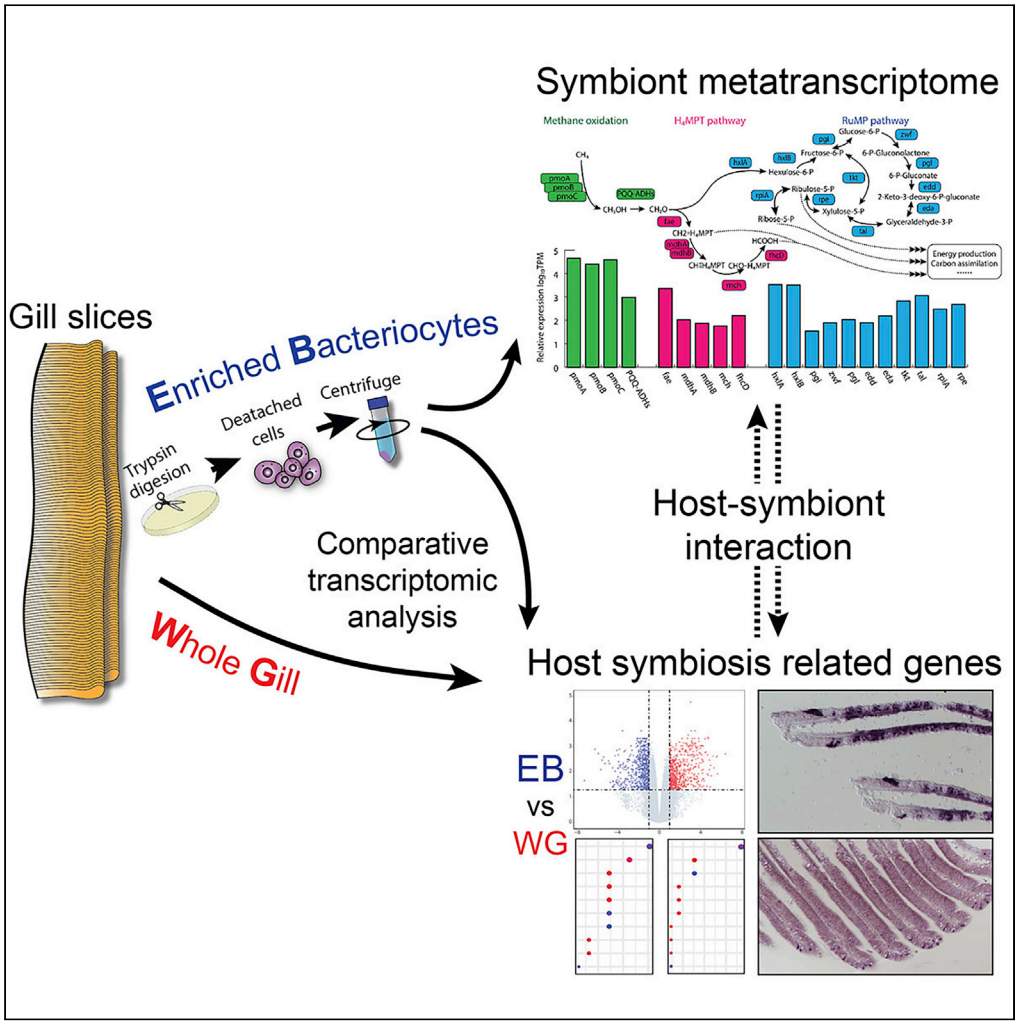
iScience

CellPress
OPEN ACCESS

Article

Molecular analyses of the gill symbiosis of the bathymodiolin mussel *Gigantidas platifrons*



Hao Wang, Huan Zhang, Zhaoshan Zhong, ..., Lei Cao, Chao Lian, Chaolun Li

lcl@qdio.ac.cn

HIGHLIGHTS

Breakdown of the symbiotic tissue to characterize the deep-sea mussel symbiosis

Bacteriocytes provide the symbiont with metabolites, shelter, and population control

Non-symbiotic parts of the gill also play crucial roles in maintaining the symbiosis

The mussel and the symbiont work together to maintain the high-efficiency symbiosis

Wang et al., iScience 24, 101894
January 22, 2021
https://doi.org/10.1016/j.isci.2020.101894

iScience

CellPress
OPEN ACCESS

Article

Molecular analyses of the gill symbiosis of the bathymodiolin mussel *Gigantidas platifrons*

Hao Wang,[1,2] Huan Zhang,[1,2] Zhaoshan Zhong,[1,2,5] Yan Sun,[2,3] Minxiao Wang,[2,3] Hao Chen,[1,2] Li Zhou,[1,2] Lei Cao,[1,2] Chao Lian,[1,2] and Chaolun Li[1,2,3,4,5,6,*]

SUMMARY

Although the deep-sea bathymodiolin mussels have been intensively studied as a model of animal-bacteria symbiosis, it remains challenging to assess the host-symbiont interactions due to the complexity of the symbiotic tissue—the gill. Using cold-seep mussel *Gigantidas platifrons* as a model, we isolated the symbiont harboring bacteriocytes and profiled the transcriptomes of the three major parts of the symbiosis—the gill, the bacteriocyte, and the symbiont. This breakdown of the complex symbiotic tissue allowed us to characterize the host-symbiont interactions further. Our data showed that the gill's non-symbiotic parts play crucial roles in maintaining and protecting the symbiosis; the bacteriocytes supply the symbiont with metabolites, control symbiont population, and shelter the symbiont from phage infection; the symbiont dedicates to the methane oxidation and energy production. This study demonstrates that the bathymodiolin symbiosis interacts at the tissue, cellular, and molecular level, maintaining high efficiency and harmonic chemosynthetic micro niche.

INTRODUCTION

Animal-microbe symbioses have played fundamental roles in the animals' adaptation, ecology, and evolution (Bang et al., 2018; McFall-Ngai et al., 2013). Many symbioses involve nutritional benefit for both partners, such that the animal provides the microbe with optimal growth conditions, necessary metabolites, and shelter, whereas the symbiont provides its host with nutrients that are otherwise lacking in the host's diet (Moya et al., 2008; Dubilier et al., 2008). To maintain the stability and efficiency of the symbiosis, the host must be able to regulate and interact with the symbiont at the tissue, cell, and molecular levels (Hinzke et al., 2019; Finlay and Medzhitov, 2007). Although in many symbioses, the host-microbe interactions are challenging to assess due to the high number of microbes potentially involved (McFall-Ngai, 2008; Dubilier et al., 2008), deep-sea bathymodiolin mussels (*Mytilidae, Bathymodiolinae*) provide us a unique low-complexity model for symbiosis studies (Dubilier et al., 2008).

The bathymodiolin mussel is one of the most common macrofaunae in deep-sea chemosynthetic ecosystems, found in cold-seeps, hydrothermal vents, sunken woods, and whale falls (Taylor and Glover, 2010), and they often dominate the biomass of these chemosynthesis-based communities and may even form dense mussel beds that serve as critical habitats for many other animals (Govenar, 2010; Vrijenhoek, 2010). Although bathymodiolin mussels retain the ability to filter-feed, their ecological successes can largely be attributed to the symbiosis in their gill: bathymodiolin mussels host intracellular thiotrophic and/or methanotrophic Gammaproteobacteria in their gill (with a few of exceptions in genera *Idas* and *Adipicola* that host gill epi-symbiont) for nutrition (Dubilier et al., 2008; Xu et al., 2019).

Numerous studies (Halary et al., 2008; Sun et al., 2017b; Zheng et al., 2017; Bettencourt et al., 2010; Bettencourt et al., 2017; Martins et al., 2014; Wang et al., 2019) have found that the bathymodiolin mussels gill has morphological adaptations and sophisticatedly regulated molecular mechanisms, keeping the gill symbiosis working as a highly efficient chemosynthetic micro niche. The gill of the bathymodiolin mussel, which comprises hundreds of filaments, has a greatly enlarged surface area and cilia ventilation (Fiala-Médioni et al., 1986). On each gill filament, the bacteriocytes, specialized gill epidermal cells that host the symbiont within the intracellular vacuoles, face the apical top of the filament, where the symbionts could easily reach

[1]Center of Deep-Sea Research, Institute of Oceanology, Chinese Academy of Sciences, 7 Nanhai Road, Qingdao 266071, P. R. China

[2]Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, P. R. China

[3]Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, P. R. China

[4]Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

[5]University of Chinese Academy of Sciences, Beijing, China

[6]Lead contact

*Correspondence: lcl@qdio.ac.cn

https://doi.org/10.1016/j.isci.2020.101894

the necessary gas such as hydrogen sulfide, methane, and oxygen from the vent or seep fluid (Halary et al., 2008). Metagenomic studies of bathymodiolin mussel symbiosis have shown that the symbiont may use a variety of vent or seep gases as energy sources, including H_2S, H_2, and CH_4 (Jillian and Nicole, 2009; Ponnudurai et al., 2017a; Ponnudurai et al., 2017b), providing the host mussel not only with a carbon source but also with necessary sterols (Takishita et al., 2017). Several RNA-seq projects on whole gill have revealed that a substantial number of genes with potential involvement in host-symbiosis interaction, such as lysozymes, nutrient transporters, stress response genes, and immune genes, are highly expressed in the gill (Zheng et al., 2017; Bettencourt et al., 2017; Barros et al., 2015; Bettencourt et al., 2009; Sayavedra et al., 2015; Wong et al., 2015).

Compared with the trophosome of *Siboglinidae* tubeworms, which is a homogeneous tissue playing a sole function of maintaining symbiosis (Hand, 1987; Monika and Angelika, 2003), the gill of bathymodiolin mussel is a more complicated tissue, which also executes other basic physiological functionalities such as respiration, transportation, growth, and immune defense (Harper, 2005; Neumann and Kappes, 2003; Nguyen et al., 2019). For example, high expression of immune genes in gills has also been reported in shallow water or intertidal non-symbiotic bivalve species, likely because the gill is always in contact with environmental bacteria during filter-feeding (Philipp et al., 2012). Furthermore, the bathymodiolin gill includes a variety of cell types. Beyond its bacteriocytes, the bathymodiolin gill filament has several nonsymbiotic parts: the tip and edge of each gill filament are covered with cilia cells and mucous cells, and on the lateral face of the gill filament, numerous intercalary cells fill the intercellular spaces among the bacteriocytes (Fiala-Médioni et al., 1986; Barry et al., 2002; Fujiwara et al., 2000). Geier et al. demonstrated that symbiont-free cells, such as the heavily ciliated tip of the gill filament, are processing different metabolites compared with the bacteriocytes (Geier et al., 2020). mRNA *in situ* hybridization studies have also shown that genes that are related or unrelated to symbiosis are specifically expressed in the bacteriocytes or other cells, respectively (Takishita et al., 2017; Ikuta et al., 2019). These accumulating data demonstrate that spatial expression information for genes is crucial for understanding symbiosis. However, because transcriptomic studies use whole-gill tissue homogenate as research material, valuable spatial information is lost.

Up to date, bathymodiolin symbiosis has not yet been stably maintained in an atmosphere lab aquarium, nor have the symbionts been cultured. These limitations greatly restrict the timing of the sampling and available technologies and make the identification of symbiosis-related genes extremely difficult. In this work, we conducted integrated molecular analyses of the gill of model deep-sea mussel *Gigantidas platifrons*, previously known as *Bathymodiolus platifrons* (Xu et al., 2019), and further analyzed its symbiosis at the molecular level. The species *G. platifrons* could be commonly found in cold seeps of the South China Sea (Feng et al., 2015), Sagami Bay (Barry et al., 2002) as well as hydrothermal vents in the Okinawa Trough (Fujiwara et al., 2000). Although these habitats are geologically distant from each other, *G. platifrons* lives in symbiosis with one single species of methane-oxidizing bacteria (Barry et al., 2002; Fujiwara et al., 2000), suggesting that this symbiosis is robust and highly adaptive. Most importantly, the genome of the host mussel (Sun et al., 2017a) and the metagenome of its symbiont (Takishita et al., 2017) have been sequenced, which significantly facilitates downstream analyses. In our repeated visits to the F-site cold seep of Jiulong reef, we found that the bacteriocytes of *G. platifrons* could be enriched by treating the gill in trypsin-EDTA solution, followed by centrifugation. We successfully enriched *G. platifrons* bacteriocytes, which were directly sequenced by Illumina RNA-seq. In comparative analyses of genes enriched in the bacteriocytes and whole gill and in metatranscriptomic analyses of the symbiont, we produced an integrated view of the molecular mechanisms that directly govern symbiosis in *G. platifrons*.

RESULTS

Gill structure of *G. Platifrons*

The gill of *G. platifrons* (Figure 1A) has a typical filibranch bivalve structure, involving cilia ventilation that provides symbiotic bacteria with necessary gases, such as methane and oxygen, from the seep fluid (Figure 1B). Double FISH analyses showed that each gill filament's lateral face was covered with bacteriocytes, with the exception of the heavily ciliated filament tip. Each bacteriocyte hosts a dense population of a single species of methane-oxidizing (MOX) symbiont (Figure 1C). Symbiont-free intercalary cells also appear between the bacteriocytes (Figure 1D).

SEM images showed that each bacteriocyte's apical top protrudes slightly above the surrounding intercalary cells (Figures 2A and 2B). Although the surface of the bacteriocytes is smooth, the intercalary cells have

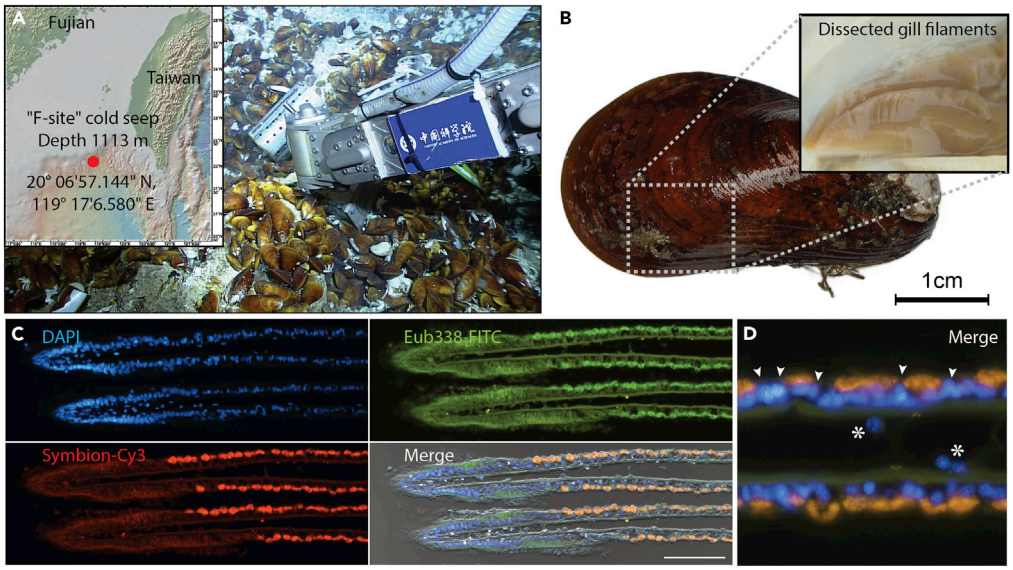


Figure 1. The symbiosis of the deep-sea mussel *Gigantidas Platifrons*

(A) Sampling site of the deep-sea mussel *G. platifrons* at the "F-site" cold seep, Jiulong reef. Image shows the high-density *G. platifrons* community. Insert: regional location map of the research site drawn by GeoMapApp (www.geomapapp.org).

(B) A healthy *G. platifrons* specimen; the inset shows the dissected gill filaments.

(C) Double-FISH (fluorescent *in situ* hybridization) with FITC-labeled eubacteria probe EU338, Cy3-labeled *G. platifrons* symbiont probe.

(D) Magnified merge image showing the bacteriocytes that host symbiont, the non-symbiotic intercalary cells (white arrowhead), and amoebocyte (white asterisk). Scale bar of panel (C): 50μm.

microvilli and cilia on their surfaces, which may help slow down water flow and improve gas exchange (Figure 2B′). A TEM image showed that the bacteriocytes have a compartmentalized internal ultrastructure (Figure 2C). The MOX symbionts (Figure 2D and 2D′), which have the characteristic intracytoplasmic membrane structure of type I methanotroph bacteria (Kalyuzhnaya et al., 2019), are engulfed in intracellular vacuoles, aggregated at the apical top of the bacteriocytes (Figures 2C and 2D). Primary and secondary lysosomes are located in the lower part of the bacteriocytes as well (Figure 2C).

The results of FISH and electron microscopy analyses have indicated the complexity of the *G. platifrons* gill. The tip, ciliary disk, and basal area of the gill filaments include a range of free symbiont cells. On the lateral face of the *G. platifrons* gill filament, the bacteriocytes and the intercalary cells appeared in a staggered pattern. We estimate that the bacteriocytes amount to less than half of the cell population of the *G. platifrons* gill.

Bacteriocytes enrichment and RNA-seq

The overall experimental procedure is given in Figure 3A. A trypsin-EDTA solution was used to rapidly detach the cells from the gill filaments. Most of the collected cells (about 70%–80%) were bacteriocytes (Figure 3B). As shown in Figures 3C and 3D, high-quality total RNA was extracted from both the enriched bacteriocytes (hereafter designated as EB samples) and the whole-gill homogenates (hereafter designated as WG samples). Noticeably, four ribosomal RNA peaks, namely, 16S and 23S ribosomal RNA from the symbionts and 18S and 28S from the host cells, were detected in the EB sample (Figure 3C). On the other hand, in the WG sample, only 18S and 28S ribosomal RNA from the host mussel were detected (Figure 3D). These results show that a significantly higher portion of symbiont RNA was present in the EB samples, suggesting successful bacteriocyte enrichment.

Nine RNA-seq libraries, namely, three from EB samples, five from WG samples, and one metatranscriptomic sample, were sequenced (Table S1). The clean reads of all nine datasets, including eight RNA-seq datasets and one metatranscriptomic dataset, were submitted to the NCBI SRA database (Bioproject: PRJNA645703).

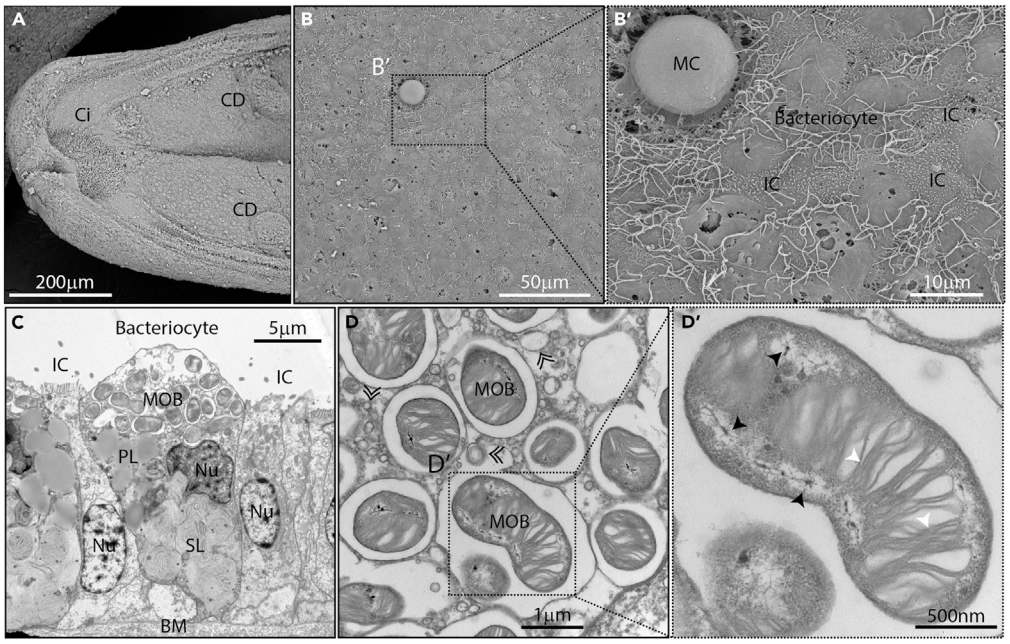


Figure 2. Electron microscopy analysis of the gill of *G. Platifrons*

(A) An SEM (scanning electron microscope) image of the heavily ciliated tip of the *G. platifrons* gill tip. Ci: ciliary, CD: ciliary disk.

(B) An SEM image of the lateral face of the *G. platifrons* gill, showing the bacteriocytes and intercalary cells. (B′) A magnified region in B showing the detailed surface structures of the bacteriocyte and intercalary cells. The surfaces of bacteriocytes are protruding slightly higher than the surrounding intercalary cells. There are cilia and micro villa on the surface of intercalary cells. MC: mucus cell, IC: intercalary cell.

(C) A TEM (transmission electron microscope) image of the longitudinal section of the bacteriocytes and adjacent intercalary cells.

(D) A TEM image of methane-oxidizing symbiotic bacteria covered in vacuoles. (D′) A magnified image showing the detailed structure of the symbiont. White arrowheads: intracytoplasmic membrane structure, which is a characteristic feature of type I methanotroph bacteria. Black arrowheads: DNA of the symbiotic bacteria.

Comparative transcriptomic analysis of the WG and EB samples

After quality control, the clean reads from the WG and EB libraries were aligned to *G. platifrons* transcript models (Sun et al., 2017a). On average, 45.92% of the WG reads and 39.25% of the EB reads were mapped to the *G. platifrons* transcripts' models. Comparative transcriptomic analyses were conducted to identify the DEGs between the WG and EB samples. The overall results of the DEG analyses are summarized in Figures 4A and 4B. Among the 22,494 transcripts, 709 (3.28% of all detected transcripts) were significantly enriched in the WG samples, and 599 (2.77% of all detected transcripts) were significantly enriched in the EB samples.

Gene Ontology (GO) analyses of the DEGs were conducted to provide an overview of the genes and pathways enriched in the EB and WG samples (Figures 4C and 4D, Table S2). Among the top 10 GO terms that were enriched in the EB samples (Figure 4C), there were 6 GO-molecular function (MF) terms—"GTP binding" (GO:0005525), "GTPase activity" (GO:0003924), "Structural constituent of cytoskeleton" (GO:0005200), "Sequence-specific DNA binding" (GO:0043465), "Potassium ion transport" (GO:0006813), and "Voltage-gated ion channel activity" (GO:0005244); 1 GO-cellular component (CC) term, "Microtubule" (GO:0005874) —and 2 GO-biological processes (BP) terms— "Protein polymerization" (GO:0051258) and "Microtubule-based process" (GO:0007017) —that were significantly enriched (Figure 4C). The enriched GO terms, such as "Potassium ion transport" and "Voltage-gated ion channel activity," clearly suggest the presence of transport activities between the bacteriocytes and non-symbiotic cells. Contrary to expectation, 6 of the top 10 enhanced GO terms in the EB samples were involved in microtubule activities (Figure 4C). Many of the DEGs in these GO terms encode a variety of tubulins (Table S2). These results suggest that microtubule activities are significantly enhanced in the bacteriocytes.

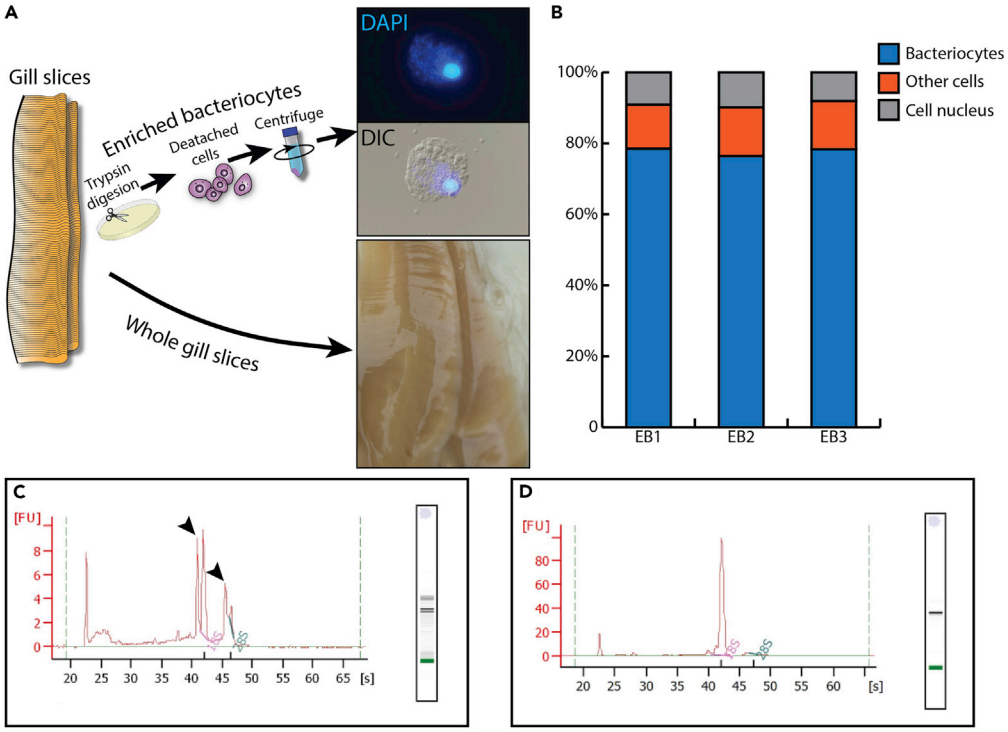


Figure 3. Enrichment of the bacteriocytes

(A) A schematic drawing of the overall experimental procedural. The freshly collected *G. platifrons* gills were treated with Trypsin-EDTA, then collected by centrifugation. DIC: differential interference contrast microscopy image of bacteriocyte; DAPI: 4′,6-diamidino-2-phenylindole stained bacteriocyte.

(B) The stacked bar chart showing the percentage of major cell types in the enriched bacteriocytes (for each sample, average of 3 counts).

(C) Agilent 2100 analysis of a typical EB sample. Both the 16S and 23S ribosomal RNA from the symbionts as well as the 18S and 28S from the host mussel cells were detected. (Black yellow heads: bacterial ribosomal RNA peaks).

(D) Agilent 2100 analysis of a typical WG sample.

Among the top 10 GO terms enriched in WG samples, there were 6 GO-BP terms, namely, "Small molecule transport" (GO:0006810), "Multicellular organism development" (GO:0007275), "DNA replication initiation" (GO:0006270), "Wnt signaling pathway" (GO:0016055), "Positive regulation of transcription" (GO:0006270), "DNA-templated" (GO:0045893), and "DNA duplex unwinding" (GO:0032508); 2 GO-MF terms including "Transporter activity" (GO:0005215) and "Chromatin binding" (GO:0003682); and 2 GO-CC terms including "Extracellular space" (GO:0005615) and "Extracellular region" (GO:0005576). These WG samples with enriched GO terms align well with the non-symbiotic part of the cellular activities of the gill. The genes enriched in "Small molecule transport" and "Transporter activity" indicate that the gill's non-symbiotic part actively shuffles or exchanges nutrients and metabolites with the bacteriocytes. In addition, several GO terms that were related to DNA-replication initiation, including genes involved in controlling cell proliferation, such as DNA-replication licensing factors, proliferating cell nuclear antigen, and eukaryotic translation initiation factor, were enriched in WG samples (Table S2). Finally, two terms related to the extracellular matrix, "Extracellular space" and "Extracellular region," were significantly enriched in WG samples, which correlate with the gill's strong secretion activities.

Furthermore, the WG- and EB-enriched DEGs provided valuable information to distinguish symbiosis-related genes from the genes involved in the background functions of the gill. Therefore, we analyzed the WG- and EB-enriched DEGs and highlighted the ones related to maintaining and regulating the symbiosis. Specifically, this involved several groups of genes, including those involved in supporting the symbiosis (such as those encoding Collagens, Figure S1; myosin, Figure S2; and C1qs Figure S3); genes involved in symbiosis-related cellular activities (such as those encoding carbonic anhydrase, Figure S4; inhibitor of apoptosis, Figure S5; and FMRFamide-like receptors, Figure S6); genes involved in immune recognition of the symbiont (such as those encoding peptidoglycan recognition proteins, Figure S7; toll-like receptors, Figure S8; C-type lectins, Figure S9; and

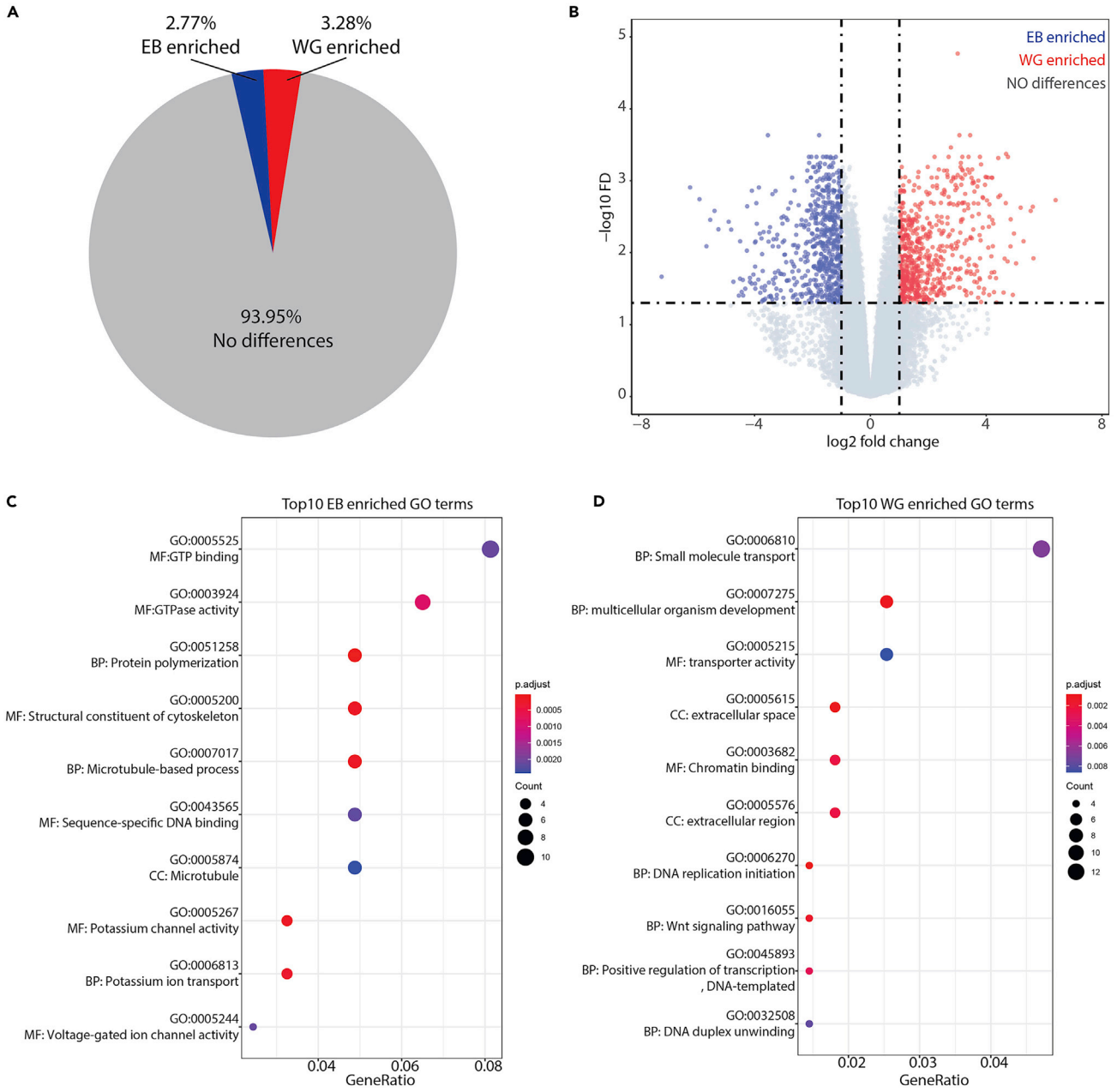


Figure 4. Comparative transcriptomic analysis of the EB and the WG samples

(A) Percentage of the transcripts that were enriched in WG (red) and EB (blue) samples.

(B) Volcano plots for the gene expression profiles in panel (A) of this figure, showing the false discovery rate ($-\log_{10}$FDR) as a function of $\log_2$ fold change.

(C and D) GO enrichment analysis of differentially expressed genes. GO terms that exhibit statistically significant differences are shown in the graph. (C) GO terms enriched in WG samples; (D) GO terms enriched in EB samples.

low-density lipoprotein receptors, Figure S10); genes involved in antiviral activities (such as those encoding tripartite motif, containing 56 proteins, Figure S11); and genes involved in immobilizing the symbiont (such as those encoding caveolin-1, Figure S12; and fibrinogens, Figure S13).

Symbiont genes expression profile analysis

A metatranscriptome library was also sequenced from the pooled EB sample to profile the symbiont's gene expression. About 80 million reads were generated from the metatranscriptomic library. About 41 million

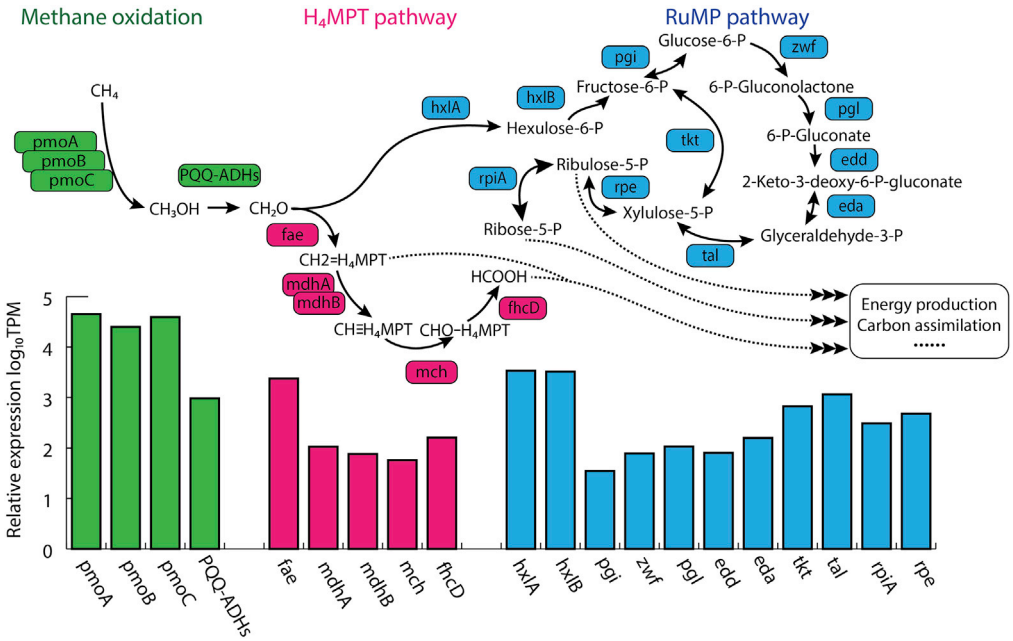


Figure 5. An overview of the symbiont methane oxidation and down-stream H_4MPT and RuMP pathways

The histogram at the bottom shows the relative gene expression levels (log_{10}TPM) of enzymes of these three pathways. Transcripts encoding enzymes in these three metabolic processes dominate the transcriptome of the *G. platifrons* symbiont. *pmoA*: gene15563 methane monooxygenase subunit A; *pmoB*: gene1562 methane monooxygenase subunit B; *pmoC*: gene1564 methane monooxygenase subunit C; *PQQ-ADHs*: gene727 PQQ-dependent dehydrogenase; *fae*: gene2129 5,6,7,8-tetrahydromethanopterin hydro-lyase; *mdhA*: gene 2604 methylenetetrahydromethanopterin dehydrogenase; *mdhB*: gene2605 methylenetetrahydromethanopterin dehydrogenase; *mch*: gene2242 methenyltetrahydromethanopterin cyclohydrolase; *fhcD*: gene337 formylmethanofuran–tetrahydromethanopterin formyltransferase; *hxlA*: gene758 3-hexulose-6-phosphate synthase; *hxlB*: gene757 3-hexulose-6-phosphate isomerase; *pgi*: gene2681 glucose-6-phosphate isomerase; *zwf*: gene1697 glucose-6-phosphate dehydrogenase; *edd*: gene2951 phosphogluconate dehydratase; *eda*: gene2950 bifunctional 4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxy-phosphogluconate aldolase; *tkt*: gene756 transketolase; *tal*: gene753 transaldolase; *rpiA*: gene2174 ribose-5-phosphate isomerase; *rpe*: gene2173 ribulose-phosphate 3-epimerase.

clean metatranscriptomic reads were mapped to the symbiont transcript models (Takishita et al., 2017). Approximately 11 million reads were mapped to protein-coding genes. Among the 4,136 predicted genes, 2,901 had a TPM value greater than 10 (Data S2). The metatranscriptomic dataset showed that the methane oxidizing process, which generates formaldehyde, is the most prominent metabolic process in the symbiont (Figure 5). The three particulate methane monooxygenases that conduct the first step of the reaction by converting methane to methanol were the most highly expressed protein-coding genes. Together, *pmoC*, *pmoB*, and *pmoA* constituted about 22.10% of the total symbiont protein-coding gene reads. Furthermore, two downstream pathways, the H4MPT (tetrahydromethanopterin) and the RuMP (ribulose monophosphate) pathways, which detoxify formaldehyde and generate energy and biomass, were highly active in the symbiont. Genes encoding enzymes of both pathways, such as *fae*, *fhcD*, *hxlA*, and *hxlB*, are all abundantly expressed (Figure 5). In addition to the genes involved in methane oxidation pathways, several groups of genes, such as those encoding DUF domain-containing proteins, transposases, antitoxins, and subunits of F0F1 ATP synthases, were highly expressed in the symbiont (Table S3).

In situ hybridization analysis

Then we selected several representative WG- and EB-enriched genes and analyzed their spatial expression patterns using *in situ* hybridization. Five WG-enriched genes were chosen for ISH analyses. Bpl_scaf_11,625–1.11, which encodes a fibril-forming collagen alpha chain-like protein, was highly expressed in the ciliated tip of the gill filament (Figures 6A–6C and A′–C′). Another collagen-encoding protein, Bpl_scaf_1314–2.22, is mainly expressed in the mucous cells, suggesting that its protein products are part of the gill's extracellular matrix system (Figures 6D–6F and 6D′–6F′). Two genes that encode DNA

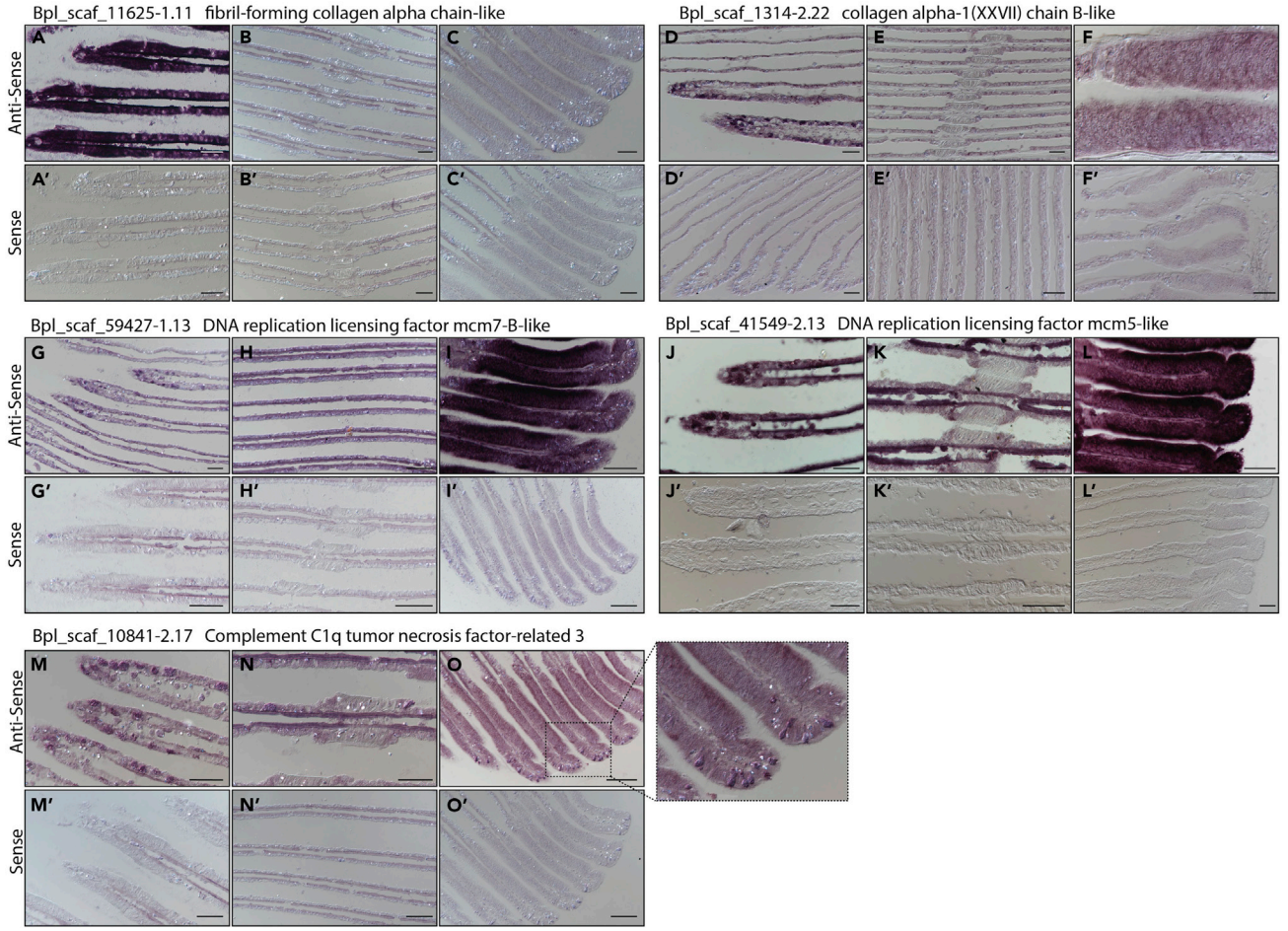


Figure 6. ISH analysis of selected WG-enriched genes

In situ hybridization analysis of selected WG enriched genes, including Bpl_scaf_11,625-1.11 (Fibril-forming collagen alpha chain-like protein); Bpl_scaf_1314–2.22 (collagen); Bpl_scaf_59,427–1.13 DNA replication licensing factor mcm7-B-like protein; Bpl_scaf_41,549–2.15 DNA replication licensing factor mcm7-B-like protein; Bpl_scaf_41,549–2.15 DNA replication licensing factor mcm5-like protein; Bpl_scaf_10841–2.17 Complement C1q tumor necrosis factor-related 3. Panel (A–O) are showing the *in situ* hybridization results by using anti-sense probes. Panel (A, D, G, J, and M) are showing the tip of the gill filaments, panel (B, E, H, K, and N) are showing the mid part of the gill filaments, and panel (C, F, I, L, and O) are showing the basal region of the gill filaments. Panels (A′–O′) are showing the *in situ* hybridization results by using control sense probes.

replicate licensing factors, Bpl_scaf_59,427–1.13 (DNA replication licensing factor mcm7-B-like protein) and Bpl_scaf_41,549–2.15 (DNA replication licensing factor mcm5-like protein), are both strongly expressed in the ventral-end growth zone, which is composed of cells with a high nucleus-to-cytoplasm ratio (Figures 6G–L and 6G′–6L′). The expression of Bpl_scaf_41,549–2.15 was also detected in the mucous cells (Figure 6J). No visible expression of these two genes was found on the bacteriocytes or intercalary cells. Finally, the ISH results showed that WG-enriched C1qDC (Figures 6M–O and 6M′–6O′) was expressed in the mucous cells, suggesting its protein product might be secreted with the mucus.

The six EB-enriched genes selected for *in situ* analyses were Bpl_scaf_21,642–0.2 (E3 ubiquitin-ligase HERC2), Bpl_scaf_54,034–5.14 (ubiquitin-conjugating enzyme E2 D3 isoform X2), Bpl_scaf_18,519–2.15 (fibrinogen C domain-containing (1), Bpl_scaf_19,969–0.58 (tubulin alpha-1A chain), and Bpl_scaf_33,659–0.9 (low-density lipo receptor-related 2). All of the EB-enriched genes showed similar expression patterns. As presented in Figure 7, on the gill filaments, the expression of the EB-enriched genes was detected in the bacteriocytes but not in the intercalary cells. At the cilia disk, the tip, and the basal area of the gill filaments, expression of these genes was not detectable. One exception to this is Bpl_scaf_33,659–0.9 (low-density lipo receptor-related (2), which was detected in two parallel clusters of cells at the tip of each gill filament (Figure 7P).

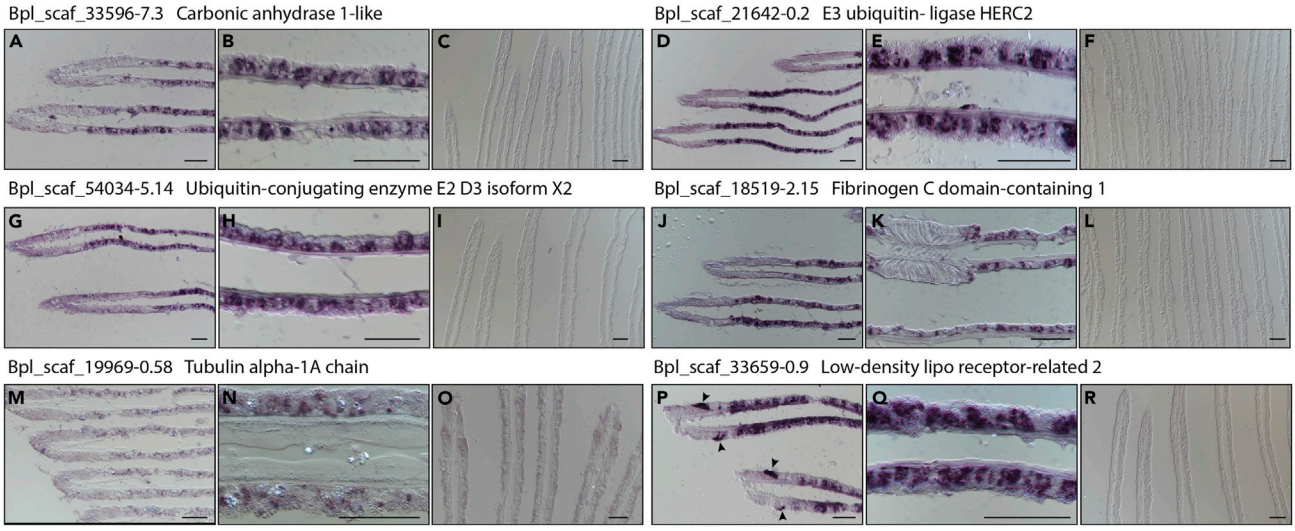


Figure 7. ISH analysis of selected EB-enriched genes

In situ hybridization analysis of selected EB-enriched genes, including Bpl_scaf_21642–0.2 (E3 ubiquitin-ligase HERC2), Bpl_scaf_54034–5.14 (ubiquitin-conjugating enzyme E2 D3 isoform X2), Bpl_scaf_18519–2.15 (fibrinogen C domain-containing 1), Bpl_scaf_19,969–0.58 (tubulin alpha-1A chain), and Bpl_scaf_33659–0.9 (low-density lipo receptor-related 2). The *in situ* hybridization results of corresponding genes by using anti-sense probes are shown in panel (A, B, D, E, G, H, J, K, M, and N). Panel (A, D, G, J, M, and P) are showing the tip of the gill filaments, whereas panel (B, E, H, K, N, and O) are showing the mid part of the gill filaments. The *in situ* hybridization results by using control sense probes are shown in panel (C, F, I, L, O, R). Black arrowhead in P: two clusters of symbiont free cells expressing Bpl_scaf_33,659–0.9 at the ciliated tip of the gill. Scale bar: 50μm.

Immunohistochemistry analysis

As shown in the GO analyses, the most significantly enriched EB genes are those encoding alpha and beta tubulins (Table S2 and Figure 8A). Therefore, immunohistochemistry (IHC) analyses were conducted to analyze the subcellular distribution of the alpha and beta tubulin proteins. As shown in Figure 8B, the anti-alpha-tubulin polyclonal antibody produced a strong immunofluorescence signal at the apical top of each bacteriocyte, where the symbionts are located. No visible subcellular distribution of alpha-tubulin was observed in the intercalary cells. Beta-actin antibody produced a similar result in the bacteriocytes, and a fluorescent signal was detected around the symbiont vacuoles (Figure 8C). The beta-tubulin antibody detected the cilia and microvilli on the intercalary cells (Figure 8C).

DISCUSSION

The transcriptomic dataset produced by this study provides a comprehensive gene-expression profile for the three major parts of *G. platifrons* gill symbiosis: the non-symbiotic parts of the gill, the bacteriocytes, and the symbiont. This, together with FISH, SEM/TEM, ISH, and IHC analyses, allowed us to understand the contribution of each part to the symbiosis for the first time and to interpret the interaction between host and symbiont at the molecular level.

Significant WG-enriched genes and functions of the non-symbiotic parts of the gill

The DEGs that were enriched in the WG samples were well aligned with the cellular function of the non-symbiotic parts of *G. platifrons* gill. For example, the bathymodiolin mussel gill is composed of hundreds of filaments that are loosely held together by the ciliary system (Wentrup et al., 2014). There are 20 genes that encode different types of collagens enriched in the WG samples (Figure S1). The ISH results showed that Bpl_scaf_11,625–1.11, which encodes a fibril-forming collagen alpha chain-like protein, was highly expressed in the ciliated tip of the gill filament (Figure 6A). Fibril-forming collagen is an adhesive protein that maintains the integrity of the extracellular matrix (Ricard-Blum and Ruggiero, 2005), suggesting that it is involved in maintaining the gill's spatial organization. The gill filaments also beat and contract to ventilate (Jørgensen, 1974), providing the symbiotic bacteria with gas from the vent or seep fluid. Thus, the genes encoding myosin, which are the basic building blocks of muscle fibers (Pette and Staron, 2000), are significantly enriched in WG samples (Figure S2).

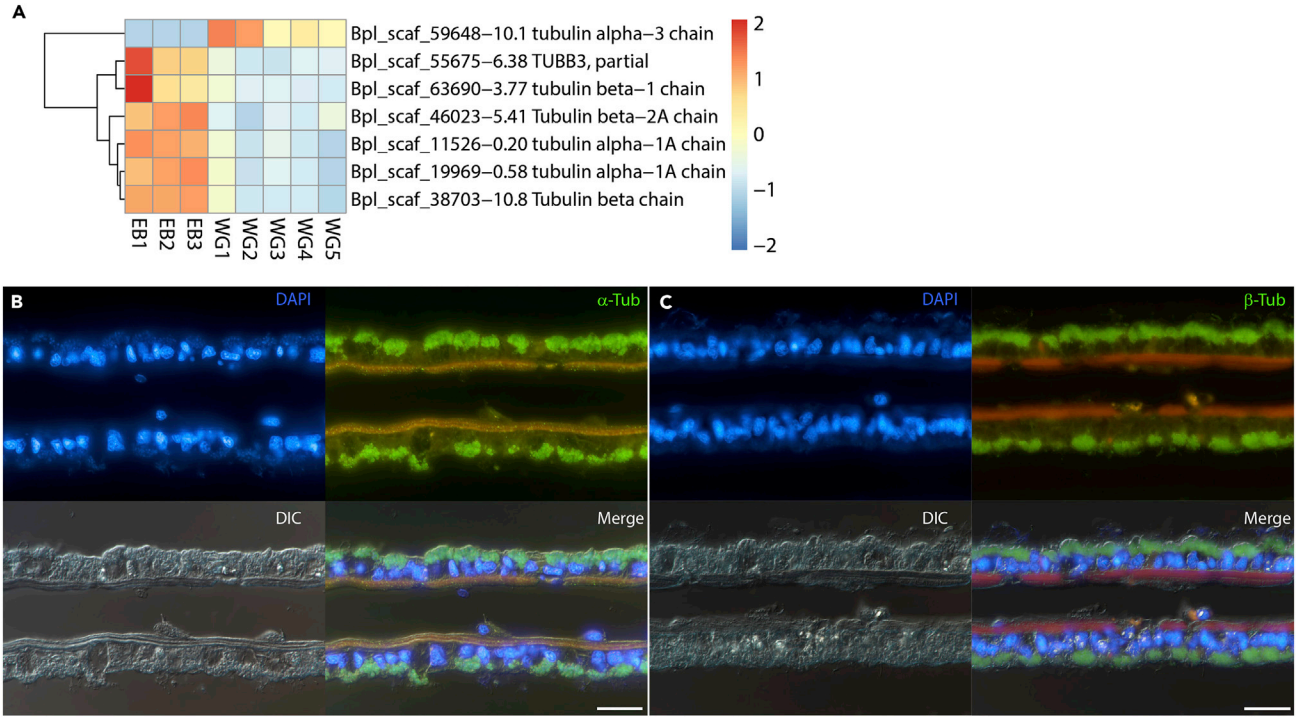


Figure 8. IHC analysis of the alpha and beta tubulins

(A) A heatmap of differentially expressed alpha and beta tubulins between EB and WG samples. The hierarchical cluster shown here was obtained by comparing the expression values of two sample groups. The blue-red scale represents the relative expression values.

(B) The IHC analysis of the protein expression pattern of alpha tubulin.

(C) The IHC analysis of the protein expression pattern of beta tubulin. Scale bar: 50μm.

As shown in Figure 4C, the most significantly enriched GO terms in the WG samples were those involved in regulating DNA replication (Figure 4D). Key genes that play crucial roles in regulating cell proliferation, such as DNA-replication licensing, proliferating cell nuclear antigen, and eukaryotic translation initiation factor, were all enriched in WG samples (Table S2). The gill filaments of adult bathymodiolin mussels show growth zones at both the ventral and dorsal ends, which proliferate throughout the life cycle of the mussel (Wentrup et al., 2014; Neumann and Kappes, 2003). Symbiont-free cells that are freshly divided from growth zones can be gradually colonized by symbionts from the surrounding bacteriocytes (Wentrup et al., 2014). It is noteworthy, as shown in ISH analyses, that the strong expression of these DNA replication-associated genes can be detected in both the growth zones and the mucous cells but not in bacteriocytes or intercalary cells. These results suggest that the recruitment of bacteriocytes is done mainly through the colonization of symbiont-free cells generated from the growth zones but not through bacteriocyte proliferation.

DEG and GO analyses also revealed that the extracellular matrix-related genes were enriched in the WG samples. The ciliated tip and lateral edge of the deep-sea bivalve's gill filaments are often covered with mucus secreted by the mucous cells (Dufour and Beninger, 2001). In shallow-water filter-feeding bivalves, the mucus functions cooperatively with the cilia to capture, process, and transport food particles to the mouth (Dufour and Beninger, 2001; Gómez-Mendikute et al., 2005; Beninger and Dufour, 1996). In deep-sea symbiotic bivalves with vestigial digestive systems, it is unlikely that mucus is involved in feeding, but it acts as a defensive barrier and/or for the maintenance of the symbiotic bacteria (Nakamura et al., 2013). Interestingly, the gill of *G. platifrons* expresses a large variety of C1q domain-containing proteins (C1qDCs), including C1q domain-containing protein, complement C1q, and complement C1q tumor-necrosis-factor-related protein, many of which are enriched in WG samples (Figure S3). In invertebrates, C1qDCs are immune molecules that bind to surface molecules of a pathogen and trigger downstream immune activities but also eliminate bacteria directly (Jiang et al., 2015; Carland and Gerwick, 2010). Unexpectedly, the *in situ* hybridization analyses showed that one WG enriched C1qDC was expressed in the

mucous cells (Figures 6M and 6O), which further indicated that the mucus of *G. platifrons* is involved in protecting the symbionts against environmental microbes.

Altogether, these results collectively show that the non-symbiotic parts of the gill play a crucial supporting role in the maintenance of symbiosis in *G. platifrons*.

Significant EB-enriched genes

Genes involved in EB cellular activates

EB samples provide valuable information for distinguishing the genes involved in host-symbiont metabolic interactions from the gill's background functionalities. For example, carbonic anhydrases (CAs) are a group of zinc-containing metalloproteins that catalyze the interconversion of carbon dioxide (CO_2) and carbonic acid (H_2CO_3) (Aspatwar et al., 2014). CAs are found in almost all animal tissues that are involved in respiration, photosynthesis, or biomineralization (Tashian, 1989). They are highly expressed in the gill of deep-sea symbiotic bivalves, suggesting that they play a crucial role in maintaining symbiosis (Ponnudurai et al., 2020; Hongo et al., 2013). Among those expressed in the gill of *G. platifrons*, two were enriched in the EB samples (Figure S4). It may be that these bacteriocyte-enriched CAs are involved in removing the CO_2 released by the methanotrophic symbiont as an end product of methane oxidation (Ponnudurai et al., 2020). The inhibitor of apoptosis proteins (IAPs) belongs to a family of anti-apoptotic proteins that promote cell survival (Silke and Meier, 2013). The whole-genome sequencing of *G. platifrons* showed that IAP genes are greatly expanded in this species compared with other lophotrochozoans, suggesting that they are involved in the process of adapting to symbiosis (Sun et al., 2017a). The non-symbiotic parts of the gill, including the ciliated cells and the hemocytes, has a higher rate of apoptosis than the bacteriocytes (Piquet et al., 2019). In this study, DEG analyses also showed that the genes encoding apoptosis inhibitors are enriched in the EB samples, suggesting that the IAPs contribute to the population control of the symbiont (Figure S5). Interestingly, several genes encoding potential FMRFamide-like receptors were also enriched in the EB samples (Figure S6). FMRFamide-like peptides (FLPs), which were first identified in the clam *Macrocallista nimbosa*, are the largest and most diverse family of neuropeptides known (Peymen et al., 2014; Zatylny-Gaudin and Favrel, 2014; Price and Greenberg, 1977). FLPs have pleiotropic activities and mediate a variety of physiological and behavioral processes, such as reproduction, feeding, and the regulation of energy balance and metabolism (Lee et al., 2004; Bigot et al., 2014). Interestingly, schistosome parasites can adjust the physiology and behavior of their intermediate molluscan hosts to their own benefit by altering the expression of regulatory neuropeptides (Hoek et al., 2005), suggesting that FLPs and their receptors serve as a cross-species regulatory signal. Among the 15 FMRFamide-like receptors encoding genes detected in our dataset, 3 were enriched in EB samples (Figure S6), indicating their involvement in the metabolic regulation of the bacteriocytes.

EB-enriched immune receptors

It has been long hypothesized that host animals use a variety of immune molecules to control and interact with the symbiont. However, the bivalve gill is also a vital immune organ, as it is continuously in contact with environmental bacteria. Highly expressed immune genes, such as the pattern recognition receptors (PRRs), are commonly found in the gill of non-symbiotic shallow water bivalve species as well (Philipp et al., 2012; Zhang et al., 2015). Comparative analyses of the EB- and WG-enriched immune genes allowed us to identify symbiosis-related immune genes. In the transcriptomic data 10 genes that encode peptidoglycan recognition proteins (PGRPs) were detected, but none of them were differentially expressed between the EB and WG samples (Figure S7). As PGRPs are largely involved in sensing bacterial peptidoglycans on the surface of Gram-positive bacteria, the homogeneous expression of PGRPs suggests that they are likely involved in responding to environmental bacteria but not in regulating the symbiont in *G. platifrons* symbiosis (Dziarski, 2004; Detree et al., 2017). Several groups of PRRs, including toll-like receptors (TLRs), C-type lectins (CTLs), and low-density lipoprotein receptors (LDLRs), were enriched in the EB samples. Three TLRs (Figure S8) and two C-type lectins (Figure S9) were enriched in the EB sample. Both TLRs and CTLs are important PRRs that activate and regulate downstream immune cascades. These EB-upregulated TLRs and CTLs have been reported to be involved in the population control of the symbiont (Zheng et al., 2017; Wong et al., 2015; Sun et al., 2017a).

Five genes that encode LDLRs were abundantly expressed and significantly enriched in the EB samples (Figure S10). The LDLRs are evolutionarily conserved cell-surface receptors that mediate the endocytosis of cholesterol-rich low-density lipoprotein and innate immune responses (Go and Mani, 2012; Liu et al., 2014). We previously showed that the expressions of several EB-upregulated LDLRs, including

Bpl_scarf_33,659–0.13, Bpl_scaf_33,569–0.9, and Bpl_scarf_2971–0.17, which are among the EB-enriched LDLRs in the present study, are suppressed by the symbiont but stimulated by environmental bacteria (Wang et al., 2019). The *in situ* hybridization analyses of the present study showed that EB-upregulated LDLR was not only expressed in the bacteriocytes but also in a cluster of cells in the ciliated tip of the gill, where it is symbiont free. The ciliated tip of the bathymodiolin gill also contains proliferating cells (Wentrup et al., 2014). The newly divided cells that are symbiont free could be colonized by symbionts released by bacteriocytes from neighboring gill filaments (Wentrup et al., 2014). These results indicate that LDLRs could play an essential role in the establishment and maintenance of the symbiosis.

Antiviral genes

The transcriptomic dataset shows that a variety of genes in the invertebrate antiviral toolkit, such as retinoic acid-inducible gene I (RIG-1), stimulator of interferon (STING), and interferon-induced protein 44-like (IFI44L), were expressed in the gill of G. *platifrons*. However, none of these were differentially expressed between the EB and WG samples. Because viruses are abundant in marine environments, the expression of invertebrate antiviral genes suggests that the mussel continuously guards against viral infection. The gill of *G. platifrons* also expresses a large variety of ubiquitin ligases, which are key players in the regulation of innate immune activities, particularly antiviral-signaling cascade (Zheng and Gao, 2019). Interestingly, several genes encoding tripartite motif-containing protein 56 (TRIM56) were upregulated in EB samples (Figure S11). TRIM56 is an E3 ubiquitin-protein ligase that plays a crucial role in innate antiviral immunity (Wang et al., 2011). It has been demonstrated in model animals that TRIM56 positively regulates dsRNA-sensing TLRs. These EB-enriched antiviral genes suggest that the host provides the symbiont with a sheltered optimal-growth condition. This allows the symbiont to reach high population density and avoid the kill-the-winner effect (Thingstad, 2000).

Genes involved in immobilizing the symbiont

Several genes that could be involved in restricting symbiont mobility, such as caveolin-1 (Figure S12) and fibrinogen (Figure S13), were highly expressed and enriched in EB samples. Caveolin-1 plays a crucial role as a structural protein in the formation of the caveolae (Zaas et al., 2009; Liu et al., 2002), which are involved in various biological processes, such as endocytosis, transcytosis, and signal transduction (Zaas et al., 2009). Many bacteria, both pathogenic and viable nonpathogenic, that are not known to process invasive capability, can enter cells by exploiting the endocytotic pathway with cholesterol-rich lipid rafts and caveolin-1 (Zaas et al., 2009). Functional analyses have revealed that intracellular bacteria are also co-localized with caveolin-1, and invasion depends on the expression of caveolin-1 (Yu, 2015). Similarly, invertebrate fibrinogens, such as the tachylectins identified in the horseshoe crab *Tachypleus tridentatus*, possess lectin-like qualities and agglutinate bacteria (Kawabata and Iwanaga, 1999). AiFREP, a fibrinogen-related protein identified in the bay scallop *Argopecten irradians*, can agglutinate chicken and human erythrocytes and those of Gram-negative and Gram-positive bacteria (Yang et al., 2014). In this study, EB-enriched fibrinogens were found to contain a single fibrinogen domain with a signal peptide (Table S4), suggesting that they are likely to be involved in binding within the symbiont's vacuoles.

An unexpected result of the DEG analyses is that the genes encoding cellular skeleton proteins, particularly alpha and beta tubulins, were enriched in the EB samples (Figure 8A). Alpha and beta tubulins are the basic building blocks of microtubules, which are the main tracks used in cells to organize organelle positioning and the cargo traffic (Fourriere et al., 2020). In this study, the results of FISH (Figures 1C and 1D) and electron microscopy (Figure 2C) showed that the symbiont is restrained at the apical top of the bacteriocytes, suggesting that its cytoskeleton is involved in holding symbiont-containing vacuoles in place and maintaining the polarity of the bacteriocytes. This hypothesis is supported by the IHC results (Figures 8B and 8C). This cell polarity has clear benefits for both the symbiont and the host. Because the symbionts are pushed closer to the gas exchange interface, they gain easy access to the necessary gas and therefore can maintain high metabolism. Within the same bacteriocyte, the host mussel maintains a dense population of symbionts and digests a portion of them to extract nutrients. This polarized and compartmentalized subcellular organization allows the host mussel to protect the symbiont from its own immune system, avoiding wiping out its whole population.

Gene expression profile of the symbiont

Methane oxidation and energy production

The evolutionary success of the association between the *G. platifrons* and the MOX symbiont is clear (Figure 1A). These mussels dominate the biomass of their chemosynthesis-based communities, and it mainly depends upon

the primary productivity of the symbionts (Feng et al., 2015; Van Dover and Fry, 1989). The metatranscriptomic dataset showed that the MOX process, which generates formaldehyde, is the most prominent metabolic process in the symbiont. The three particulate methane monooxygenases, which conduct the first step of the reaction by converting methane to methanol, were the most highly expressed protein-coding genes. Furthermore, two downstream pathways, the tetrahydromethanopterin (H_4MPT) pathway and the ribulose monophosphate (RuMP) pathway, which detoxify formaldehyde and generate energy as well as biomass, were highly active in the symbiont (Figure 5). The genes that encode subunits of both the F1 and F0 motors of F-type ATPase were all highly expressed (Table S3), suggesting the active production of ATP by the symbiont. ATP molecules, which serve as the universal energy currency of life, could also be utilized by the host, either by milking (translocation of nutrient from symbiont tissue to the host cells) or by farming (direct digestion of symbionts by the host) (Ponnudurai et al., 2017a; Ponnudurai et al., 2017b; Streams et al., 1997). Currently, the best-known symbiotic MOX bacteria from marine invertebrates are found within a single lineage in the Gammaproteobacteria and are related to type I methanotrophs (Jillian and Nicole, 2009), with the exception of the recently discovered fan worm-type II methanotroph association (Goffredi et al., 2020). This preference for type I methanotrophs could be because the RuMP pathway used by type I methanotrophs requires less energy and therefore, is more efficient than the serine pathway used by type II methanotrophs (Jillian and Nicole, 2009). The metatranscriptomic dataset of this study shows that MOX and two downstream detoxification/energy production pathways dominate the *G. platifrons* symbiont transcriptome, providing evidence at the transcription level that supports the efficiency-driven hypothesis.

Interaction between symbiont and host

In many symbioses, the symbionts act as pathogens in their interactions with their hosts. For example, the symbiont of the cold-seep Siboglinidae tubeworm *Paraescarpia echinospica* abundantly expresses secreted esterase (PSE), which plays a crucial role in bacterial virulence and pathogenesis and may aid in the digestion of the host animals' cells (Yang et al., 2020). In the sulfur-oxidizing (SOX) symbiont of the deep-sea mussel *B. azoricus*, up to 7.6% of the genome represents toxins or virulence genes (Sayavedra et al., 2015). By contrast, no high percentage of toxin or virulence genes has been found in clam symbionts or closely related free-living SOX bacteria (Newton et al., 2008). We found that the *G. platifrons* symbiont expresses outer membrane protein A (OmpA), which plays a critical pathogenic role in bacterial adhesion and invasion (Pines and Inouye, 1999). However, highly expressed toxin or virulence genes were not detected in the metatranscriptomic dataset, and genes encoding antitoxins that counteract bacterial toxins (Riffaud et al., 2020) were abundantly expressed in the symbiont (Table S3).

Genes encoding preprotein translocase subunit SecY (Bieker et al., 1990) and its regulatory protease htpX (Sakoh et al., 2005) were highly expressed in the *G. platifrons* symbiont. SecY is the central subunit of the protein translocation channel SecYEG, which plays a crucial role in moving proteins across cell membranes (Veenendaal et al., 2004). The expression of secY indicates an active exchange of protein products between the symbiont and mussel host. Interestingly, proteins in the DUF domain, particularly the group of DUF1566 domain-containing proteins, were abundantly expressed in the symbiont (Figure S4). Although the function of the DUF1566 domain is still largely unknown, most of these proteins contain a signal peptide (Table S5), suggesting that they are secreted proteins that travel from the symbiont to the host bacteriocyte.

Transposases

Unexpectedly, metatranscriptomic data showed that the genes encoding a variety of transposases were highly expressed in the symbiont (Table S3). Transposase genes, together with some short flanking sequences necessary for transposition, are transposable elements (TEs) (Aziz et al., 2010). TEs are considered to be parasite DNA elements with no purpose other than self-reproduction (Doolittle and Sapienza, 1980). However, accumulating data suggest that TEs are involved in gene duplication, genome rearrangement, gene regulation, and other activities, all of which could generate genomic diversity and therefore are beneficial to the host bacteria (Touchon and Rocha, 2007; Werren, 2011). TEs and thus transposase genes are enriched in the genomes of mutualistic symbionts and some pathogens that have recently transited or are transitioning to an obligate host-associated life (Kleiner et al., 2013; Schmitz-Esser et al., 2011). It has been proposed that TEs can help relax the purification of the selection on the symbiont genome, as the effective population size of the symbiont may decrease during transmission to obligate intracellular lifestyle (Newton and Bordenstein, 2011).

TEs are absent in the reduced genome of most obligate, host-restricted symbionts with long-term association with its host (Sakoh et al., 2005). It has been well documented that bathymodiolin mussels acquire

their symbionts horizontally from the environment during the early juvenile stage (Dubilier et al., 2008). However, the high expression of transposases indicates that the symbiont of *G. platifrons* might be transitioning to an obligate host-associated life and therefore experiencing a genetic bottleneck. Overall, these data collectively suggest that the symbiont has a mutualistic relationship closely associated with its host and has less frequent dispersal.

Conclusion

In this work, we conducted integrated molecular analyses of the gill tissue of the deep-sea symbiotic mussel *G. platifrons*. Our data showed that the non-symbiotic parts of the *G. platifrons* gill express the genes essential to regulating the gill's spatial structure and movement, which ventilates seep water to provide the symbiont with necessary gases. The gill secretes mucus that contains immune molecules to eliminate environmental microbes. The gill also has actively proliferating growth zones that potentially supplement the bacteriocyte population. At the cellular level, our data suggest that bacteriocytes supply the symbiont with necessary metabolites, regulate the symbiont population with the innate immune system, and protect the symbiont from phage with its antiviral toolkits. We also found that the bacteriocytes immobilize and separate the symbiont from the remaining parts of the cell, which may help it reach the gases in the seep fluid and allow the host to simultaneously milk and farm it. Finally, metatranscriptomic data for the symbiont show that MOX and downstream energy-generating pathways dominate the symbiont's metabolism, suggesting that it benefits from an ideal micro niche created by the host. Although the symbiont expresses several genes that could disarm the host's immune molecules, it behaves as a mutualistic symbiont, not a pathogen.

The results presented in this study demonstrate that different parts of *G. platifrons* symbiosis work together at the tissue, cellular, and molecular levels, maintaining high efficiency and a harmonic, chemosynthetic micro niche. Noticeably, although gill symbiosis is the essential evolutionary innovation of *G. platifrons* relative to its shallow-water homolog species, the symbiosis-regulating genes identified in this study are mostly conserved genes that have conserved functions across the animal kingdom. These results indicate that genes' spatial expression is also crucial for the adaptation and evolution of symbiosis in animals.

Limitations of the study

In this work, we identified the mussel host's symbiotic and non-symbiotic genes with a comparative transcriptomic approach. However, some genes involved in host-symbiont metabolic interchange and regulation may not be bacteriocyte specific or differentially expressed and thus may have evaded detection in our DEG analyses. Therefore, we did not attempt to identify host-symbiont metabolic exchange networks or regulator pathways in the present work. Our results demonstrate the complexity of the gill filaments of *G. platifrons*. Although we have identified a few gene markers that could represent major cell types in the gill filament, this is not sufficient to fully comprehend its cellular functionalities, particularly in the non-symbiotic parts. With the rapid development and advances in RNA-seq and deep-sea sampling technologies, we look forward to tackling these limitations in our future deep-sea cruises.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Chaolun Li (lcl@qdio.ac.cn).

Materials availability

This study did not generate new unique reagents.

Data and code accessibility

The data have been deposited with links to BioProject Accession no. PRJNA645703 in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/).

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101894.

ACKNOWLEDGMENTS

This study was supported by the National Key R&D Program of China (Project Number 2018YFC0310800), the Scientific and Technological Innovation Project (No.2016ASKJ14) from Qingdao National Laboratory for Marine Science and Technology, Strategic Priority Research Program of the Chinese Academy of Sciences (XDB42020401 and XDA22050303), National Natural Science Foundation of China (Grant No. 41806200), and the Senior User Project of RV KEXUE (KEXUE2016G06). We appreciate all the assistance provided by the crews on RV Kexue.

AUTHOR CONTRIBUTIONS

W.H., Z.H., and L.C.L designed the study; Z.Z.S. and L.C collected the samples; W.H., Z.H., Z.Z.S., and S.Y. finished the laboratory work. W.H., W.M.X., C.H., Z.L., and C.L. performed the bioinformatics analysis. W.H., Z.H., and L.C.L drafted the manuscripts. The manuscript was revised by all of the authors. All authors read, approved, and contributed to the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 14, 2020
Revised: October 7, 2020
Accepted: December 2, 2020
Published: January 22, 2021

REFERENCES

Aspatwar, A., Tolvanen, M.E., Ortutay, C., and Parkkila, S. (2014). Carbonic anhydrase related proteins: molecular biology and evolution. Subcell. Biochem. *75*, 135–156.

Aziz, R.K., Breitbart, M., and Edwards, R.A. (2010). Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Res. *38*, 4207–4217.

Bang, C., Dagan, T., Deines, P., Dubilier, N., Duschl, W.J., Fraune, S., Hentschel, U., Hirt, H., Hulter, N., Lachnit, T., et al. (2018). Metaorganisms in Extreme Environments: Do Microbes Play a Role in Organismal Adaptation? Zoology (Jena) *127*, 1–19.

Barros, I., Divya, B., Martins, I., Vandeperre, F., Santos, R.S., and Bettencourt, R. (2015). Post-capture immune gene expression studies in the deep-sea hydrothermal vent mussel Bathymodiolus azoricus acclimatized to atmospheric pressure. Fish Shellfish Immunol. *42*, 159–170.

Barry, J.P., Buck, K.R., Kochevar, R.K., Nelson, D.C., Fujiwara, Y., Goffredi, S.K., and Hashimoto, J. (2002). Methane-based symbiosis in a mussel, Bathymodiolus platifrons, from cold seeps in Sagami Bay, Japan. Invertebr. Biol. *121*, 47–54.

Beninger, P.G., and Dufour, S. (1996). Mucocyte distribution and relationship to particle transport on the pseudolamellibranch gill of Crassostrea virginica (Bivalvia:Ostreidae). Mar. Ecol. Prog. Ser. *137*, 133–138.

Bettencourt, R., Barros, I., Martins, E., Martins, I., Cerqueira, T., Colaco, A., Costa, V., Rosa, D., Froufe, H., Egas, C., Stefanni, S., Dando, P., and Santos, R. (2017). An insightful model to study innate immunity and stress response in deep-sea vent animals: profiling the mussel Bathymodiolus azoricus. Organismal and Molecular Malacology *8*, 161–187.

Bettencourt, R., Dando, P., Collins, P., Costa, V., Allam, B., and Serrao Santos, R. (2009). Innate immunity in the deep sea hydrothermal vent mussel Bathymodiolus azoricus. Comp. Biochem. Physiol. A Mol. Integr. Physiol. *152*, 278–289.

Bettencourt, R., Pinheiro, M., Egas, C., Gomes, P., Afonso, M., Shank, T., and Santos, R.S. (2010). High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel Bathymodiolus azoricus. BMC Genomics *11*, 559.

Bieker, K.L., Phillips, G.J., and Silhavy, T.J. (1990). The sec and prl genes ofEscherichia coli. J. Bioenerg. Biomembranes *22*, 291–310.

Bigot, L., Beets, I., Dubos, M.P., Boudry, P., Schoofs, L., and Favrel, P. (2014). Functional characterization of a short neuropeptide F-related receptor in a lophotrochozoan, the mollusk Crassostrea gigas. J. Exp. Biol. *217*, 2974–2982.

Carland, T.M., and Gerwick, L. (2010). The C1q domain containing proteins: where do they come from and what do they do? Dev. Comp. Immunol. *34*, 785–790.

Detree, C., Lallier, F.H., Tanguy, A., and Mary, J. (2017). Identification and gene expression of multiple peptidoglycan recognition proteins (PGRPs) in the deep-sea mussel Bathymodiolus azoricus, involvement in symbiosis? Comp. Biochem. Physiol. B Biochem. Mol. Biol. *207*, 1–8.

Doolittle, W.F., and Sapienza, C. (1980). Selfish genes, the phenotype paradigm and genome evolution. Nature *284*, 601–603.

Dubilier, N., Bergin, C., and Lott, C. (2008). Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. Nat. Rev. Microbiol. *6*, 725.

Dufour, S.C., and Beninger, P.G. (2001). A functional interpretation of cilia and mucocyte distributions on the abfrontal surface of bivalve gills. Mar. Biol. *138*, 295–309.

Dziarski, R. (2004). Peptidoglycan recognition proteins (PGRPs). Mol. Immunol. *40*, 877–886.

Feng, D., Cheng, M., Kiel, S., Qiu, J.-W., Yang, Q., Zhou, H., Peng, Y., and Chen, D. (2015). Using Bathymodiolus tissue stable carbon, nitrogen and sulfur isotopes to infer biogeochemical process at a cold seep in the South China Sea. Deep Sea Res. Oceanogr. Res. Pap. *104*, 52–59.

Fiala-Médioni, A., Métivier, C., Herry, A., and Le Pennec, M. (1986). Ultrastructure of the Gill of the Hydrothermal-Vent Mytilid Bathymodiolus Sp*92* (Marine Biology), pp. 65–72.

Finlay, B.B., and Medzhitov, R. (2007). Host-microbe interactions: fulfilling a niche. Cell Host Microbe *1*, 3–4.

Fourriere, L., Jimenez, A.J., Perez, F., and Boncompain, G. (2020). The role of microtubules in secretory protein transport. J. Cell Sci *133*, jcs237016.

Fujiwara, Y., Takai, K., Uematsu, K., Tsuchida, S., Hunt, J.C., and Hashimoto, J. (2000). Phylogenetic characterization of endosymbionts in three hydrothermal vent mussels influencehost distributions, influence on host distributions. Mar. Ecol. Prog. Ser. *208*, 147–155.

Geier, B., Sogin, E.M., Michellod, D., Janda, M., Kompauer, M., Spengler, B., Dubilier, N., and Liebeke, M. (2020). Spatial metabolomics of in situ host–microbe interactions at the micrometre scale. Nat. Microbiol. *5*, 498–510.

Go, G.W., and Mani, A. (2012). Low-density lipoprotein receptor (LDLR) family orchestrates cholesterol homeostasis. Yale J. Biol. Med. *85*, 19–28.

Goffredi, S.K., Tilic, E., Mullin, S.W., Dawson, K.S., Keller, A., Lee, R.W., Wu, F., Levin, L.A., Rouse, G.W., Cordes, E.E., et al. (2020). Methanotrophic bacterial symbionts fuel dense populations of deep-sea feather duster worms (Sabellida, Annelida) and extend the spatial influence of methane seepage. Sci. Adv. *6*, eaay8562.

Gómez-Mendikute, A., Elizondo, M., Venier, P., and Cajaraville, M.P. (2005). Characterization of mussel gill cells in vivo and in vitro. Cell Tissue Res. *321*, 131–140.

Govenar, B. (2010). Shaping vent and seep communities: habitat provision and modification by foundation species. In The Vent and Seep Biota: Aspects from Microbes to Ecosystems, S. KIEL, ed. (Springer Netherlands), pp. 403–432.

Halary, S., Riou, V., Gaill, F., Boudier, T., and Duperron, S. (2008). 3D FISH for the quantification of methane- and sulphur-oxidizing endosymbionts in bacteriocytes of the hydrothermal vent mussel Bathymodiolus azoricus. ISME J. *2*, 284–292.

Hand, S.C. (1987). Trophosome ultrastructure and the characterization of isolated bacteriocytes from invertebrate-sulfur bacteria symbioses. Biol. Bull. *173*, 260–276.

Harper, E.M. (2005). Fossil invertebrates | bivalves. In Encyclopedia of Geology, R.C. Selley, L.R.M. Cocks, and I.R. Plimer, eds. (Elsevier), pp. 369–378.

Hinzke, T., Kleiner, M., Breusing, C., Felbeck, H., Hasler, R., Sievert, S.M., Schluter, R., Rosenstiel, P., Reusch, T.B.H., Schweder, T., et al. (2019). Host-Microbe Interactions in the Chemosynthetic Riftia Pachyptila Symbiosis. mBio *10*, e02243–19.

Hoek, R.M., Li, K.W., Van Minnen, J., Lodder, J.C., De Jong-Brink, M., Smit, A.B., and Van Kesteren, R.E. (2005). LFRFamides: a novel family of parasitation-induced -RFamide neuropeptides that inhibit the activity of neuroendocrine cells in Lymnaea stagnalis. J. Neurochem. *92*, 1073–1080.

Hongo, Y., Nakamura, Y., Shimamura, S., Takaki, Y., Uematsu, K., Toyofuku, T., Hirayama, H., Takai, K., Nakazawa, M., Maruyama, T., and Yoshida, T. (2013). Exclusive localization of carbonic anhydrase in bacteriocytes of the deep-sea clam Calyptogena okutanii with thioautotrophic symbiotic bacteria. J. Exp. Biol. *216*, 4403–4414.

Ikuta, T., Tame, A., Saito, M., Aoki, Y., Nagai, Y., Sugimura, M., Inoue, K., Fujikura, K., Ohishi, K., Maruyama, T., and Yoshida, T. (2019). Identification of cells expressing two peptidoglycan recognition proteins in the gill of the vent mussel, Bathymodiolus septemdierum. Fish Shellfish Immunol. *93*, 815–822.

Jiang, S., Li, H., Zhang, D., Zhang, H., Wang, L., Sun, J., and Song, L. (2015). A C1q domain containing protein from Crassostrea gigas serves as pattern recognition receptor and opsonin with high binding affinity to LPS. Fish Shellfish Immunol. *45*, 583–591.

Jillian, P.M., and Nicole, D. (2009). Methanotrophic symbioses in marine invertebrates. Environ. Microbiol. Rep. *1*, 319–335.

Jørgensen, C.B. (1974). On gill function in the mussel Mytilus Edulis L. Ophelia *13*, 187–232.

Kalyuzhnaya, M.G., Gomez, O.A., and Murrell, J.C. (2019). The methane-oxidizing bacteria (methanotrophs). In Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes, T.J. Mcgenity, ed. (Springer International Publishing), pp. 1–34.

Kawabata, S., and Iwanaga, S. (1999). Role of lectins in the innate immunity of horseshoe crab. Dev. Comp. Immunol. *23*, 391–400.

Kleiner, M., Young, J.C., Shah, M., Verberkmoes, N.C., and Dubilier, N. (2013). Metaproteomics reveals abundant transposase expression in mutualistic endosymbionts. mBio *4*, e00223–13.

Lee, K.S., You, K.H., Choo, J.K., Han, Y.M., and Yu, K. (2004). Drosophila short neuropeptide F regulates food intake and body size. J. Biol. Chem. *279*, 50781–50789.

Liu, P., Rudick, M., and Anderson, R.G. (2002). Multiple functions of caveolin-1. J. Biol. Chem. *277*, 41295–41298.

Liu, R., Wang, L., Sun, Y., Wang, L., Zhang, H., and Song, L. (2014). A low-density lipoprotein receptor-related protein (LRP)-like molecule identified from Chlamys farreri participated in immune response against bacterial infection. Fish Shellfish Immunol. *36*, 336–343.

Martins, E., Figueras, A., Novoa, B., Santos, R.S., Moreira, R., and Bettencourt, R. (2014). Comparative study of immune responses in the deep-sea hydrothermal vent mussel Bathymodiolus azoricus and the shallow-water mussel Mytilus galloprovincialis challenged with Vibrio bacteria. Fish Shellfish Immunol. *40*, 485–499.

Mcfall-Ngai, M. (2008). Are biologists in 'future shock'? Symbiosis integrates biology across domains. Nat. Rev. Microbiol. *6*, 789–792.

Mcfall-Ngai, M., Hadfield, M.G., Bosch, T.C., Carey, H.V., Domazet-Loso, T., Douglas, A.E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. Proc. Natl. Acad. Sci. U S A *110*, 3229–3236.

Monika, B., and Angelika, S. (2003). Ultrastructural reinvestigation of the trophosome in adults of Riftia pachyptila (Annelida, Siboglinidae). Invertebr. Biol. *122*, 347–368.

Moya, A., Peretó, J., Gil, R., and Latorre, A. (2008). Learning how to live together: genomic insights into prokaryote–animal symbioses. Nat. Rev. Genet. *9*, 218–229.

Nakamura, Y., Konishi, M., Ohishi, K., Kusaka, C., Tame, A., Hatada, Y., Fujikura, K., Nakazawa, M., Fujishima, M., Yoshida, T., and Maruyama, T. (2013). Mucus glycoproteins selectively secreted from bacteriocytes in gill filaments of the deep-sea clam Calyptogena okutanii. Open J. Mar. Sci. *03*, 167–174.

Neumann, D., and Kappes, H. (2003). On the growth of bivalve gills initiated from a lobule-producing budding zone. Biol. Bull. *205*, 73–82.

Newton, I.L.G., and Bordenstein, S.R. (2011). Correlations between bacterial ecology and mobile DNA. Curr. Microbiol. *62*, 198–208.

Newton, I.L.G., Girguis, P.R., and Cavanaugh, C.M. (2008). Comparative genomics of vesicomyid clam (Bivalvia: Mollusca) chemosynthetic symbionts. BMC Genomics *9*, 585.

Nguyen, T.V., Alfaro, A.C., Young, T., and Merien, F. (2019). Tissue-specific immune responses to Vibrio sp. infection in mussels (Perna canaliculus): a metabolomics approach. Aquaculture *500*, 118–125.

Pette, D., and Staron, R.S. (2000). Myosin isoforms, muscle fiber types, and transitions. Microsc. Res. Tech. *50*, 500–509.

Peymen, K., Watteyne, J., Frooninckx, L., Schoofs, L., and Beets, I. (2014). The FMRFamide-like peptide family in nematodes. Front. Endocrinol. (Lausanne) *5*, 90.

Philipp, E.E., Kraemer, L., Melzner, F., Poustka, A.J., Thieme, S., Findeisen, U., Schreiber, S., and Rosenstiel, P. (2012). Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan Mytilus edulis. PLoS One *7*, e33091.

Pines, O., and Inouye, M. (1999). Expression and secretion of proteins in E. coli. Mol. Biotechnol. *12*, 25–34.

Piquet, B., Shillito, B., Lallier, F.H., Duperron, S., and Andersen, A.C. (2019). High rates of apoptosis visualized in the symbiont-bearing gills of deep-sea Bathymodiolus mussels. PLoS One *14*, e0211499.

Ponnudurai, R., Heiden, S.E., Sayavedra, L., Hinzke, T., Kleiner, M., Hentschker, C., Felbeck, H., Sievert, S.M., Schlüter, R., Becher, D., et al. (2020). Comparative proteomics of related symbiotic mussel species reveals high variability of host–symbiont interactions. ISME J. *14*, 649–656.

Ponnudurai, R., Kleiner, M., Sayavedra, L., Petersen, J.M., Moche, M., Otto, A., Becher, D., Takeuchi, T., Satoh, N., Dubilier, N., et al. (2017a). Metabolic and physiological interdependencies in the Bathymodiolus azoricus symbiosis. ISME J. *11*, 463–477.

Ponnudurai, R., Sayavedra, L., Kleiner, M., Heiden, S.E., Thurmer, A., Felbeck, H., Schluter, R., Sievert, S.M., Daniel, R., Schweder, T., et al. (2017b). Genome sequence of the sulfur-oxidizing Bathymodiolus thermophilus gill endosymbiont. Stand. Genomic Sci. *12*, 50.

Price, D.A., and Greenberg, M.J. (1977). Structure of a molluscan cardioexcitatory neuropeptide. Science *197*, 670–671.

Ricard-Blum, S., and Ruggiero, F. (2005). The collagen superfamily: from the extracellular matrix to the cell membrane. Pathol. Biol. *53*, 430–442.

Riffaud, C., Pinel-Marie, M.-L., and Felden, B. (2020). Cross-regulations between bacterial toxin–antitoxin systems: evidence of an interconnected regulatory network? Trends Microbiol. *28*, 851–866.

Sakoh, M., Ito, K., and Akiyama, Y. (2005). Proteolytic activity of HtpX, a membrane-bound and stress-controlled protease from Escherichia coli. J. Biol. Chem. *280*, 33305–33310.

Sayavedra, L., Kleiner, M., Ponnudurai, R., Wetzel, S., Pelletier, E., Barbe, V., Satoh, N., Shoguchi, E., Fink, D., Breusing, C., et al. (2015). Abundant toxin-related genes in the genomes of beneficial symbionts from deep-sea hydrothermal vent mussels. eLife *4*, e07966.

Schmitz-Esser, S., Penz, T., Spang, A., and Horn, M. (2011). A bacterial genome in transition–an exceptional enrichment of IS elements but lack of evidence for recent transposition in the symbiont Amoebophilus asiaticus. BMC Evol. Biol. *11*, 270.

Silke, J., and Meier, P. (2013). Inhibitor of apoptosis (IAP) proteins-modulators of cell death and inflammation. Cold Spring Harb. Perspect. Biol. *5*, a008730.

Streams, M.E., Fisher, C.R., and Fiala-Médioni, A. (1997). Methanotrophic symbiont location and fate of carbon incorporated from methane in a hydrocarbon seep mussel. Mar. Biol. *129*, 465–476.

Sun, J., Zhang, Y., Xu, T., Zhang, Y., Mu, H., Zhang, Y., Lan, Y., Fields, C.J., Hui, J.H.L., Zhang, W., et al. (2017a). Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. Nat. Ecol. Evol. *1*, 121.

Sun, Y., Wang, M., Li, L., Zhou, L., Wang, X., Zheng, P., Yu, H., Li, C., and Sun, S. (2017b). Molecular identification of methane monooxygenase and quantitative analysis of methanotrophic endosymbionts under laboratory maintenance in Bathymodiolus platifrons from the South China Sea. PeerJ *5*, e3565.

Takishita, K., Takaki, Y., Chikaraishi, Y., Ikuta, T., Ozawa, G., Yoshida, T., Ohkouchi, N., and Fujikura, K. (2017). Genomic evidence that methanotrophic endosymbionts likely provide deep-sea Bathymodiolus mussels with a sterol intermediate in cholesterol biosynthesis. Genome Biol. Evol. *9*, 1148–1160.

Tashian, R.E. (1989). The carbonic anhydrases: widening perspectives on their evolution, expression and function. Bioessays *10*, 186–192.

Taylor, J.D., and Glover, E.A. (2010). Chemosymbiotic bivalves. In The Vent and Seep Biota: Aspects from Microbes to Ecosystems, S. Kiel, ed. (Springer Netherlands), pp. 107–135.

Thingstad, T.F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnol. Oceanogr. *45*, 1320–1328.

Touchon, M., and Rocha, E.P. (2007). Causes of insertion sequences abundance in prokaryotic genomes. Mol. Biol. Evol. *24*, 969–981.

Van Dover, C.L., and Fry, B. (1989). Stable isotopic compositions of hydrothermal vent organisms. Mar. Biol. *102*, 257–263.

Veenendaal, A.K.J., Van Der Does, C., and Driessen, A.J.M. (2004). The protein-conducting channel SecYEG. Biochim. Biophys. Acta *1694*, 81–95.

Vrijenhoek, R.C. (2010). Genetic diversity and connectivity of deep-sea hydrothermal vent metapopulations. Mol. Ecol. *19*, 4391–4411.

Wang, H., Zhang, H., Wang, M.X., Chen, H., Lian, C., and Li, C. (2019). Comparative transcriptomic analysis illuminates the host-symbiont interactions in the deep-sea mussel Bathymodiolus platifrons. Deep-Sea Res. I Oceanogr. Res. Pap. 151.

Wang, J., Liu, B., Wang, N., Lee, Y.-M., Liu, C., and Li, K. (2011). TRIM56 is a virus- and interferon-inducible E3 ubiquitin ligase that restricts pestivirus infection. J. Virol. *85*, 3733–3745.

Wentrup, C., Wendeberg, A., Schimak, M., Borowski, C., and Dubilier, N. (2014). Forever competent: deep-sea bivalves are colonized by their chemosynthetic symbionts throughout their lifetime. Environ. Microbiol. *16*, 3699–3713.

Werren, J.H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. Proc. Natl. Acad. Sci. U S A *108*, 10863–10870.

Wong, Y.H., Sun, J., He, L.S., Chen, L.G., Qiu, J.W., and Qian, P.Y. (2015). High-throughput transcriptome sequencing of the cold seep mussel Bathymodiolus platifrons. Sci. Rep. *5*, 16597.

Xu, T., Feng, D., Tao, J., and Qiu, J.-W. (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 Res. Part Oceanogr. Res. Pap. *146*, 79–90.

Yang, C., Wang, L., Zhang, H., Wang, L., Huang, M., Sun, Z., Sun, Y., and Song, L. (2014). A new fibrinogen-related protein from Argopecten irradians (AiFREP-2) with broad recognition spectrum and bacteria agglutination activity. Fish Shellfish Immunol. *38*, 221–229.

Yang, Y., Sun, J., Sun, Y., Kwan, Y.H., Wong, W.C., Zhang, Y., Xu, T., Feng, D., Zhang, Y., Qiu, J.W., et al. (2020). Genomic, transcriptomic, and proteomic insights into the symbiosis of deep-sea tubeworm holobionts. ISME J. *14*, 135–150.

Yu, L.C. (2015). Commensal bacterial internalization by epithelial cells: an alternative portal for gut leakiness. Tissue Barriers *3*, e1008895.

Zaas, D.W., Swan, Z., Brown, B.J., Wright, J.R., and Abraham, S.N. (2009). The expanding roles of caveolin proteins in microbial pathogenesis. Commun. Integr. Biol. *2*, 535–537.

Zatylny-Gaudin, C., and Favrel, P. (2014). Diversity of the RFamide peptide family in mollusks. Front. Endocrinol. (Lausanne) *5*, 178.

Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., and Zhang, G. (2015). Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. *5*, 8693.

Zheng, P., Wang, M., Li, C., Sun, X., Wang, X., Sun, Y., and Sun, S. (2017). Insights into deep-sea adaptations and host-symbiont interactions: a comparative transcriptome study on Bathymodiolus mussels and their coastal relatives. Mol. Ecol. *26*, 5133–5148.

Zheng, Y., and Gao, C. (2019). E3 ubiquitin ligases, the powerful modulator of innate antiviral immunity. Cell Immunol. *340*, 103915.

iScience, Volume *24*

Supplemental Information

Molecular analyses of the gill symbiosis of the bathymodiolin mussel *Gigantidas platifrons*

Hao Wang, Huan Zhang, Zhaoshan Zhong, Yan Sun, Minxiao Wang, Hao Chen, Li Zhou, Lei Cao, Chao Lian, and Chaolun Li

Transparent Methods:

Animal collection:

The deep-sea mussels *G. platifrons* were collected during R/V *Kexue* "2017 South China Sea Cold-seep expedition" at the "F-site" of Jiulong reef (Depth 1113m, 20° 06'57.144" N, 119° 17'6.580" E). The mussels were stored in a storage apparatus (Du et al., 2018) during the ascending of the ROV to avoid pressure and temperature fluctuations.

Bacteriocyte enrichment:

All mussels were processed immediately after the ROV was retrieved on deck. For bacteriocyte enrichment, six *G. platifrons* mussels were dissected, and their gill slices were distributed into three groups as independent replicates. The gill slices were washed three times with pre-chilled FAISSW (Filtered and Autoclaved *in situ* seawater. The *In situ* seawater was collected from previous ROV dives, which was then filtered by 0.22μM membrane, autoclaved, and pre-chilled to 4°C) to remove cell debris, and then incubated in pre-chilled Trypsin solution (0.25% Trypsin, 1mM EDTA, prepared in FAISSW and pre-chilled to 4°C) for 30 minutes. The supernatant of the trypsin solution was first filtered through gaze and then 50 μm nylon filter. A small portion (3 X 1mL) of filtered supernatant was collected from each replicate, fixed with formaldehyde, and stored at 4°C. These formaldehyde-fixed samples were then used to calculate the proportion of bacteriocytes in each sample. The remaining cells were spun down by centrifuging at 800 g for 10 min at 4°C. The cell pellets were fixed in cell fixative (4% PFA, freshly prepared with FAISSW, pre-chilled to 4°C) for 15 min on ice. The pellet was then washed

three times with ice-cold PBS, then resuspended in storage buffer (1×PBS, 1% RNase-free BSA [Ambion], and 1U/μL Recombinant RNase Inhibitor [Takara]) and stored at -80°C until use.

Meanwhile, gill slices from 5 individual mussels were dissected. The gill slices were washed three times with FAISSW, then fixed, washed, and stored with the same method described above.

RNA extraction:

Total RNA of enriched bacteriocytes and gill slices were extracted by using the miRNeasy FFPE kit (Qiagen). The RNA extractions were conducted according to the manufacturer's protocol except that the first de-waxing step was skipped. The quantity and quality of each RNA sample were examined by using Nanodrop ND-1000 and Agilent 2100 Bioanalyzer.

Library construction and high-throughput sequencing:

All eight RNA samples, including three enriched bacteriocyte samples and five whole gill slices samples, were sent to Novogene (Beijing, China) for messenger RNA enrichment, RNA fragmentation, cDNA synthesis, adaptor ligation, RNA-seq library construction by using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, USA) according to standard protocol.

The remaining RNA of three EB samples was pooled together, from which a metatranscriptomic library was constructed and sequenced. The ribosomal RNA of this sample was removed by using a Ribo-Zero rRNA Removal Kit (Illumina) according to the manufacturer's protocol. The RNA was then fragmented, and the first-strand cDNA was synthesized by random

hexamer priming. The second strand cDNA synthesis and sequencing library construction were also conducted according to standard protocol.

All nine libraries were sequenced on the Illumina Hiseq4000 platform to produce pair ends reads of 150 bp read length.

Bioinformatics:

The raw reads were firstly processed through QC-Chain (Zhou et al., 2013) to remove the adaptor and low-quality reads. For enriched bacteriocytes VS whole gill analysis, gene expression levels were quantified by directly mapping clean reads to the *G. platifrons* transcripts model using bioinformatics tool Salmon v1.2.1 with "—seqBias," "—gcBias," "—posBias," and "—validateMappings" flag selected (Patro et al., 2017). R package edgeR v3.3 was used to normalize (with TMM method) read counts and identify the bifacially expressed transcripts (Robinson et al., 2010). The transcripts with logFC (log2 fold-change) >1 and FDR (false detection rate) <0.05 were considered as differentially expressed (the read counts and detailed DEG analysis results were provided in Supplementary Data S1). The GO enrichment analysis of the WG and EB enriched genes was performed using R package clusterProfiler 3.10.1 (Yu et al., 2012).

For symbiont transcriptome analysis, the gene expression level was quantified by mapping clean reads to *G. platifrons* symbiont metagenomic dataset (Takishita et al., 2017) by using Salmon. The symbiont gene expression profile was provided in Supplementary data S2.

FISH, ISH, and IHC:

The gills of freshly collected mussels were fixed in 4% paraformaldehyde (freshly prepared in FAISSW) at 4°C overnight. The gills were washed three times with ice-cold PBS to remove remaining fixative and then dehydrated in 100% methanol. The samples were stored at -20°C until use.

For FISH (fluorescent *in situ* hybridization) analysis, the gill sections were dehydrated, Paraplast plus (Sigma-Aldrich) embedded according to standard protocol. Sections of 7μm thickness were cut using a microtome (Leica). The double FISH with FITC labeled EU338 eubacteria probe, and Cy3 labeled symbiont specific probe (Supplementary Tab. S6) was performed according to the method described by Halary et al., (2008).

For ISH (*in situ* hybridization) analysis, a fragment of each targeted transcripts was amplified using standard PCR reaction with gene-specific primer pairs and *G. platifrons* gill cDNA as template (The primers were list in Supplementary Tab. S6). The PCR fragment was ligated into PCR II dual promoter vector (Invitrogen), and Sanger sequenced to confirm the insertion. The DNA template for *in vitro* mRNA transcription was amplified by using T7 or Sp6 primer combined with either forward or reversed gene-specific primer. Digoxigenin labeled anti-sense probes or sense control probe were transcribed using the DIG labeling mix (Roche) and Maxi Script kit with either T7 or Sp6 enzyme (Ambion) according to the manufacturer's protocol, and then cleaned up using SigmaSpin post reactive cleanup column (Sigma-Aldrich). The *in situ* hybridizations were performed according to the methods described by Cheng et al., (2016).

For IHC (immunohistochemistry) analysis, the sections were first permeabilized by being washed three times in PBST (1XPBS, 0.1% Tween20), 15 min each at room temperature. The sections were blocked by incubating in Blocking buffer (PBST, 2% BSA, 2% sheep serum) for 1 hr at room temperature. Then, the sections were incubated with primary antibody: α-Tubulin

Rabbit pAb or β-Tubulin Rabbit pAb (ABclonal) 1:100 diluted with Blocking buffer at 4°C overnight. The sections were washed with PBST three times and incubated with a fluorescent-labeled secondary antibody (Alexa Fluor 488 labeled goat anti-rabbit secondary antibody, Invitrogen) 1:1000 diluted with Blocking buffer for 2 hr at room temperature. The sections were washed three times with PBST and then stained with DAPI.

All sections were observed and imaged with a Nikon Eclipse Ni microscope with a DS-Ri2 camera.

Electron microscopy analysis:

The gill slices of the *G. platifrons* were dissected and fixed in electron microcopy fixative (2.5% glutaraldehyde and 2% paraformaldehyde) at 4°C. For SEM (scanning electron microscope) analysis, the samples were dehydrated in a graded ethanol series, and then critical point dried. The samples were then coated with gold (Sputter/Carbon Thread, EM ACE200) and observed under SEM (VEGA3, TESCAN). For TEM (transmission electron microscope) analysis, the samples were rinsed with double distilled water, post-fixed with 1% osmium tetroxide, then washed with double distilled water. The samples were then rinsed, dehydrated, and embedded in Ep812 resin. Ultra-thin sections were obtained with an ultramicrotome (thickness 70nm, Reichert-Jung ULTRACUT E). The sections were then double-stained with lead-citrate and uranyl acetate. The sections were observed under the TEM (JEM1200, JEOL) operated under 100KV.

Supplementary Table S1: Summary of the illumina sequencing. Related to Main text and Figure 4.

Library	*Number of reads*	*Accession number*	*Data size*	*Mapping rates*
EB1	20.6 M	SAMN15514069	14.1 Gb	40.37%
EB2	29.4 M	SAMN15514070	20.1 Gb	38.94%
EB3	30.3 M	SAMN15514071	20.7 Gb	38.43%
WG1	27.4 M	SAMN15514072	18.8 Gb	46.07%
WG2	25.9 M	SAMN15514073	17.7 Gb	45.05%
WG3	27.2 M	SAMN15514074	18.6 Gb	45.45%
WG4	29.9 M	SAMN15514075	20.5 Gb	46.23%
WG5	24.0 M	SAMN15514076	16.4 Gb	46.82%
Meta	80.3 M	SAMN15514077	55.0 Gb	Symbiont: 51.3455%

Supplementary Table S3: Classification of the gene groups among the top10% highest expressed symbiont genes. Related to Main Text and Figure 5.

	Number of genes	*Total TPM*	*Percentage of protein coding genes' TPM*
Methane oxidation	10	122220.1	23.51%
Ribosomal protein	45	55074.04	10.60%
DUF1566 domain containing protein	21	35476.59	6.83%
Other DUF domain containing protein	9	41279.49	7.94%
IS family transposases	15	13597.51	2.62%
DDE domain-containing protein	7	6497.873	1.25%
Transposases	5	3125.463	0.60%
Resolvase	2	1850.266	0.36%
Cold-shock protein	4	10547.12	2.03%
Transcription	9	4068.963	0.78%
Antitoxin	6	3163.57	0.61%
Toxin	2	673.3065	0.13%
Subunits of FATP synthase	4	2406.56	0.46%
Hypothetical protein	159	189664.3	36.49%
Totals	298	489645.2	94.21%

Supplementary Table S4: Signal peptide cleavage site of the EB enriched Fibrogens, predicted by SignalP 5.0 (Almagro Armenteros et al., 2019). Related to Main Text and Figure 7.

GeneID	*Prediction*	*SP(Sec/SPI)*	*OTHER*	*CS Position*
Bpl_scaf_18519-1.3	OTHER	0.001845	0.998155	
Bpl_scaf_18519-2.15	SP(Sec/SPI)	0.994174	0.005826	CS pos: 17-18. ALS-CP. Pr: 0.6507
Bpl_scaf_56025-0.28	SP(Sec/SPI)	0.986977	0.013023	CS pos: 17-18. ALS-CR. Pr: 0.4836
Bpl_scaf_18519-2.13	SP(Sec/SPI)	0.992812	0.007188	CS pos: 17-18. ALS-CP. Pr: 0.6542

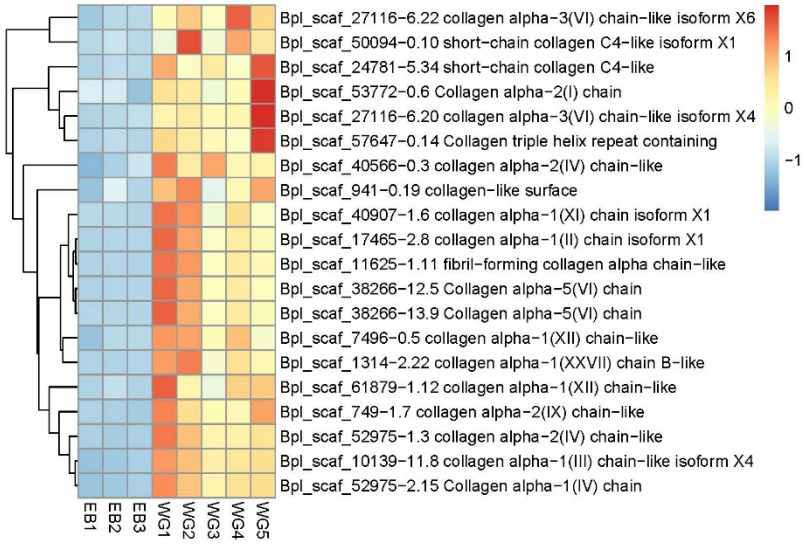
Supplementary Table S5: Signal peptide cleavage site of the Symbiont DUF1566 domain-containing proteins, predicted by SignalP 5.0 (Almagro Armenteros et al., 2019). Related to Main Text and Figure 5.

GeneID	*Prediction*	*SP(Sec/SPI)*	*OTHER*	*CS Position*
gene2595	SP(Sec/SPI)	0.996868	0.001058	CS pos: 20-21. ALA-IN. Pr: 0.9889
gene1713	SP(Sec/SPI)	0.993655	0.000163	CS pos: 24-25. ANA-DF. Pr: 0.9820
gene4084	SP(Sec/SPI)	0.971581	0.000361	CS pos: 28-29. ANA-DF. Pr: 0.9494
gene3511	SP(Sec/SPI)	0.994084	0.001929	CS pos: 23-24. VMA-YG. Pr: 0.7611
gene3533	SP(Sec/SPI)	0.958338	0.004127	CS pos: 20-21. CNA-GL. Pr: 0.8266
gene2457	SP(Sec/SPI)	0.985113	0.010161	CS pos: 20-21. VQA-KA. Pr: 0.5727
gene1712	SP(Sec/SPI)	0.819436	0.032167	CS pos: 21-22. FYS-AP. Pr: 0.4883
gene1710	SP(Sec/SPI)	0.956896	0.031043	CS pos: 18-19. SFA-YV. Pr: 0.9407
gene3501	SP(Sec/SPI)	0.972752	0.004006	CS pos: 20-21. SNA-GL. Pr: 0.7222
gene2466	SP(Sec/SPI)	0.737482	0.020708	CS pos: 29-30. VNA-AG. Pr: 0.6260
gene2933	SP(Sec/SPI)	0.998393	0.000411	CS pos: 19-20. VHA-SE. Pr: 0.9954
gene1521	SP(Sec/SPI)	0.9964	0.000361	CS pos: 24-25. VNA-NE. Pr: 0.9886
gene1969	SP(Sec/SPI)	0.998478	0.000516	CS pos: 20-21. ANA-ED. Pr: 0.9975
gene3073	SP(Sec/SPI)	0.926808	0.032259	CS pos: 25-26. AYG-DA. Pr: 0.4953
gene3062	SP(Sec/SPI)	0.98877	0.003792	CS pos: 22-23. AYG-WE. Pr: 0.5101
gene3499	SP(Sec/SPI)	0.968858	0.002964	CS pos: 20-21. ASA-SG. Pr: 0.6497
gene3617	SP(Sec/SPI)	0.989745	0.004321	CS pos: 20-21. ASA-YG. Pr: 0.5120
gene3628	SP(Sec/SPI)	0.990786	0.00028	CS pos: 23-24. SQA-KT. Pr: 0.8853
gene3629	SP(Sec/SPI)	0.997335	0.000642	CS pos: 22-23. VHA-KT. Pr: 0.9789
gene3640	SP(Sec/SPI)	0.969017	0.000398	CS pos: 22-23. ASA-SQ. Pr: 0.8815
gene3487	SP(Sec/SPI)	0.97014	0.023506	CS pos: 29-30. AHA-TS. Pr: 0.9181
gene3638	SP(Sec/SPI)	0.995368	0.001713	CS pos: 18-19. AFA-DA. Pr: 0.8587
gene2101	SP(Sec/SPI)	0.953553	0.002318	CS pos: 21-22. AVA-SQ. Pr: 0.8519
gene3060	SP(Sec/SPI)	0.877447	0.036259	CS pos: 20-21. SFS-DV. Pr: 0.8094

gene2099	SP(Sec/SPI)	0.986336	0.005238	CS pos: 19-20. AFA-GF. Pr: 0.9112
gene1959	SP(Sec/SPI)	0.996298	0.001155	CS pos: 22-23. AMA-EK. Pr: 0.9747
gene4010	OTHER	0.041643	0.941794	
gene2028	SP(Sec/SPI)	0.994558	0.00099	CS pos: 20-21. ANA-GH. Pr: 0.8245
gene1957	SP(Sec/SPI)	0.928525	0.021597	CS pos: 31-32. ASA-IS. Pr: 0.7007
gene1902	SP(Sec/SPI)	0.991628	0.001903	CS pos: 20-21. ANA-GH. Pr: 0.8431
gene2533	OTHER	0.015623	0.979661	
gene3820	SP(Sec/SPI)	0.977618	0.012297	CS pos: 30-31. VNA-SG. Pr: 0.8510

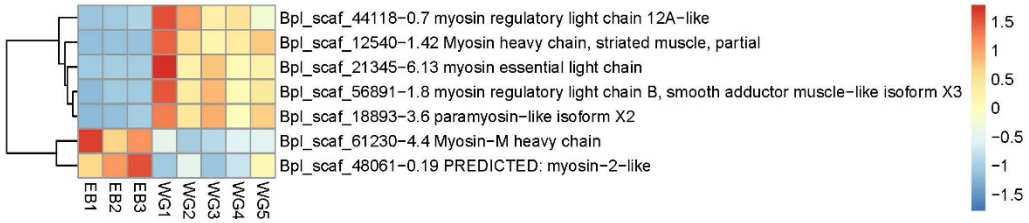
Supplementary Table S6: Probes and Primers. Related to Transparent Methods.

Probes or Primers	*Sequences 5'-3'*	*Reference*
EUB338 (5' FITC labeled probe)	GCTGCCTCCCGTAGGAGT	(Amann et al., 1990)
Symb (5' Cy3 labeled probe)	CGAGATATTATCCTCGCCTG	Present study
Bpl_scaf_11625-1.11-F	GGTGAGAGAGGACGAGATGG	Present study
Bpl_scaf_11625-1.11-R	CCTAATTCTCCCCTCGGTCC	Present study
Bpl_scaf_1314-2.22-F	TGCAAAGGAGCCAGAAGAGG	Present study
Bpl_scaf_1314-2.22-R	ACCGACCACCAATGACTGTT	Present study
Bpl_scaf_59427-1.13-F	TTAGCTGACCAGGGAGTGTG	Present study
Bpl_scaf_59427-1.13-R	GTGCAACCCAGCCCATAAAA	Present study
Bpl_scaf_41549-2.13-F	TGCCTCCTCCATTAAAGCCA	Present study
Bpl_scaf_41549-2.13-R	GTACACACTATTGGCCGCTG	Present study
Bpl_scaf_10841-2.17-F	GGTGTGTGTTGTTGTGTCTGA	Present study
Bpl_scaf_10841-2.17-R	AATGACGACCAAGGCTGTTC	Present study
Bpl_scaf_33596-7.3-F	GGCTAGTCTGCTTTGTTCCG	Present study
Bpl_scaf_33596-7.3-R	ACAGTGCTGCTATGTGTTGC	Present study
Bpl_scaf_21642-0.2-F	AGGAAACAACGAACCGGATG	Present study
Bpl_scaf_21642-0.2-R	TGTCTTGGTGGTCTGCTAGG	Present study
Bpl_scaf_54034-5.14-F	ATATCCAGAAATCCACCTCCAG	Present study
Bpl_scaf_54034-5.14-R	ATTTGTGTGTCCAATCTGAGGC	Present study
Bpl_scaf_18519-2.15-F	AACCTGACATAGCCTGACGT	Present study
Bpl_scaf_18519-2.15-R	CAACGTCCTTGTCATGTGCA	Present study
Bpl_scaf_19969-0.60-F	CCAACAACTACGCCAGAGGT	Present study
Bpl_scaf_19969-0.60-R	TGTGATCAAGACGAGCCCAG	Present study
Bpl_scaf_33659-0.9-F	TATGGGGTGTGAGACTGACG	Present study
Bpl_scaf_33659-0.9-R	CCTGGGTTGGTATTGTGACG	Present study



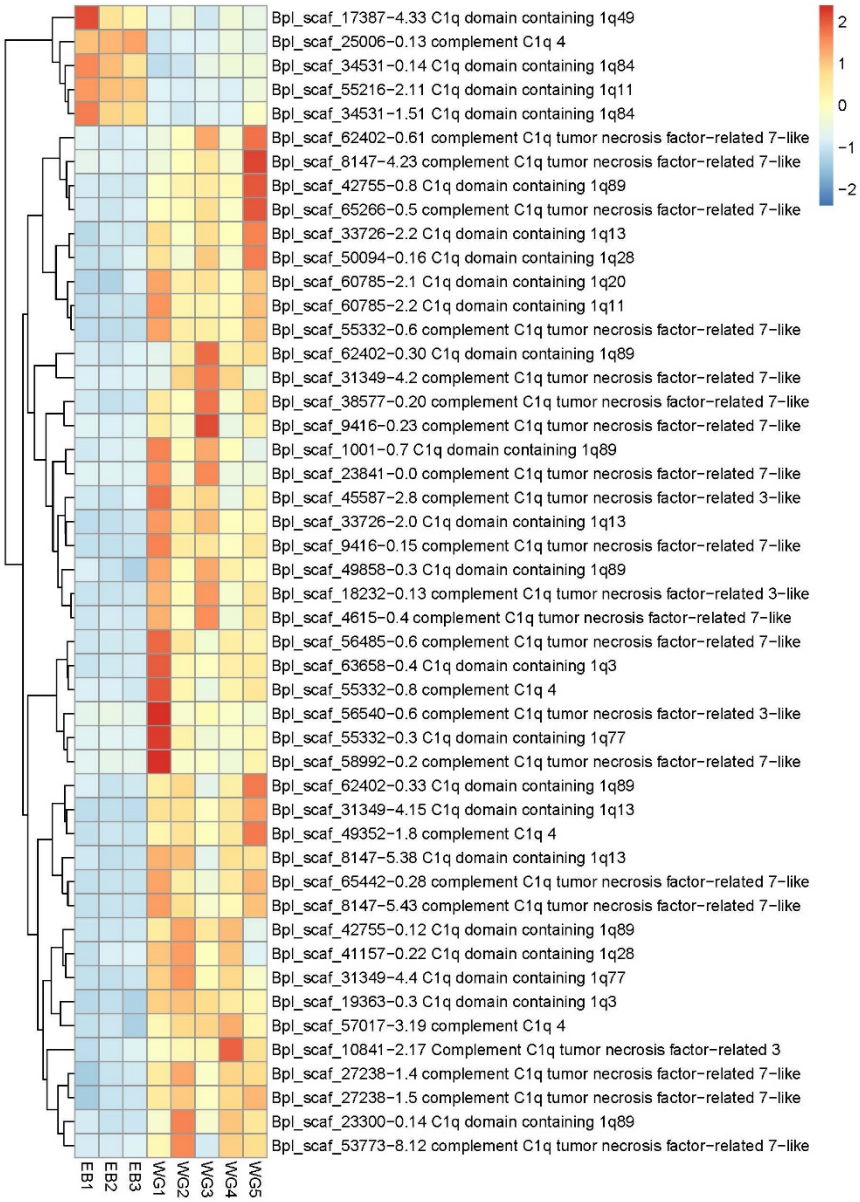
Supplementary Figure S1: Heatmap showing the expression pattern of the differentially expressed transcripts encoding Collagens. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



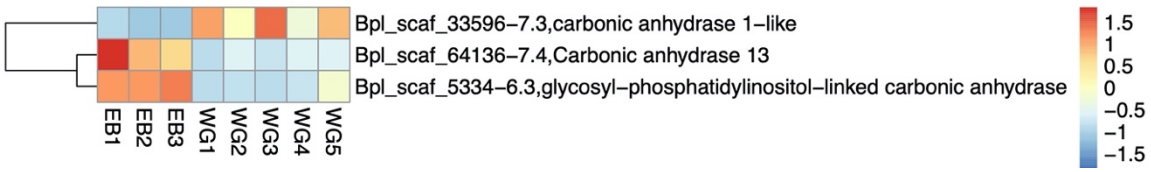
Supplementary Figure S2: Heatmap showing the expression pattern of the differentially expressed transcripts encoding myosin. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



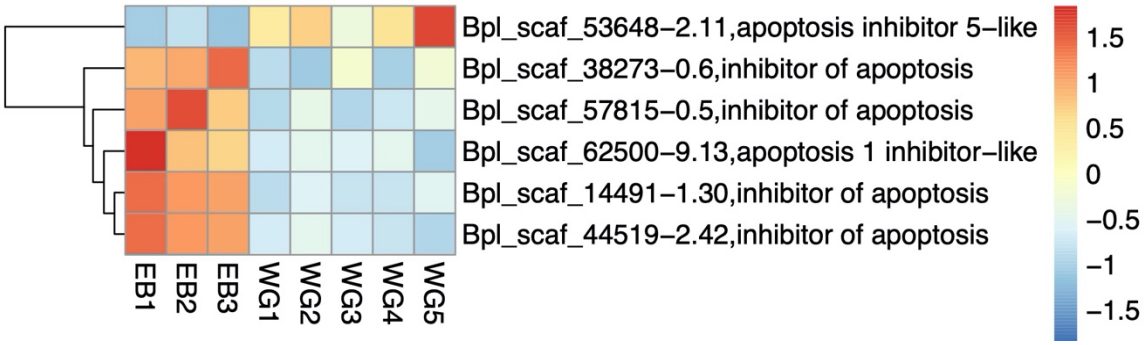
Supplementary Figure S3: Heatmap showing the expression pattern of the differentially expressed transcripts encoding C1q domain containing proteins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



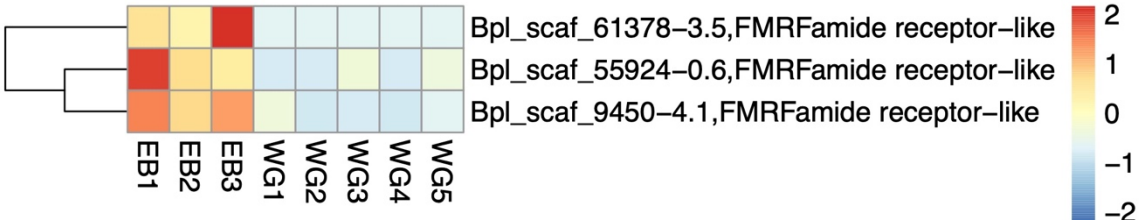
Supplementary Figure S4: Heatmap showing the expression pattern of the differentially expressed transcripts encoding Carbonic anhydrase (CAs) proteins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



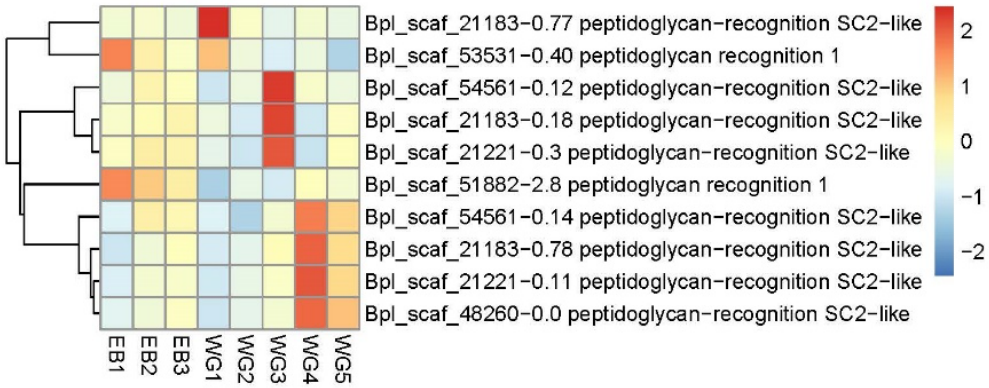
Supplementary Figure S5: Heatmap showing the expression pattern of the differentially expressed transcripts encoding inhibitor of apoptosis (IAs) proteins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



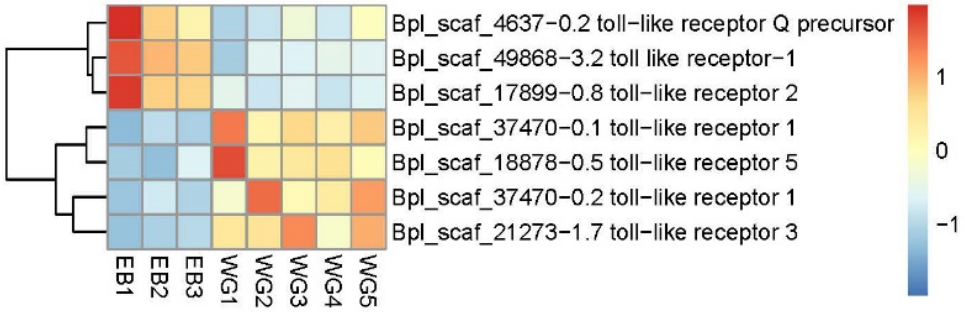
Supplementary Figure S6: Heatmap showing the expression pattern of the differentially expressed transcripts encoding FMRFamide receptor-like proteins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



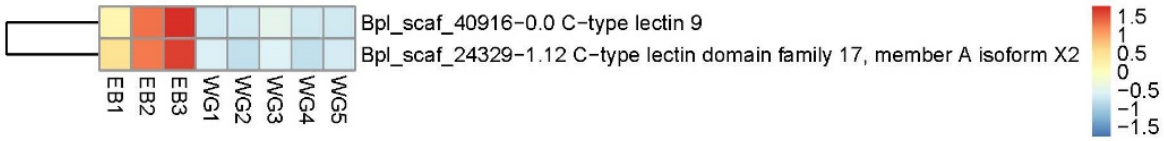
Supplementary Figure S7: Heatmap showing the expression pattern of the transcripts encoding Peptidoglycan recognition proteins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



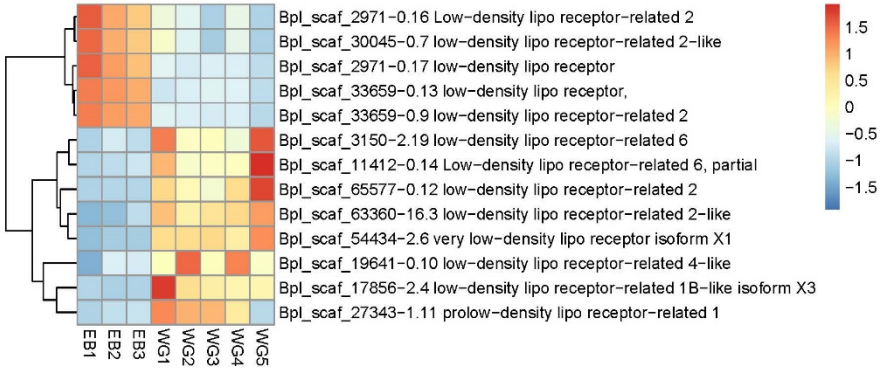
Supplementary Figure S8: Heatmap showing the expression pattern of the differentially expressed transcripts encoding Toll-like receptors. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



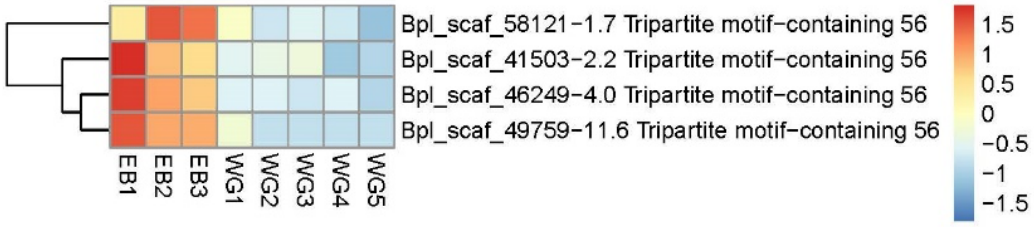
Supplementary Figure S9: Heatmap showing the expression pattern of the differentially expressed transcripts encoding C-type lectins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



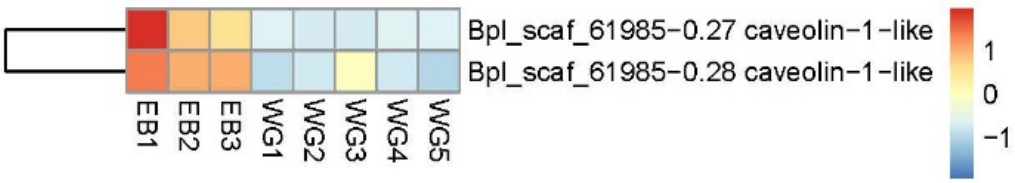
Supplementary Figure S10: Heatmap showing the expression pattern of the differentially expressed transcripts encoding Low-density lipo receptors and Low density lipo receptor-related proteins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



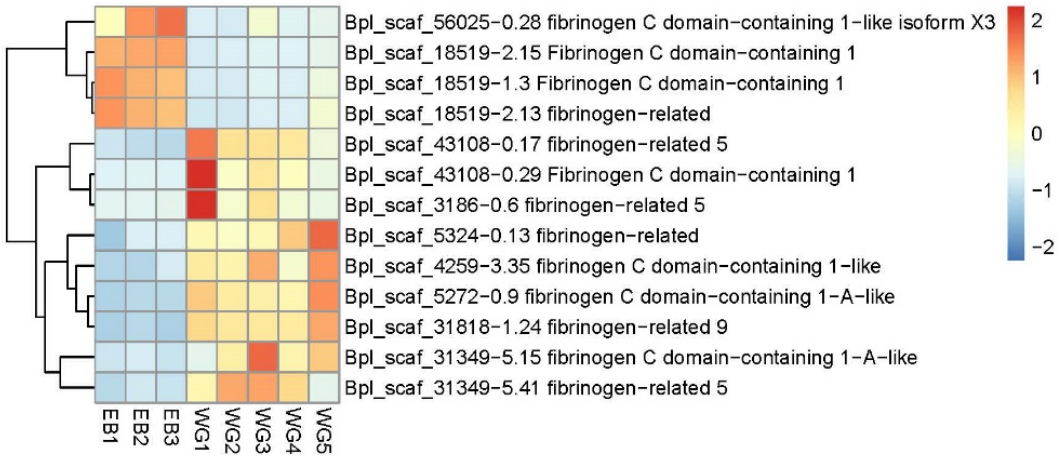
Supplementary Figure S11: Heatmap showing the expression pattern of the differentially expressed transcripts encoding Tripartite motif-condaing 56 proteins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



Supplementary Figure S12: Heatmap showing the expression pattern of the differentially expressed transcripts encoding Caveolins. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.



Supplementary Figure S13: Heatmap showing the expression pattern of the differentially expressed transcripts encoding Fibrinogens. Related to Main Text and Figure 4.

The hierarchical cluster shown here was obtained by comparing the expression values of all five sample groups. The blue-red scale represents the relative expression values.

References:

Almagro Armenteros, J. J., Tsirigos, K. D., Sonderby, C. K., Petersen, T. N., Winther, O., Brunak, S., Von Heijne, G. & Nielsen, H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol,* 37**,** 420-423.

Amann, R. I., Binder, B. J., Olson, R. J., Chisholm, S. W., Devereux, R. & Stahl, D. A. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Applied and environmental microbiology,* 56**,** 1919-1925.

Cheng, Q., Wang, H., Jiang, S., Wang, L., Xin, L., Liu, C., Jia, Z., Song, L. & Zhu, B. 2016. A novel ubiquitin-protein ligase E3 functions as a modulator of immune response against lipopolysaccharide in Pacific oyster, Crassostrea gigas. *Developmental & Comparative Immunology,* 60**,** 180-190.

Du, Z., Zhang, X., Xi, S., Li, L., Luan, Z., Lian, C., Wang, B. & Yan, J. 2018. In situ Raman spectroscopy study of synthetic gas hydrate formed by cold seep flow in the South China Sea. *Journal of Asian Earth Sciences,* 168**,** 197-206.

Halary, S., Riou, V., Gaill, F., Boudier, T. & Duperron, S. 2008. 3D FISH for the quantification of methane- and sulphur-oxidizing endosymbionts in bacteriocytes of the hydrothermal vent mussel Bathymodiolus azoricus. *ISME J,* 2**,** 284-92.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods,* 14**,** 417-419.

Robinson, M. D., Mccarthy, D. J. & Smyth, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.

Takishita, K., Takaki, Y., Chikaraishi, Y., Ikuta, T., Ozawa, G., Yoshida, T., Ohkouchi, N. & Fujikura, K. 2017. Genomic Evidence that Methanotrophic Endosymbionts Likely Provide Deep-Sea Bathymodiolus Mussels with a Sterol Intermediate in Cholesterol Biosynthesis. *Genome Biol Evol,* 9**,** 1148-1160.

Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology,* 16**,** 284-287.

Zhou, Q., Su, X., Wang, A., Xu, J. & Ning, K. 2013. QC-Chain: fast and holistic quality control method for next-generation sequencing data. *PLoS One,* 8**,** e60234.