Deciphering deep-sea chemosynthetic symbiosis by single-nucleus RNA-sequencing

Hao Wang¹²³#, Kai He⁴#, Huan Zhang¹#, Quanyong Zhang⁵#, Lei Cao¹, Jing Li¹⁶, Zhaoshan Zhong¹, Hao Chen¹, Li Zhou¹, Chao Lian¹, Minxiao Wang¹, Kai Chen⁵, Pei-Yuan Qian²³*, and Chaolun Li¹⁶,⁷*

¹Center of Deep-Sea Research, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China
²Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Nansha, Guangzhou, China
³Department of Ocean Science, Hong Kong University of Science and Technology, Hong Kong S.A.R, China.
⁴Key Laboratory of Conservation and Application in Biodiversity of South China, School of Life Sciences, Guangzhou University, Guangzhou, Guangdong 510006, China
⁵State Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, China
⁶South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China
⁷University of Chinese Academy of Sciences, Beijing, China

Correspondence authors: Prof. Pei-yuan Qian and Prof. Chaolun Li

Email addresses: boqianpy@ust.hk and lcl@qdio.ac.cn

#These authors contributed equally to this work

The authors declare no conflict of interest.
Abstract:

Bathymodioline mussels dominate deep-sea methane seep and hydrothermal vent habitats and obtain nutrients and energy primarily through chemosynthetic endosymbiotic bacteria in the bacteriocytes of their gill. However, the molecular mechanisms that orchestrate mussel host-symbiont interactions remain unclear. Here, we constructed a comprehensive cell atlas of the gill in the mussel *Giantiidus platifrons* from the South China Sea methane seeps (1100m depth) using single-nucleus RNA sequencing (snRNA-seq) and whole-mount *in situ* hybridisation. We identified 13 types of cells, including three previously unknown ones, uncovered unknown tissue heterogeneity. Every cell type has a designated function in supporting the gill's structure and function, creating an optimal environment for chemosynthesis, and effectively acquiring nutrients from the endosymbiotic bacteria. Analysis of snRNA-seq of *in situ* transplanted mussels clearly showed the shifts in cell state in response to environmental oscillations. Our findings provide insight into principles of host-symbiont interaction and the bivalves' environmental adaption mechanisms.
Introduction:

Mutualistic interactions between multicellular animals and their microbiota play a fundamental role in animals' adaptation, ecology, and evolution. By associating with symbionts, host animals benefit from the metabolic capabilities of their symbionts and gain fitness advantages that allow them to thrive in habitats they could not live in on their own. Prime examples of such symbioses are bathymodioline mussels and gammaproteobacteria endosymbionts. The bathymodioline mussels occur worldwide at deep-sea chemosynthetic ecosystems, such as cold-seeps, hydrothermal vents, and whale falls. The host mussel acquired sulfur-oxidising (SOX) and/or methane-oxidising (MOX) symbiont through horizontal transmission at their early life stage. Their symbionts, which are hosted in a specialised gill epithelial cell, namely, the bacteriocytes, utilise the chemical energy from the reduced chemical compounds such as CH₄ and H₂S released from cold seeps or hydrothermal vents to fix carbon and turn into carbon source for the host mussel. The ecological success of the bathymodioline symbioses is apparent: the bathymodiolin mussels are often among the most dominant species in the deep-sea chemosynthetic ecosystems. Thus, it is critical to know how the bathymodioline mussels interact with the symbionts and maintain the stability and efficiency of the symbiosis.

The gill of bathymodiolin mussels has adapted to deep-sea chemosynthetic life at the levels of molecule, cell and tissue. Compared to shallow water mussels, the gill filaments of bathymodiolin mussels have enlarged surfaces allowing them to hold more symbionts per filament (Supplementary Fig. S1). This adaption requires not only novel cellular and molecular mechanisms to maintain and support the enlarged gill filament structure but also a strong ciliary ventilation system to pump vent/seep fluid to fuel the symbiotic bacteria. Previous studies based on whole-genome sequencing and bulk RNA-seq projects have shown that genes in the categories of nutrient transporters, lysosomal proteins, and immune receptors are either expanded in the host mussel's genome or up-regulated in the gill, suggesting that these genes involved in host-symbiont interaction. Though providing deep molecular insights, these studies mainly used homogenised tissues that average genes' expression levels amongst different cell types and eliminate cell and gene spatial distribution information. In addition,
the broad expression and function of potentially ‘symbiosis-related’ proteins also greatly
limited data interpretation. Therefore, a systemic atlas of gill cell types and the descriptions of
cell-type-specific gene expressional profiles are warranted to a better understanding of the
host-symbiont interaction and environmental adaptation mechanisms of the bathymodioline
symbiosis.

In recent years, single cell/nucleus RNA-sequencing (sc/sn RNA-seq) technologies have
become one of the preferred methods for investigating the composition of complex tissue at
the transcriptional level\textsuperscript{22}. snRNA-seq has several advantages, such as compatibility with frozen
samples, elimination of dissociation-induced transcriptional stress responses, and reduced
dissociation bias\textsuperscript{23}. These advantages are significant for deep-sea and other ecological studies
because cell and molecular biology facilities are not available in the field. To examine the
cellular and molecular mechanisms of the environmental adaptations and host-symbiont
interactions in bathymodioline mussels, we conducted a single nucleus RNA-sequencing-based
transcriptomic study. We analysed the gill symbiosis using deep-sea mussels Gigantidas
platifrons\textsuperscript{24}. Our work provides a proof-of-principle for environmental adaptation mechanisms
study in non-traditional but ecologically important organisms with single-nucleus RNA-
sequencing technologies.
Results and discussion

1. *G. platifrons* deep-sea *in situ* transplant experiment and single-cell transcriptomic analysis:

We conducted a *G. platifrons* deep-sea *in situ* transplant experiment at the “F-Site” Cold-seep (~1117 m depth) and retrieved three groups of samples (Fig. 1A), as follows: the ‘Fanmao’ (meaning prosperous) group, which comprised the mussels collected from methane-rich (~40 μM) site of the cold-seep, where the animals thrived; the ‘starvation’ group, which comprised the mussels collected from the methane-rich site and then moved to a methane-limited (~0.054 μM) starvation site for 14 days before retrieval; and the ‘reconstitution’ group, which comprised the methane-rich site mussels moved to the ‘Starvation’ site for 11 days and then moved back to the methane-rich site for another 3 days.

We next performed single nucleus RNA-sequencing (snRNA-seq) of the gill posterior tip of the three groups of samples using the microwell-based BD Rhapsody platform™ Single-Cell Analysis System (Fig. 1A). After quality control, we obtained 9717 (Fanmao), 21614 (Starvation), and 28928 (Reconstitution) high-quality single nuclei transcriptomic data.

To minimise the dropout effect in snRNAseq, we utilised a reciprocal PCA (RPCA) strategy to project cells in the three samples onto the same space based on conserved expressed genes among them\(^{25}\). This strategy maximises the number of cells per cluster regardless of the organism’s state and therefore maximises genes per cluster\(^{26}\). We revealed 14 cell clusters and their corresponding marker genes (Figs. 1B-D) with Seurat\(^{27,28}\). Because there are few available canonical marker genes per cell type for *G. platifrons* and molluscs, we carefully characterised each cell cluster by 1) examining the cluster marker genes' functions, 2) identifying the expression pattern of cluster marker gene using whole-mount *in situ* hybridisation (WISH) and/or double-fluorescent *in situ* hybridisation (FISH) analyses, 3) conducting scanning electron microscopy (SEM) and transmission electron microscopy (TEM) analyses, and 4) evaluating the stability per cluster using a bootstrapping algorithm. We then successfully identified 13 of the 14 cell clusters, which could be generally divided into four groups, including 1) the supportive cells, 2) the ciliary cells, 3) the proliferation cells, and 4) the bacteriocytes. One cell cluster remained ambiguous. Notably, a bootstrapping strategy suggested that all clusters were highly supported when using all three samples (Supplementary Fig. S2A). When using each sample...
individually, most clusters were also well supported except of the three ciliary cells (Supplementary Fig. S2 B-D). This could be due to that these cells were derived from the same precursor cells, and at least partially because they were represented by fewer numbers of nuclei and thus fewer available genes, especially in the food groove ciliary cell (Supplementary Table S1). The spatial and functional annotation of the cell clusters revealed that the cell clusters worked in close cooperation to support the symbiont and maintain a high-efficiency chemosynthetic system.

2. Cell atlas of G. platifrons gill

Supportive cells: We have identified four supportive cell types, namely, inter lamina cells, basal membrane cells (BMC)1, BMC2, and mucus cells (Figs 2A and 2B). The inter lamina cells, marked by high expression of fibrillar-forming collagens, were the cells located between the two layers of basal membranes. These cells were previously identified as amoeboid cells\textsuperscript{13}, and their function has not been explored. The inter lamina cells were densely distributed around the food groove and at the rim of the gill filament. While on the middle part of the gill filament, the inter lamina cells distributed in parallel rows. It might help to connect the two sheets of basal membranes and maintain the spatial integrity of the gill filament. In addition, we observed the enrichment of ribosomal proteins involved in protein synthesis (Supplementary Table S2), indicating a high metabolic rate\textsuperscript{29} of inter lamina cells. This implies that the inter lamina cells may process the nutrients acquired from the symbiont.

We have also identified two types of BMC populations which have not been recognised before. The BMC1 expressed genes encoding extracellular matrix and adhesive proteins (Fig. 2B and Supplementary Table S2), such as SCO-spondin (Bpl_scaf_27638-4.18), Zonadhesin (Bpl_scaf_1371-2.4), Cartilage matrix proteins (Bpl_scaf_16371-0.14 and Bpl_scaf_3387-5.47), and Collagen alpha-1(XXI) chain protein (Bpl_scaf_51350-0.19). BMC2 (Fig. 2B and Supplementary Table S2) was distinctive from BMC1 in the context of expression (Fig 2A) and marker genes (Fig 2B). Genes encoding hemicentins (Bpl_scaf_38731-0.50 and Bpl_scaf_7293-0.14), which are known to stabilise the cells’ contact with the basal membrane\textsuperscript{30}, were highly expressed in BMC2. Thus, these two cell types are likely to help building the basal lamina, and
stabilising the epithelial-derived cells, such as bacteriocytes and intercalary cells, on the surface of the basal membrane.

Mucus cells are specialised secretory cells with intracellular mucus vacuoles. In shallow-water filter-feeding bivalves, mucus cells secret mucus, which cooperates with the ciliary ventilation system to capture, process, and transport food particles to their mouths, known as filter feeding mechanism. Deep-sea mussels may not necessarily retain this function since they do not normally acquire food resources such as phytoplankton and planktonic bacteria but obtain most of nutrient through their chemosynthetic symbiont. Thus, it was hypothesised that mucus cells are involved in other biological functions such as immune responses to pathogens. Herein, genes encoding proteins with microbe-binding functionalities were enriched in mucus cells (Fig. 2B and Supplementary Table S2), such as C-type lectins (Bpl_scaf_58694-0.22 and Bpl_scaf_38133-0.8), C1q domain-containing protein (Bpl_scaf_55216-2.11), and Fibrinogen C domain-containing protein (Bpl_scaf_18519-2.15). The expression of similar immune genes was upregulated in a shallow water mussel *Mytilus galloprovincialis* when challenged by a pathogenic bacteria. Our WISH and double-FISH analyses showed that mucus cells were embedded within the outer rim cilia and scattered on the gill lamella alongside the bacteriocytes (Figs. 2G and 2K). These data collectively suggest that mucus cells may help mussel maintain the immune homeostasis of gill. Interestingly, WISH and double-FISH analyses showed that mucus cells were also distributed alongside the gill lamella's food groove and the inner edge, where the density of bacteriocytes is low (Fig. 2G). Because the food groove is the main entry of food practical to the labial palps and mouth, this distribution pattern implies that mucus cells may be also involved in capturing planktonic bacteria and sending them to the mouth.

**Ciliary and smooth muscle cells:** A remarkable feature of bathymodioline mussel's gill is its ciliary ventilation system, which constantly agitates the water and provides the symbiont with the necessary gas. We identified four types of ciliary cells (Figs. 3A and 3B): apical ridge ciliary cells (ARCCs), food groove ciliary cells (FGCCs), lateral ciliary cells (LCCs), and intercalary cells (ICs), as well as a newly identified type of smooth muscle cells. All ciliary cells were marked by canonical cilia genes, such as genes encoding flagellar proteins, ciliary motor kinesin proteins,
ciliary dynein proteins, and ciliary microtubules and were clearly distinguishable by specifically expressed genes (Supplementary Table S3 and Supplementary Fig. S3). ARCCs were characterized by expression of Bpl_scaf_20631-1.16, which encodes homeobox Dlx6a-like protein (Fig 3D and 3F) which is a marker of the apical ectodermal ridge. FGCCs expressed genes encoding primary cilia development regulator Tubby-related protein 3 and primary ciliary cell structural protein TOG array regulator of axonemal microtubules protein 1 and primary ciliary cell structural protein TOG array regulator of axonemal microtubules protein 1, suggesting the FGCCs are the sensory ciliary cells that gather information from the surrounding environment. Both ARCCs and FGCCs were located around the ventral tip of the gill filament. LCCs were distributed as two parallel rows along the gill lamella's outer rim and ciliary disks' outer rim (Figs. 3G and 3H) and had highly expressed genes involved in cilia structure, cilia movement and ATP synthases (Fig. 3B and Supplementary Table S3), indicating that LCCs may have a strong ability to beat their cilia. Interestingly, we identified a group of previously unreported smooth muscle cells (SMCs) co-localised with the LCCs as showed by WISH (Fig. 3I and Supplementary Fig. S4). SMCs strongly expressed several Low-density lipoprotein receptors (LDL receptor) and LDLR-associated proteins (Fig. 3B). In addition, these cells also expressed the angiotensin-converting enzyme-like protein (Bpl_scaf-29477-5.10) and the ‘molecular spring’ titin-like protein (Bpl_scaf_56354-7.8) (Supplementary Table S3). The expression of these genes could be commonly found in human vascular smooth muscle cells. Collectively, we suspect that SMCs are involved in lateral cilium movement and the gill slice contraction.

ICs are the specialised ciliary cells surrounding the bacteriocytes (Figs. 3B and 3J). WISH analysis showed that the expressions of IC markers had apparent spatial variations (Fig. 3J). The expression of ICs marker rootletin (Bpl_scaf_22362-6.9) was considerably higher in the ICs close to the frontal edge of the gill filament, and the expression gradually decreased along with the direction of inter lamina water flow (Fig. 3K and Supplementary Fig. S5), implying that ICs ventilate the water flow and the mucus through the gill filaments. Furthermore, compared with the other three types of ciliary cells, the ICs expressed several genes encoding transcription factors involved in determining cell fate (Fig. 3B and Supplementary Table S3), such as transcription factors Sox 8 and Wnt pathway cell polarity regulator.
secreted frizzled-related protein 5 (Bpl_scaf_57424-6.1)\textsuperscript{53}, suggesting that the ICs might also play regulatory roles\textsuperscript{54-56}.

**Proliferation cells:** It has long been known that the bathymodioline mussel gill has three types of proliferation cells that are conserved throughout all filibranchia bivalves: the budding zone at the posterior end of the gill where new gill filaments are continuously formed, and the dorsal and ventral ends of each gill filament\textsuperscript{57,58}. These “cambial-zone”-like cell populations could continuously proliferate throughout the whole life span of the mussel\textsuperscript{58}. Our snRNA-seq data recognised three types of proliferation cells, which is consistent with previous findings (Fig. 4A, Fig. 4B and Supplementary Table S4). The gill posterior end budding zone cells (PEBZCs) are located on the first few freshly proliferated filaments of the posterior tip of gill (Figs. 4C and 4D). The PBEZCs marker gradually disappeared at around the 11th -12th role of the gill filament, suggesting the maturation of the gill filaments, which is similar to the developmental pattern reported in another deep-sea mussel *Bathymodiolus azoricus*\textsuperscript{59}. The dorsal end proliferation cells (DEPCs), which expressed the hallmarks genes of muscular tissue, as well as cell proliferation and differentiation regulators, were the proliferation cells in connective tissue at the dorsal end of the gill slice (Fig. 4A, 4E). The ventral end proliferation cells (VEPCs) were two symmetrical triangle-shaped cell clusters of small symbiont-free cells (Fig. 4F). In VEPCs, genes encoding ribosomal proteins, chromatin proteins, RNA and DNA binding proteins, and cell proliferation markers were all up-regulated (Fig. 4B and Supplementary Table S4), indicating that VEPCs are meristem-like cells\textsuperscript{59,60}.

It has been hypothesised that new proliferation cells will be colonised by symbionts, serving as the vital mechanism of bacteriocyte recruitment\textsuperscript{59}. We then determined which proliferation cell type gives rise to the mature cells, especially the bacteriocytes, which has little available information regarding their precursor and transmission mode in bathymodioline mussels\textsuperscript{59,61}. Slingshot trajectory analysis suggested that the PEBZCs might be the origin of all gill epithelial cells, including the other two proliferation cells (VEPC and DEPC) and bacteriocytes (Supplementary Fig. S6). The sole exception was BMC2, which may be derived from VEPC rather than PEBZC. This result is consistent with previous studies which suggested new gill filaments of the filibranch mussels are formed in the gill’s posterior budding zones\textsuperscript{59}.
colonisation by the symbiont might play a crucial role in determining the fate of the bacteriocytes. Noticeably, in *G. platifrons*, only the pillar-shaped first row of gill filament, comprised of small meristem-like cells, was symbiont-free (Figs. 4D and D'), whereas all the other gill filaments were colonised by symbionts. This pattern of symbiosis establishment is different from that of *B. azoricus*, in which PEBZCs are symbiont free and are gradually colonised by symbionts released from the bacteriocytes on the adjacent mature gill filaments after maturation [59]. On mature gill filaments, the DEPCs and VEPCs are seemingly the sources of new cells that sustain the growth of the gill filament from both dorsal and ventral directions, respectively [57].

3. Bacteriocytes and host-symbiont interaction:

We conducted the whole-mount FISH using a bacterial 16S rRNA probe for symbiont to determine the spatial distribution of bacteriocytes, and their positions relative to the other cell types on the gill filament (Figs. 5A, B). Bacteriocytes covered the majority of the surface of the gill filament, except the ventral tip, the ciliary disk, and the frontal edge (lateral ciliary). The bacteriocytes were surrounded by intercalary cells with microvilli and cilia on the surface (Fig. 5C). Noticeably, the bacteriocytes’ size and FISH signal intensity, which reflected the symbiont number in each bacteriocyte, gradually decreased from the heavily ciliated frontal edge to the inner edge, demonstrating that the symbiont abundance in each bacteriocyte was affected by the environmental gradient [11].

The gene expression profile of bacteriocytes aligned well with the morphological analysis, which suggested that the bacteriocytes have structural, metabolic, and regulatory adaptations to cope with the symbiont (Figs. 5D, 5F and 5I). It has been hypothesised that the bacteriocytes extract nutrients from the symbiont through two pathways: lysing the symbiont through the lysosome (the ‘farming’ pathway) or directly utilising the nutrient produced by the symbiont (the ‘milking’ pathway) [62,63]. Our TEM observations clearly detected intracellular vacuole and lysosome system of the bacteriocyte that harbour, transport, and digest the symbionts (Fig. 5D), which is consistent with previous studies [18,64]. The snRNA data showed bacteriocytes expressed cellular membrane synthesis enzyme (phosphoethanolamine N-methyltransferase,
Bpl_scaf_15282-3.10) and a series of lysosomal proteins such as lysosomal proteases (cathepsins; Bpl_scaf_61711-5.12, Bpl_scaf_46838-6.40 and Bpl_scaf_59648-4.5; lysosomal alpha-glucosidase-like), protease regulators 56 (TMPRSS15 proteins; Bpl_scaf_52188-1.15 and Bpl_scaf_15410-0.9), and lysosomal traffic regulator proteins (rabenosyn-5, Bpl_scaf_54816-0.3 and Bpl_scaf_52809-1.6, lysosomal-trafficking regulator-like isoform X1). Among the proteases, cathepsins were thought to be evolutionary conserved molecular tools that host utilized to control the residence of their symbiont microbes. They were also highly expressed in symbiotic tissue of other deep-sea chemosynthetic animals, such as vesicomyid clams and vestimentiferan tubeworms. Bacteriocytes also expressed genes encoding cellular vesicle transports (kinesins, Bpl_scaf_14819-0.11, Bpl_scaf_54265-1.13, and Bpl_scaf_4784-1.40), potential amino acid transporter (Bpl_scaf_36159-5.5), and genes involved in intracellular vesicle transport, such as the FYVE and coiled-coil domain-containing protein 1 (Bpl_scaf_33726-5.7). In addition to form and mobilise early endosomes, the protein products of these genes could transport symbiont-secreted nutrient vesicles to the host (Fig. 5F and Supplementary Table S5), supporting the “milking” pathway.

The symbiont of G. platifrons belongs to type I methanotrophy, of which the core metabolic function is linked with the development of intracytoplasmic membranes leading to a high lipid/biomass content. Recent lipid biomarker analyses showed that the gill of G. platifrons contains a high amount of bacterial lipids, which are directly utilised by the host to synthesise most of its lipid contents. Downstream of the bacteriocyte’s metabolic cascade, genes encoding proteins that may be involved in fatty acid/lipid metabolism, such as perilipin2 (Bpl_scaf_27158-3.8), which is the critical protein to form intracellular lipid droplets, and a variety of fatty acid metabolism enzymes (acetyl carboxylase 2, Bpl_scaf_55250-5.11; fatty acid desaturase 1-like isoform X1, Bpl_scaf_35916-0.6; long-chain fatty acid-ligase ACSBG2-like isoform X1, Bpl_scaf_28862-1.5), were up-regulated, suggesting that the fatty acid could be a major form of nutrients passing from the symbiont to the host mussel.

Additionally, bacteriocytes expressed several solute carriers, including sodium/ascorbate cotransporter (solute carrier family 23 members 1-like, Bpl_scaf_32311-1.19), sodium/potassium/calcium exchanger (sodium potassium calcium exchanger 3-like,
Bpl_scaf_14503-1.28), sodium/chloride ion cotransporter (solute carrier family 12 member 3-like isoform X1, Bpl_scaf_44604-3.7), ferrous iron transporter (solute carrier family 40 member 1-like, Bpl_scaf_63447-0.12) and zinc transporter (zinc transporter ZIP14-like, Bpl_scaf_44428-5.3). The solute carriers are a large family of ATP-dependent transporters that shuttle a variety of small molecules across the cellular membrane. In several model symbiotic systems, solute carriers play a vital role in host-symbiont interaction by either providing the symbiont with substrates or transporting symbiont-produced nutrients to the host. Previous studies demonstrated that solute carrier genes were expanded in deep-sea chemosynthetic animals' genome and highly expressed in symbiotic organs including bathymodioline mussel's gill. In bathymodioline mussels, previous bulk RNA-seq studies detected up-regulated expression of a large variety of solute carriers in the gill, which is consistent with the present study, suggesting that solute carriers may play crucial roles in shuttling nutrients in and out of bacteriocytes and in maintaining the suitable intracellular micro-environment for the symbiont.

4. Cell-type-specific response to environmental stresses:

To examine cell-type-specific acclimatisation to environmental changes the expression profiles of differentially expressed genes (DEGs) were compared between the three states (Fanmao, starved and reconstituted animals) of samples collected in our in situ transplantation experiment (full lists of DEGs are shown in Supplementary Table S6-S18). Notably, for each state three mussels were processed but pooled for nuclei extraction before snRNA sequencing (see Method for detail). We tested our hypothesis here that at the starvation site, a relatively low concentration of methane shall upset symbiont metabolism and thus substantially affect symbiont-hosting bacteriocytes by assessing the transcriptional changes per cell type. We calculated the distances between the three sets of samples on the UMAP plot (Fig. 6A) estimating the centroid coordinates per sample per cell type, and calculating the Euclid distance between the centroid coordinates of each pair of states per cell type. The impact of starvation was variable across cell types, as reflected by the cross-state distances (Fig. 6B, green bar). Starvation resulted in the most significant transcriptional changes in bacteriocytes reflected by...
a large Fanmao-vs-starvation distance, followed by VEPC and inter lamina cells. On the other hand, starvation had little impact on the transcriptions of ciliary cells and most supportive cells such as the food groove ciliary cell, BMC2, and mucus cells. Fig 6B also shows that after reconstitution (moving the mussels back to the methane-rich site Fanmao for 3 days), the expressional profile of bacteriocytes rapidly changed back, reflected by a large starvation-vs-reconstitution distance and much smaller Fanmao-vs-reconstitution distance. This result coincided with and was supported by our pseudo-time analysis for bacteriocytes (Fig. 6C), showing that bacteriocytes in the reconstitution are in intermediate and transitional states between Fanmao and starvation.

For the bacteriocytes population, we conducted cell trajectory pseudo-timing and detailed DEG analysis to interpret the mechanism of host-symbiont interactions. The branched heatmap showed distinct gene expression patterns amongst the three treatments (Fig. 6D and supplementary Fig. S7). The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the bacteriocytes’ DEGs provides an overall view of the pathways enriched in each environmental state (Supplementary Fig. S8 A-C). The genes encoding ribosomal proteins are highly expressed in the bacteriocytes under the methane-rich ‘Fanmao’ state (Supplementary Fig. S8A), suggesting an active protein synthesis and cellular metabolism. Moreover, organic ion transporters (A BCB P-glycol, Bpl_scaf_18613-16.10; canalicular multispecific organic anion transporter 2, Bpl_scaf_52110-4.43; sodium- and chloride-dependent glycine transporter 1-like, Bpl_scaf_13642-5.8), which may be involved in transporting symbiont-produced nutrients, are also highly expressed.

In starved G. platifrons, negative regulators of cell proliferation, such as autocrine proliferation repressors and receptor-type tyrosine phosphatase beta, were up-regulated in bacteriocytes, suggesting repression of cell growth. We also observed the enrichment of genes in the apoptosis pathway (Supplementary Fig. S8B), which is attributed to the upregulation of caspases (Bpl_scaf_25165-3.8 and Bpl_scaf_24225-0.12) and cathepsins (Bpl_scaf_64706-0.16, Bpl_scaf_61711-5.12) which can trigger caspases-dependent cell death, suggesting cellular stress condition. Correspondingly, a gene encoding baculoviral IAP repeat-containing protein (Bpl_scaf_6172-0.3), which can bind caspase and inhibit apoptosis, was up-regulated in
bacteriocytes (Fig. 6E). In *Drosophila*, the baculoviral IAP proteins are important in animal’s response to cellular stress and promote cell survival\(^{88,89}\). Similarly, the gene encoding MAP3K14 (Bpl_scaf_60908-4.14) and potential E3 ubiquitin ligase RNF213 (Bpl_scaf_25983-1.26, Bpl_scaf_1894-1.45, Supplementary Fig. S9), which were apoptosis suppressor or regulator, were up-regulated\(^90\). These results suggest that starved bacteriocytes were carrying out cell-type-specific adjustments to cope with stresses. The E3 ubiquitin ligases could also work as intracellular immune sensors of bacterial lipopolysaccharides\(^91\). Thus, the encoded protein may be able to active downstream immunological toolkit to digest the symbiont population for nutrients.

As mentioned above, bacteriocytes obtain nutrients from endosymbionts. KEGG pathway analyses suggested that phagosome, lysosome-related pathways were upregulated in the starvation state, indicating that bacteriocytes were actively digesting the endosymbionts in the starved *G. platifrons*. Interestingly, protein synthesis activity was more activated in inter lamina cells and VEPCs as supported by the up-regulation of genes encoding ribosomal proteins in both cell populations (Supplementary Tables S6 and S17). This was contrary to the situation in bacteriocytes and may be a consequence of the high activity of “farming” in bacteriocytes.

We anticipated that after moving the starved *G. platifrons* back to Fanmao site, bacteriocytes and their endosymbionts would “reconstitute” leading to the partly restored “farming” and “milking” pathways. This was confirmed by higher expression of fatty acid metabolic genes in the mussels in the reconstitution state than those in the starvation state, such as long-chain-fatty-acid-ligase ACSBG2-like (Bpl_scaf_28862-1.5), elongation of very long-chain fatty acids 7-like (Bpl_scaf_5959-1.9), and fatty acid desaturase 1-like (Bpl_scaf_35916-0.6). These findings were also consistent with the result of KEGG analyses. The mitochondrial trifunctional enzyme (Bpl_scaf_28862-1.5) was also highly expressed, suggesting a high level of energy-producing activity. The KEGG analyses showed that glutamatergic synthase was enriched in the reconstitution state in comparison with those in both the Fanmao and starvation states. In the insect aphid–Buchnera endosymbiosis model, the host-produced glutamate could be transported and directly utilised by the symbiont\(^{92,93}\). Similar glutamate-based host–symbiont metabolic interaction mechanisms were also proposed in the deep-sea...
mussel *Bathymodiolus thermophilus* and *B. azoricus*. As the pseudotime analyses showed above (Fig. 6C), reconstitution was the intermediate state between Fanmao and starvation. Metabolic and gene regulatory functions of bacteriocytes were still different from that in Fanmao. The KEGG analyses suggested regulatory (Rap1 signalling, retrograde endocannabinoid signalling, chemokine signalling, glucagon signalling, thyroid hormone synthesis, mRNA surveillance pathway, etc.) and metabolic pathways were enriched in the reconstitution state (fatty acid metabolism, salivary secretion, aldosterone synthesis and section, insulin secretion and pancreatic section) (Supplementary Fig. S8C). For example, we detected up-regulation of the genes encoding carbonic anhydrases (Bpl_scaf_33596-7.3 and Bpl_scaf_48274-0.3), electroneutral sodium bicarbonate exchanger 1 (Bpl_scaf_61230-5.11), and globin-like proteins (Bpl_scaf_50392-1.14 and Bpl_scaf_24370-0.5), which could provide the symbiont with carbon dioxide and oxygen necessary for the symbiont’s chemosynthetic metabolism.

5. Summary and outlook:

Using the deep-sea mussel *G. platifrons* as a model organism, we demonstrated the power of integrating snRNA-seq and WISH data in unravelling the mechanism behind animal-microbe symbiosis. The robustness of this strategy showed stable and highly distinguishable expression patterns of each cell type regardless of the different environmental states. We successfully profiled the specific roles of different types of cells, including the previously unknown cell types, in maintaining the structure and function of the gill. The supportive cells that located in between (inter lamina cells) and on the basal membrane’s surface (BMC1 and BMC2) helped maintain the anatomical structure of the basal membrane. Ciliary and smooth muscle cells involved in gill slice contraction and cilia beat allowing the gathering of material and information from the surrounding environment. Proliferation cells (PBEZCs, DEPCs and VEPCs) gave rise to new cells, including bacteriocytes which obtained nutrients from endosymbionts using intracellular vacuoles and lysosomes.

The snRNA analyses also revealed that different cell types collaborated to support bacteriocytes' functionality. The PBEZCs gave rise to new bacteriocytes that allowed symbiont colonisation. Bacteriocytes attached to the basal membrane and were stabilised by supportive
Mucus cells co-localise with bacteriocytes and help to maintain immune homeostasis. The beating ciliary cells controlled the water flow, providing bacteriocytes with necessary inorganic substances from the environment. This new information on cell-cell interaction certainly advanced our overall understanding of how endosymbiotic microbes and host cells communicate and collaborate, which cannot be easily achieved through other methods.

Moreover, the analysis of snRNA data from bacteriocytes has provided insight into the molecular mechanisms employed by the host to maintain and regulate the symbiosis. Notably, the bacteriocytes-enriched transcripts involved in harbouring, digesting symbionts, and transporting nutrients produced by the symbionts were identified and characterised. Our in situ transplant experiments by moving mussels between methane-rich and methane-limited sites also provided clues of cell-type specific responses to environmental change. Under a methane-limited environment, the staved mussels more actively consumed endosymbionts through the "farming" pathway. After being moved back to the methane-rich environment, mussels produced more glutamates which sustained the regrowth of symbionts. These preliminary findings showed that the deep-sea mussels were able to control their endosymbionts using a set of genes in response to environmental change. Due to the limitation of remotely operated vehicle (ROV) dives and sampling capacity, we had to have pooled samples of each state for nuclei extraction and sequencing. Thus, the cells per cluster could be considered as technical than biological replicates. Although this sampling process strategy has been broadly used in snRNA/scRNA sequencing, we recognise the possible violation of assumptions in p-value calculation for DEGs between the three states.

Overall, the single-cell spatial and functional atlas developed in the present work will decipher some common principles of symbiosis and environmental adaption mechanisms of animals. The workflow developed in the present study could provide insightful references for researchers focusing on the mechanistic study of the biological adaptation of biologically and ecologically important non-model animals.

Acknowledgements: This study was supported by the Science and Technology Innovation Project of Laoshan Laboratory (Project Number No. LSKJ202203104), the National Natural
Author contributions: W.H., K.C., P.Y.Q. and L.C.L designed the study. Z.H., L. Cao, Z.Z.S., L. Chao and W.M.X conducted the in situ transplant experiments and collected the samples; Z.Q.Y. and W.H. constructed the snRNA-seq library. H.K. performed the bioinformatics analysis. W.H., Z.H., L. Cao, J.L., H.C. and L.Z. performed the experiments. This paper was written by W.H., H.K., K.C., P.Y.Q. and L.C.L. All authors revised the manuscript. All authors read, approved, and contributed to the final manuscript.

Competing interests: The authors declare no competing interests.
**Figure legends**

**Figure 1:** Identification of 14 cell types in the gill of deep-sea symbiotic mussel *Gigantidae platifrons*. A: Overall experimental scheme describing the deep-sea in situ transplant experiment, the sample preparation procedures and the single-cell analysis and validation pipeline. Three *G. platifrons* samples were included in the present study: 'Fanmao,' Starvation and Reconstitution. The cell nucleus was extracted from each sample, which included a pool of gill posterior tip of three mussels. The snRNA-seq libraries were constructed according to the BD Rhapsody single-nuclei 3′ protocol. Cell population-specific markers were validated by WISH and ISH. B: The image shows the posterior end of the gill of *G. platifrons*. C: UMAP representation of *G. platifrons* gill single cells. Cell clusters are coloured and distinctively labelled. D: Heat map profile of markers in each cluster. The colour gradient represents the expression level of every single cell.

**Figure 2:** Supportive cell populations of *G. platifrons* gill. A: UMAP representation of the four supportive cell populations. B: Expression profiles of the cell markers that are specific or enriched in the supportive cell populations. The sizes of the circles represent the percentages of cells in those clusters that expressed a specific gene. Genes shown in red were validated by WISH or double FISH. C and C': Schematics demonstrating the overall structural (panel C) and supportive cell distribution (C'). D-F: WISH characterisation of the selected representative cell population markers. H-K: Double FISH characterisation of the selected representative cell population markers. The marker genes confirmed by WISH and ISH in the current study are indicated in red. Scale bar: 50µm.

**Figure 3:** Ciliary cell populations of *G. platifrons* gill. A: UMAP representation of the four ciliary cell populations and potential smooth muscle cell population. B: Expression profiles of the cell markers that are specific or enriched in the ciliary cell populations. The sizes of the circles represent the percentages of cells in those clusters that expressed a specific gene. The genes shown in red were validated by WISH or double FISH. C-E, G and I-J: WISH characterisation of the selected representative cell population markers. F and H: SEM analysis of the ciliary cells of
**G. platifrons gill.** K: Schematic of the water flow agitated by different ciliary cells. The marker genes confirmed by WISH in the current study are indicated in red.

**Figure 4:** Proliferation cell populations of *G. platifrons gill.* A: UMAP representation of the three proliferation cell populations. B: Expression profiles of the cell markers that are specific or enriched in the supportive cell populations. The sizes of the circles represent the percentages of cells in those clusters that expressed a specific gene. Genes shown in red were validated by WISH or double FISH. C and D: Photographic and schematic analyses of the spatial position of the three Proliferation cell populations. E, F and J: FISH and WISH characterisation of the selected population markers. The marker genes confirmed by ISH or WISH in the current study are indicated in red. Scale bar: 50 µm.

**Figure 5:** Characterisation of the bacteriocytes of *G. platifrons.* A: Schematic of the overall structure of *G. platifrons gill* filaments. B: Whole-mount FISH analyses of the overall distribution of bacteriocytes on the *G. platifrons gill* filament. C: SEM analysis of the bacteriocytes. D: TEM analysis of a bacteriocyte. E: UMAP representation of *G. platifrons* bacteriocytes. F: Expression profiles of the cell markers that are specific or enriched in the bacteriocytes. The sizes of the circles represent the percentages of cells in those clusters that expressed a specific gene. Double FISH validated the genes shown in red. I: Schematic of the host-symbiont interaction based on the single-cell transcriptome of *G. platifrons* bacteriocytes. The marker genes confirmed by ISH in the current study are indicated in red.

**Figure 6:** Analysis of cell population-specific DEGs. A: UMAP representation of the impact of deep-sea *in situ* transplant treatments on the gene expression pattern of each cell population. The cells from different treatments were labelled with different colours. The dashed line encircled bacteriocyte populations have a considerably altered expression profile. B: Histogram of cross-state distances between the centroids of the Fanmao, Starvation and Reconstitution groups per cell type on UMAP. The black dashed lines indicate the bacteriocyte populations whose expression profile was remarkably altered. C: Visualization of bacteriocytes onto the
pseudo time map using monocle. The black lines indicate the main path of the pseudotime ordering of the cells. D: Bifurcation of selected gene expression along two branches in response to environmental perturbation. Genes are clustered hierarchically into two groups, illustrating up- (cluster 1) and down- (cluster 2) regulated genes in the starvation state compared with Fanmao. Genes in red colour were discussed in Section 4. The heat map showing the gene expression profiles of all bacteriocytes' DEGs is shown in supplementary Fig. S7; E: Proposed model for the molecular mechanisms of host-symbiont interactions in response to environmental changes.

Materials and methods:

Deep-sea in situ transplant experiment and sample collection: in situ transplant experiment was conducted at the 'F-site' cold seep during the R/V 'Kexue' 2020 South China Sea cold seep cruise. The overall design of the in situ transplant experiment and environmental states are shown in Figure 1. First, mussels in the methane-rich 'Fanmao' site (meaning 'prosperous' site, 22°06′55.425″N, 119°17′08.287″E, depth 1117 m) were scoped into three nylon bags with approximately 10 mussels in each bag. Then, two bags of mussels were transplanted to the low methane Starvation site (Figure 1, 22°07′00″N, 119°17′07.02″E, depth 1147.42 m).

After 11 days of transplantation, one bag of mussels in the Starvation site was moved back to the 'Fanmao' site. On the 14th day of the transplant, three bags of mussels: one bag from the Fanmao site (designated as the 'Fanmao' sample), one bag from the starvation site (designated as the 'Starvation' sample) and one bag of mussels which were first transplanted to the Starvation site for 11 days and moved back to the Fanmao site for 3 days (designated as the 'Reconstitution' sample) were all retrieved by ROV Fanxian. The mussels were kept in a hydraulic pressure-sealed biobox during the ascending of the ROV. The biobox is made of heat-insulated material, which will prohibit heat exchange with warm surface water. The samples were immediately processed once onboard R/V Kexue. For snRNA-seq, the posterior end tip of the mussel's gill was dissected (Figure 1C), snap-frozen with liquid nitrogen and then stored at −80°C until use. For WISH, ISH and FISH, the gills of the mussels were first fixed with 4%
paraformaldehyde (PFA, prepared with autoclaved 0.22 μM membrane filtered in situ seawater) at 4 °C overnight. Then, the gill was washed with ice-cold 1× PBS three times, dehydrated and stored in 100% methanol at −20 °C.

Single-nucleus RNA-sequencing of *G. platifrons*: The gill nucleus was extracted using the Nuclei PURE prep nuclei isolation kit (Sigma-Aldrich). For each sample, posterior end tips from 3 individual mussels were randomly selected and pooled together. The posterior tips were homogenised in 10 mL of ice-cold lysis solution (Nuclei PURE lysis buffer with 0.1% Triton X100 and 1 mM DTT) on ice. The cell nuclei were then separated by sucrose gradient centrifugation according to the manufacturer's protocol. The cell nuclei pellets were washed by re-suspending in ice-cold DPBS–BSA solution (1× DPBS, 0.04% nuclease-free BSA, 0.01% RNase inhibitor; Takara). The nucleus was spun down by centrifugation at 500×g for 5 min at 4 °C. This step was repeated to remove the contaminants from the cell plasma. Finally, the nucleus was re-suspended in the DPBS–BSA solution. The concentration of the cell nucleus was counted by Cell Countess II. Single-nucleus RNA-seq libraries were then constructed with the BD Rhapsody™ single-cell analysis system using the BD Rhapsody WTA Amplification Kit according to the manufacturer's protocol. The libraries were then sequenced with the Illumina HiSeq 4000 platform. The clean reads of three datasets were submitted to the National Centre for Biotechnology Information Sequence Read Archive database (Bioproject: PRJNA779258).

Bioinformatics: Three raw datasets (Fanmao, Starvation and Reconstitution) were processed individually by following the BD single-cell genomics analysis setup user guide (Doc ID: 47383 Rev. 8.0) and BD single-cell genomics bioinformatics handbook (Doc ID: 54169 Rev. 7.0). In brief, a *G. platifrons* genome reference v1.0 (the genome could be downloaded at Dryad Digital Repository http://dx.doi.org/10.5061/dryad.h9942) was indexed and built using STAR 2.5.2b. Gene names were prepared according to the BD-single cell guideline. The sequencing data were then mapped to the indexed genome using the BD Rhapsody single-nucleus pipeline v1.9 with default parameters to generate an expression matrix per dataset.
Each data matrix was converted to a SingleCellExperiment data format, and empty barcodes were removed using the emptyDrops function of DropletUtils v3.14. Then, we converted and processed the data in Seurat v3.28. We removed cells that had <100 or > 2500 genes and <100 or > 6000 unique molecular identifiers (UMIs). We also removed genes that had <10 UMIs in each data matrix. Then, we log-normalised the data and used DoubletFinder v2.0 to remove potential doublet, assuming a 7.5% doublet formation rate. The numbers of retained nuclei were 9717, 21614 and 28928 for Fanmao, Starvation and Reconstitution data, respectively (Supplementary Table S19). We used the top 3000 highly variable genes for principal component analysis (PCA) and a reciprocal PCA approach to integrating the three datasets.

The first 40 principal components (PCs) were used for uniform manifold approximation and projection (UMAP) dimensional reduction and following clustering (Supplementary Fig. S10). We used an empirical parameter, a resolution of 0.2, in the FindClusters function. The average expression of each gene per cluster were also provided as supplementary data S1. Based on the WISH results (see below), we manually combined the clusters with similar gene expressions and WISH patterns. Eventually, we assigned all the cells to 14 reliable clusters (Supplementary Fig. S11). We adopted a bootstrap sampling strategy to evaluate the stability for each cluster. We ran bootstrapping with 100 iterations using cells in the three samples together and in each sample individually. We determined marker genes per cluster using the FindAllMarkers function based on the Wilcoxon rank sum test. The marker genes were used for KEGG enrichment using ClusterProfiler v4.2.

Cell trajectory analysis: We conducted slingshot trajectory analyses using Slingshot v2.2.1 for a subset of clusters ('intercalary cells,' 'dorsal-end proliferation cells,' 'ventral-end proliferation cells,' 'mucus cells,' 'basal membrane cells 1,' 'basal membrane cells 2' and 'bacteriocytes') to explore the developmental trajectory of cells, assuming that all these cells were developed from the same precursor ('posterior-end budding zone'). We performed PCA for the subsampled data set and conducted dimensional reduction using phateR v1.0.7 for slingshot analyses. We used the Potential of Heat-diffusion for Affinity-based Trajectory Embedding (PHATE) because it could better reveal developmental branches than other tools.
We also examined the effect of deep-sea transplant experiments on shaping gene expression patterns by comparing the expression levels amongst the three different states of a given cluster. This comparison was done for the bacteriocytes. The biased number of cells per state could affect the results of the dimensional reduction and calculation of marker genes, and the sequenced nucleus per state was unbalanced; therefore, we first downsampled the cells per cluster per state to about the same number. Thereafter, we performed PCA and dimensional reduction using UMAP and PHATE, calculated marker genes and conducted slingshot trajectory and KEGG enrichment analyses as mentioned above for each cell type.

**Phylogenetic estimation:** Phylogenetic estimation was conducted for E3 ubiquitin ligase RNF213 genes. We downloaded RNF213 genes from GenBank for representative vertebrate and invertebrate species across the tree of life. The amino acid sequences were aligned with the G. *platifrons* sequences annotated as E3 ubiquitin ligase (Bpl_scaf_25983-1.26 and Bpl_scaf_1894-1.45) using MAFFT v7.450. The alignment was used to estimate a maximum-likelihood best gene tree and to calculate bootstrap values on each node using RAxML v8.2.

**G. platifrons gill fixation and storage:** The gill tissues of *G. platifrons* were dissected within minutes after the ROV, and samples were retrieved on board R/V Kexue. The gill tissues were briefly washed with ice-cold filtered and autoclaved *in situ* seawater (FAISW) and then fixed in 4% PFA prepared in FAISW at 4 °C overnight. The gill tissues were washed three times with ice-cold PBST, dehydrated in 100% methanol and stored at -20 °C until use.

**Whole-mount in situ hybridisation:** For WISH and double FISH analyses, the DNA fragments (~1000bp) of the targeted genes were first PCR amplified with gene-specific primers (GSPs) pairs using *G. platifrons* gill cDNA as template (the sequences of targeted genes and gene-specific primers were provided in Supplementary Data S2). The amplified fragments were ligated into the pMD18-simpleT vector (Takara) and transformed into *E. coli*. Individual colonies were picked up, and their plasmids were sequenced to confirm the inserts. The templates for in vitro mRNA transcription were amplified using T7 forward GSP (sense probe control) or Sp6
reversed GSPs (antisense probe) combined with either forward or reversed gene-specific primer.

Labelled probes and control probes were generated using digoxigenin (DIG)-12-UTP (Roche) or fluorescein-12-UTP (Roche) according to the protocol described by Thisse et al. with Sp6 and T7 RNA polymerase, respectively.

For WISH, the connective region at the end of the W-shaped gill filament was cut off, and each gill slice was carefully peeled off with fine-tip tweezers. We dissected gill tissues from 5 individual mussels and pooled all the gill slices together. The gill slices were then rehydrated in 75%, 50% and 25% methanol-PBST (1×PBS with 0.1% Tween 20) for 15 min each, followed by 3×5 min PBST washes. The gill slices were then permeabilised with 2 μg/mL proteinase K in PBST for 30 min at 37 °C. Post-digestion fixation was conducted by fixing the gill slices with 4% PFA in PBST for 30 min at room temperature (RT). After 3×5 min PBST wash to remove the residual fixative, the gill slices were pre-hybridised with a hybridisation mix (HM; containing 50% formamide, 5×saline–sodium citrate (SSC), 0.1% Tween 20, 10 μg/mL heparin, 500 μg/mL yeast tRNA) for 1 h at 65 °C. For each hybridisation, 5-10 gill slices were added to 400 μL of fresh HM containing ~0.5 ng/μL DIG-labelled probe. Hybridisation was conducted in a 55 °C shaking water bath overnight. Post-hybridisation washes were performed according to the following steps: the gill slices were first washed 3×15 min with hybridisation washing buffer (50% formamide, 5×SSC, 0.1% Tween 20), followed by 3×15 min 2×SSC with 0.1% Tween 20 and 3×15 min 0.2×SSC with 0.1% Tween 20. The washings were also conducted in a shaking water batch, and all the washing buffers were pre-heated to 55 °C. The samples were washed three times with PBST and then blocked in blocking buffer (2.5% sheep serum, 2% BSA in PBST) for 1 h at RT. Each sample was incubated with 1:10,000 diluted anti-DIG-AP antibody (Roche) at 4 °C overnight. The samples were incubated with the antibody, then washed for 6×15 min with PBST, followed by washing in 3×15 min alkaline Tris buffer (100 mM Tris-HCl, pH 9.5; 100 mM NaCl; 50 mM MgCl$_2$). The samples were incubated in nitro blue tetrazolium/5-Bromo-4-chloro-3-indolyl phosphate staining solution (Sangon). After the desired expression pattern was revealed, the staining reaction was stopped by 3×15 min PBST–EDTA wash (PBST, 1 mM EDTA). The gill slices were cleared by incubating in 100% glycerol overnight at 4 °C and then mounted on glass slides. The results of control hybridisations (with sense probes) were provided in Supplementary Fig.
S12-S15. WISH analyses were repeated with another batch of *G. platifrons* gill slices samples collected during the R/V ’Kexue’ 2017 South China Sea cold seep cruise to confirm the consistency in expression patterns.

**Paraffin embedding and double fluorescent in situ hybridisation:** The methanol-dehydrated gill slices were incubated in 100% ethanol, a 1:1 mixture of 100% ethanol and xylene and Xylene twice for 1 h each at RT. The samples were embedded by incubating in Paraplast Plus (Sigma-Aldrich) for 2 h at 65 °C and then cooled down to RT. Sections with 5 μM thickness were cut using a microtome (Leica).

For double FISH, sections were dewaxed by incubating in xylene twice, a 1:1 mixture of 100% ethanol and xylene, 100% ethanol twice, 95% ethanol, 85% ethanol and 75% ethanol for 15 min each at RT. The sections were washed with PBST three times for 10 min each and then permeabilised by 2 μg/mL proteinase K (NEB) in PBST for 15 min at RT. Post-digestion fixation was conducted by incubating the sections in 4% PFA in PBST for 30 min at RT. The sections were washed three times with PBST for 15 min each. Pre-hybridisation was conducted by incubating the sections in HM for 1 h at 55 °C. Then, the *in situ* hybridisation was performed by incubating the sections in ~0.5 ng/μL fluorescein (Roche)-labelled probe prepared in fresh HM overnight at 55 °C. The sections were washed three times with 2×SSC for 15 min each at 55 °C, cooled down to room temperature and washed three times with PBST.

A second-round hybridisation was conducted on the DIG-labelled oligonucleotides to label the symbiont. The slices were hybridised for 1 h at 46 °C with 100 ng DIG-labelled *G. platifrons* symbiont-specific probe in FISH buffer (0.9 M NaCl, 0.02 M Tris-HCl, 0.01% sodium dodecyl sulphate [SDS] and 30% formamide). The slices were then washed with FISH washing buffer (0.1 M NaCl, 0.02 M Tris HCl, 0.01% SDS and 5 mM EDTA) three times at 5 min each at 48 °C. The slices were washed with PBST three times and then blocked with blocking buffer (2.5% sheep serum and 2% BSA in sheep serum) for 1 h at room temperature.

The slices were then incubated with 1:1000 diluted anti-fluorescein–peroxidase (POD) (Roche) overnight at 4 °C, then washed six times with PBST for 15 min each and three times with TNT buffer (100 mM Tris-HCl, pH 7.5; 100 mM NaCl; 0.1% Tween 20) for 15 min each. Afterwards,
the fluorescent signal of the *G. platifrons* gene expression pattern was developed by the TSA fluorescent kit (Akoya Biosciences) according to the manufacturer’s protocol. The slices were washed three times, and the remaining POD activity was quenched by incubation in 1% hydrogen peroxide solution for 1 h at room temperature. Then, the slices were washed three times with PBS, blocked with blocking buffer for 30 min at room temperature, incubated with 1:2500 diluted anti-DIG–POD (Roche) for 2 h at RT and washed with PBST for six times and TNT for three times. The symbiont FISH signal was developed using the TSA Cy3 kit (Akoya Biosciences). Finally, the slices were washed with PBST, stained with DAPI, and mounted with ProLong Diamond Antifade Mountant (Thermo Fisher).

**Microscopy imaging:** All the WISH samples and whole-mount 16S FISH images were observed and imaged with a Nikon Eclipse Ni microscope with a DS-Ri2 camera. The double FISH slides were imaged with a Zeiss LSM710 confocal microscope.

** Electron microscopy analysis:** The gill slices of the *G. platifrons* were dissected and fixed in electron microscopy fixative (2.5% glutaraldehyde and 2% PFA) at 4 °C. For SEM analysis, the samples were dehydrated in a graded ethanol series and then dried at the critical point. The samples were then coated with gold (sputter/carbon Thread, EM ACE200) and observed under a scanning electron microscope (VEGA3, Tescan). For TEM analysis, the samples were rinsed with double distilled water, post-fixed with 1% osmium tetroxide and 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 (70 nm thickness, Reichert-Jung Ultracut E). The sections were then double-stained with lead citrate and uranyl acetate. The cells were observed under a transmission electron microscope (JEM1200, Jeol) operated under 100 kV.
References


42 Louka, P. et al. Proteins that control the geometry of microtubules at the ends of cilia. *J. Cell Biol.* **217**, 4298-4313 (2018). [https://doi.org/10.1083/jcb.201804141](https://doi.org/10.1083/jcb.201804141)


