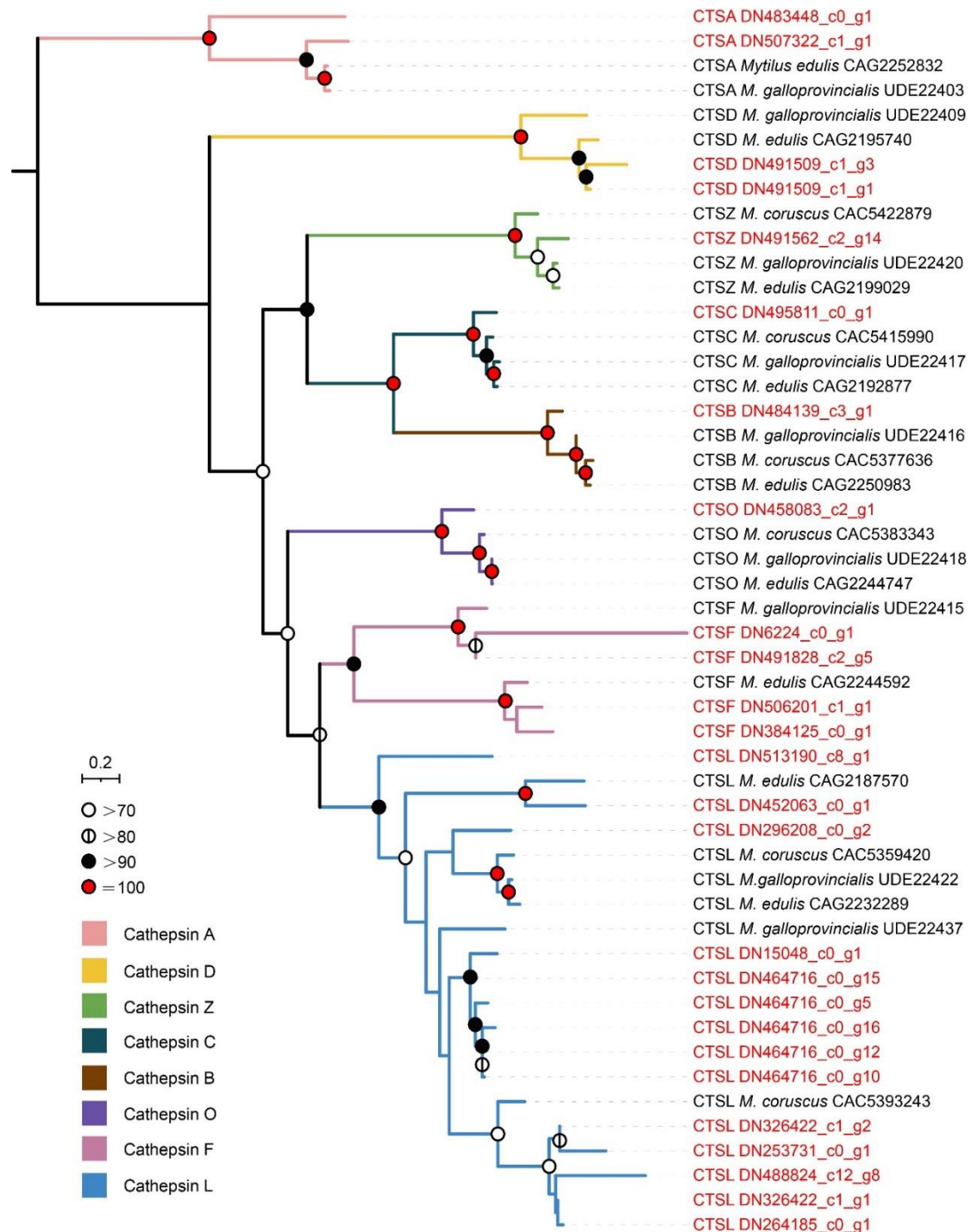
A: Venn diagram showing unigenes expressed in gill and foot (TPM \geq 0.5). B: Volcano plot of DEGs between gill and foot ($q < 0.01$). C: Heat map of annotated DEGs between foot and gill ($q < 0.01$). G1–G3 and F1–F3 indicate gill and foot tissues from mussels 1–3, respectively.



Supplementary Figure S21 Heat map of pattern recognition receptor (PRR) transcripts in host transcriptome

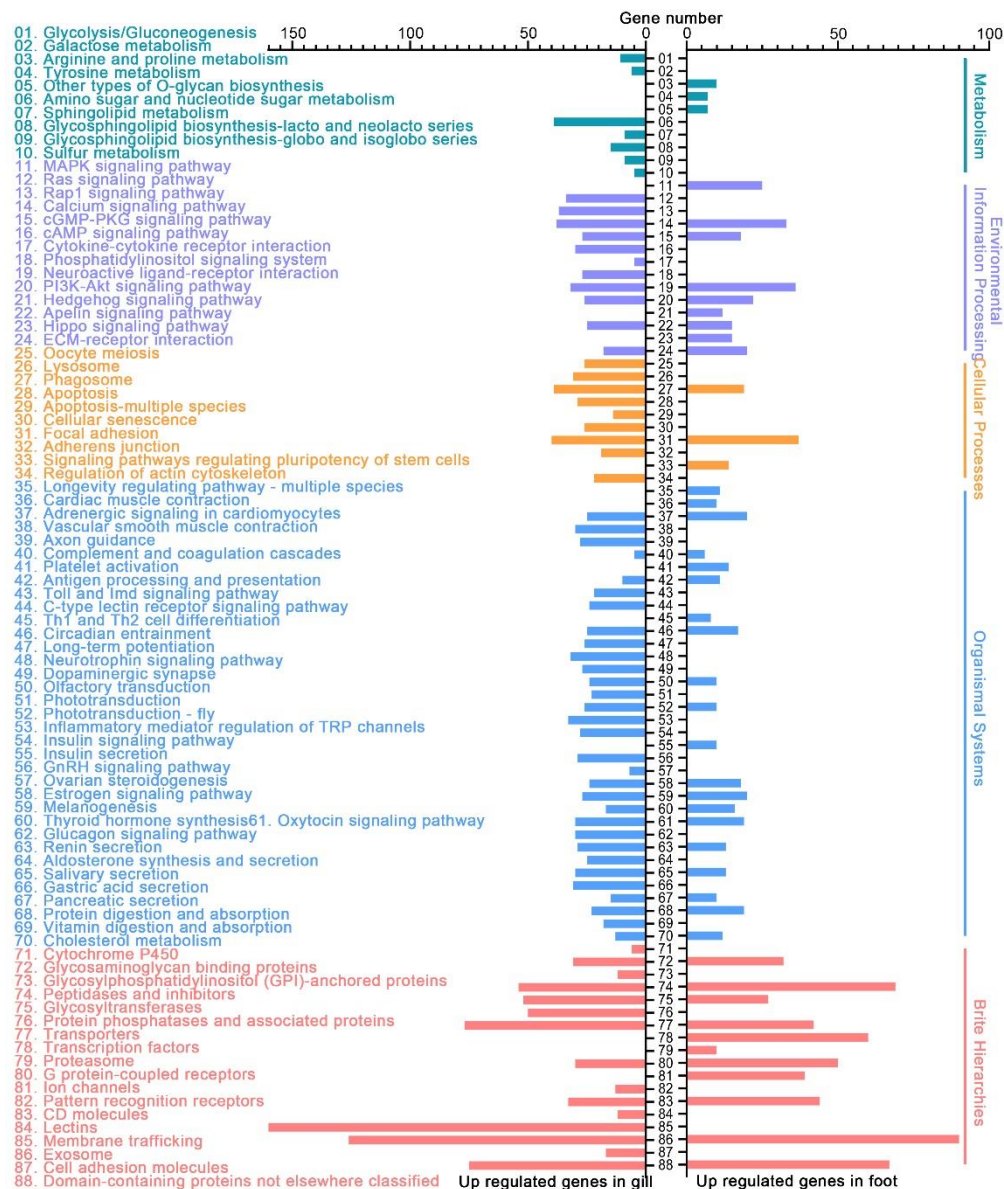
G1–G3 and F1–F3 indicate gill and foot tissues from mussels 1–3, respectively. Color

gradient represents transcription levels based on \log_{10} -transformed transcripts per million ($\log_{10}(\text{TPM})$).



Supplementary Figure S22 ML tree of amino acid sequences of cathepsins (CTS) in selected bivalves

Tree was constructed based on WAG+F+G4 model with 1 000 bootstrap replicates, sequences identified in this study are indicated by red letters.



Supplementary Figure S23 KEGG enrichment of DEGs between *G. haimaensis* gill and foot tissues

Different groups are labeled by different colors. Unknown functions or those related to human disease were excluded.

Supplementary Tables

Supplementary Table S1 List of datasets included in custom database for metatranscriptomic split and protein identification

Supplementary Table S2 Summary of data output for DNA and RNA sequencing

Supplementary Table S3 Summary of de novo assembly of DNA and RNA sequences

Supplementary Table S4 Summary of functional annotation of *G. haimaensis* gill-associated bacterial genomes

Supplementary Table S5 Summary of annotation results for assembled transcriptome of host gill and foot tissues

Supplementary Table S6 Summary of protein identification results in gill metaproteome

Supplementary Table S7 Distribution of genes involved in selected central metabolic pathways among bathymodioline symbiont genomes and their free-living relatives

Supplementary Table S8 Genes involved in secretion system and transporters in *G. haimaensis* symbionts

Supplementary Table S9 List of genes related to biosynthesis of amino acids, vitamins, and cofactors in *G. haimaensis* holobiont

Supplementary Table S10 Symbiosis-related genes and their transcription levels in *G. haimaensis*

Supplementary Table S11 Symbiosis-related proteins and their abundances in *G. haimaensis*

Supplementary Table S12 Gene abbreviations in Figure 7

Supplementary Tables S1–S12 are listed as a separate Excel file.