

Function and Development of Deep-sea Mussel Bacteriocytes Revealed by snRNA-seq and Spatial Transcriptomics

Hao Chen (陈浩)^{1,*,#}, Mengna Li (李梦娜)^{1,2,#}, Zhaoshan Zhong (钟兆山)^{1,*,#}, Inge Seim³,
Minxiao Wang (王敏晓)^{1,*,#}, Chao Lian (连超)¹, Lianhong Zhuo (卓联鸿)¹, Xinjiang Wan (宛新江)¹,
Hao Wang (王昊)¹, Guanghui Han (韩广辉)¹, Li Zhou (周丽)¹, Huan Zhang (张垲)¹,
Lei Cao (曹磊)¹, Chaolun Li (李超伦)^{4,5,*}

¹Center of Deep Sea Research, Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

²National Deep Sea Center, Qingdao 266071, China

³Marine Mammal and Marine Bioacoustics Laboratory, Institute of Deep-sea Science and Engineering, Chinese Academy of Sciences, Sanya 572000, China

⁴South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

⁵Laoshan Laboratory, Qingdao 266071, China

*Corresponding authors: lcl@qdio.ac.cn (Li C), wangminxiao@qdio.ac.cn (Wang M), chen hao@qdio.ac.cn (Chen H).

[#]Equal contribution.

Handling Editor: Fangqing Zhao

Abstract

Deep-sea chemosynthetic ecosystems are among the most unusual ecosystems on Earth, where most megafauna form close symbiotic associations with chemosynthetic microbes to obtain nutrition and shelter from the toxic environment. Despite the diverse forms of symbiotic organs in these deep-sea holobionts, the function and development of bacteriocytes, the host cells harboring symbionts, are still largely uncharacterized. Here, we conducted an *in situ* decolonization assay and state-of-the-art single-nucleus and spatial transcriptomic analyses to reveal the function and development of deep-sea mussel bacteriocytes. Bacteriocytes appear to optimize immune processes to facilitate the recognition, engulfment, and elimination of endosymbionts. They also interact directly with endosymbionts in carbohydrate and ammonia metabolism by exchanging metabolic intermediates via transporters such as SLC37A2 and RHBG-A. Bacteriocytes arise from three different proliferation cell types, and their successive development trajectories were delineated using multi-omics data and 3D reconstruction analyses. The molecular functions and developmental processes of bacteriocytes are guided by the same set of molluscan-conserved transcription factors and may be influenced by endosymbionts through sterol metabolism. The coordination in the functions and development of bacteriocytes, and between the host and symbionts, highlights the phenotypic plasticity of symbiotic cells, and underpins host-symbiont interdependence in adaptation to the deep sea.

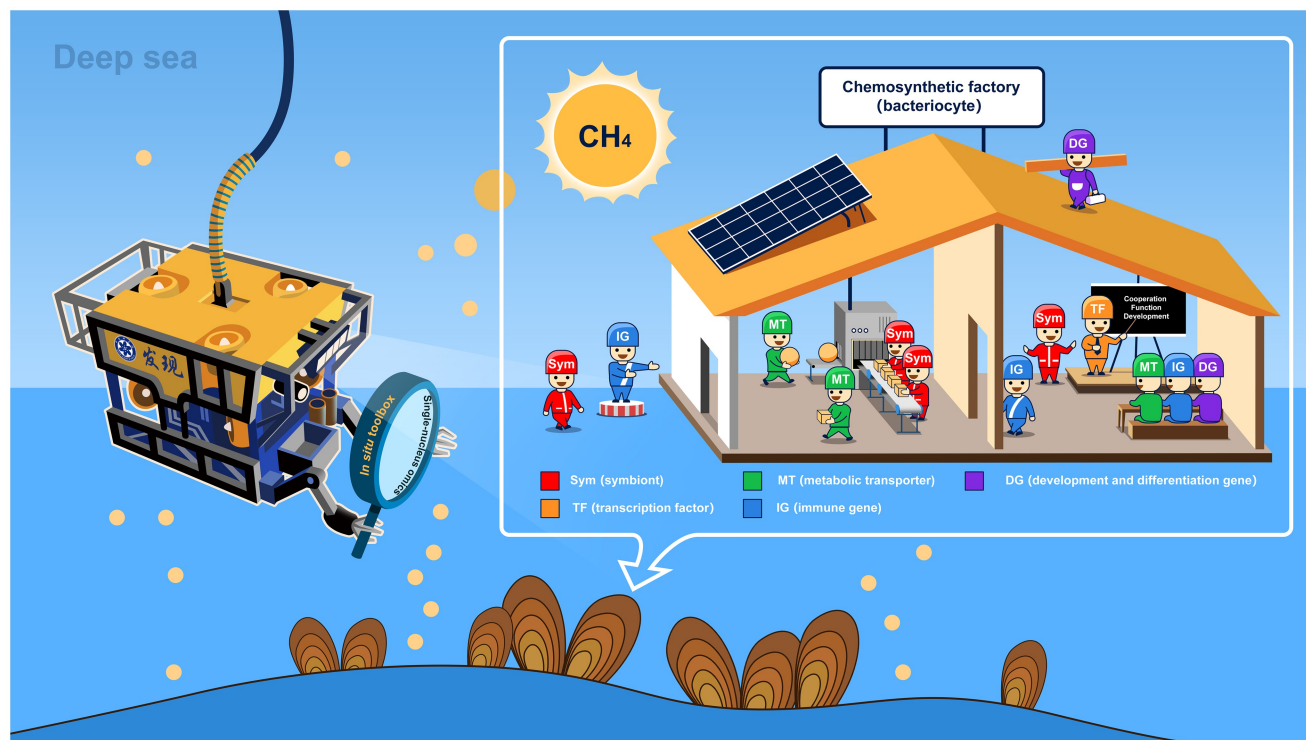
Key words: Chemosymbiosis; Host-symbiont interaction; Metabolic interaction; Immune regulation; Development plasticity.

Received: 20 October 2024. Revised: 3 July 2025. Accepted: 16 November 2025.

© The Author(s) 2025. Published by Oxford University Press and Science Press on behalf of the Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Graphical abstract



Introduction

Symbiosis is a major driver of the acquisition of novel adaptive traits, expansion of ecological ranges, and shaping of biodiversity and evolution in eukaryotes [1]. To establish symbiotic associations, most animal hosts have evolved highly complex symbiotic cells and organs (e.g., the aphid bacteriocytes, the bobtail squid light organ, and the tubeworm trophosome) that enable mutualistic relationships. Understanding the functions and development of symbiotic cells and organs is necessary to comprehend the formation and evolution of symbiosis [2]. With a growing number of studies, the functions and symbiotic interactions within symbiotic cells and organs are becoming clear in some model holobionts [3–5]. However, characterizing the function and development of non-model holobiont symbiotic cells and organs is challenging [6]. Furthermore, whether symbionts influence these functions and developmental programs, and by what mechanisms, remains debated and may vary greatly across species [7–9].

Since their discovery in 1977, chemosynthetic ecosystems in deep-sea cold seeps and hydrothermal vents have attracted considerable attention [10]. Notably, most endemic invertebrates in these ecosystems have evolved close symbiotic associations with chemosynthetic bacteria to obtain nutrients [11]. Mollusks, especially bivalves from the Mytilidae, Vesicomyidae, Solemyidae, Thyasiridae, and Lucinidae families, are particularly interesting because their chemosymbioses vary with the location of symbionts (extracellular or intracellular), type of symbiont (methanotroph, thiotroph, or both), and mode of transmission (vertical or horizontal) [12]. Moreover, whereas some mollusks develop bacteriocytes in their gills to host chemosynthetic endosymbionts, others may lose

their symbionts in certain environments and undergo a reversal toward a non-symbiotic state, akin to their non-symbiotic shallow water counterparts [13]. For these reasons, mollusks are considered promising models for investigating chemosymbiosis — particularly the function and development of symbiotic cells and organs.

Among deep-sea mollusks, Bathymodiolinae mussels are known to harbor endosymbiotic methanotrophs and/or thiotrophs in their specialized gill epithelial cells (bacteriocytes) [14,15]. Adult deep-sea mussels contain 100 million to 1 billion bacteriocytes, each harboring hundreds of methanotrophic or thousands of thiotrophic endosymbionts [16]. Fine-scale metagenomic and meta-transcriptomic sequencing has revealed within-species diversity of symbiotic bacteria and functional differences among different symbiont strains [17,18]. Pathways and genes related to endocytosis, lysosomes, and central carbon metabolism have also been implicated by bulk RNA sequencing (RNA-seq) or proteomics of mussel gills, highlighting the potential metabolic and immune interactions between mussels and endosymbionts [19–21]. Additionally, new bacteriocytes can arise from proliferation cells adjacent to growth zones, outlining their developmental trajectory in the gill [22]. Interestingly, the structure of newly formed bacteriocytes changes drastically after symbiont infection, suggesting a potential role for symbionts in bacteriocyte function and development [22,23]. Although mussel bacteriocytes have been isolated in some studies, their precise functions and developmental processes remain insufficiently characterized due to the extensive heterogeneity of gill tissue and limitations in the isolation and *in vitro* culture of bacteriocytes and endosymbionts [24]. Recent

advances in single-cell and spatial transcriptomics have allowed these issues to be addressed in non-model organisms without defined host cell lineages and culturable symbionts [25–27]. Recently, we constructed a comprehensive cell atlas of the gill of the methanotrophic deep-sea mussel *Gigantidas platifrons* using single-nucleus RNA sequencing (snRNA-seq) on the BD Rhapsody platform and characterized the expression pattern of mussel bacteriocytes [28]. While our data highlighted the unique functions of mussel bacteriocytes, how proliferation cells differentiate into bacteriocytes and whether a coordinated network controls their functions and development remain to be further addressed. Using high-resolution single-nucleus and spatial transcriptomic analyses, we conducted a comprehensive analysis, including a phagocytosis assay, 5-ethynyl-2'-deoxyuridine (EdU) labeling, and 3D electron microscopy to characterize the molecular functions and developmental trajectory of bacteriocytes. We also employed an *in situ* decolonization assay to compare our data with decolonized mussels and reveal the regulatory network governing the functions and development of symbiotic cells, as well as the potential influence of symbionts on this network.

Results

Gill cell lineage landscape

To fully capture the expression patterns of key cell lineages, such as bacteriocytes and proliferation cells, we performed snRNA-seq on the 10x Genomics Chromium platform and spatial transcriptomics sequencing (ST-seq) on the 10x Genomics Visium platform using two groups of deep-sea mussels (Figure 1A). Mussels collected from the seepage region, where methane concentration is up to 31,227 ppm ($n = 14$), were used to represent mussels in a fully symbiotic state (designated as the InS group). Mussels that were transplanted *in situ* to a low methane geographic region for 604 days, with a methane concentration of approximately 800 ppm (39-fold lower than the seepage region) and an average 76.5% decrease in endosymbionts (calculated by quantitative real-time PCR of the methane monooxygenase gene in comparison with the relative DNA content) were used to represent partially decolonized mussels (designated as the DeC group; $n = 5$; Figure 1A, Figure S1A and B). Four mussels from the InS (three mussels) and DeC (one mussel) group with similar body size were randomly selected and subjected to snRNA-seq; two of these mussels (one mussel from each group) were also subjected to ST-seq. To include more possible cell types, the dorsal-middle region of the gill, which contains both descending and ascending gill filaments and is adjacent to the posterior end of the mussel, was employed in snRNA-seq. The cross-sectioned middle region of the gill, viewed dorsally, was used for ST-seq (Figure 1A).

Principal coordinate analysis (PCoA) of the meta-transcriptomic data for the two groups showed strong variation among individuals (Figure S1C and D). Approximately 57,776 cells (average 32,458 reads per cell) and over 3600 barcoded spots (average 172,705 reads per spot) were obtained for the snRNA-seq samples (four in total) and ST-seq samples (two in total), respectively (Table S1). Transcript coverage profiling

showed that the sequencing depth of all samples was saturated on both sequencing platforms (Figure S1E–H), which was critical for subsequent analyses.

As expected, given our sampling strategy, more cell clusters (9 out of 13) were observed in both sample groups (InS and DeC) in the snRNA-seq dataset than in the ST-seq dataset (Figure 1B and C, Figure S2A). Among the 13 snRNA-seq cell clusters, 12 shared marker genes were reported in our previous study [28]. By comparing the shared marker genes with those from our previous study, we successfully identified bacteriocytes (cluster 0), three types of proliferation cells [dorsal end proliferation cells (DEPCs, cluster 1), posterior end budding zone cells (PEBZCs, cluster 2) and ventral end proliferation cells (VEPCs, cluster 6)], four types of ciliary and smooth muscle cells [intercalary cells (ICCs, cluster 3), lateral ciliary cells (LCCs, cluster 8), food groove ciliary cells (FGCCs, cluster 10), and smooth muscle cells (SMCs, cluster 12)], and four types of supportive cells [basal membrane cell 1 (BMC1, cluster 5), BMC2 (cluster 9), mucus cell (MC, cluster 4), and inter lamina cells (ILC, cluster 11)]. In addition, the median number of identified genes per cell and the total number of candidate marker genes both improved substantially in comparison with our previous study — where only 631 identified genes per cell per sample and 92 marker genes for bacteriocytes with at least 1.28-fold upregulated expression levels and $P < 0.01$ in the target cell cluster compared to the rest clusters were reported and with scRNA-seq of other deep-sea invertebrates [28,29]. In particular, the median number of identified genes per cell reached up to 900 and 2600 genes per cell/spot for snRNA-seq and ST-seq of the InS sample, respectively (Table S1); a total of 321, 81, 88, 173, 119, 147, 264, 113, 155, 177, 188, and 57 genes were identified as the marker genes of bacteriocyte, DEPC, PEBZC, VEPC, ICC, LCC, FGCC, SMC, BMC1, BMC2, ILC, and MC, respectively (Table S2).

To further verify the 13 identified cell types, particularly bacteriocytes, we combined the snRNA-seq and ST-seq data from the InS group (mussels in the fully symbiotic state) using Seurat v3.2's anchor-based integration and projected the identified cell clusters from the snRNA-seq dataset onto the ST-seq data of the gill tissue (Figure 1D and E). While the fluorescence *in situ* hybridization (FISH) assay for eubacteria and methane oxidation symbionts confirmed that the cells of cluster 0 in the ST-seq data were bacteriocytes (Figure 1F), we further noticed that the majority of the annotated bacteriocyte marker genes from snRNA-seq data (219 out of 321) were also abundantly expressed in bacteriocytes from the ST-seq data (annotated as marker genes; Figure 1G and H; Table S2). Besides, the shared marker genes were found mainly expressed by bacteriocytes in the middle part of the gill filament, rather than by cells in the abfrontal region of the gill, where immature bacteriocytes are located (Figure 1G). Functional enrichment analysis based on Gene Ontology (GO) and double-label FISH of marker genes showed that bacteriocytes largely express genes involved in metabolic processes (e.g., biosynthesis and transport of carbohydrates, lipids, amino acids, and vitamins) and symbiosis-related immune responses (including GO:0002376, GO:0009617, GO:0044111, and GO:0044403) (Figure S2B). Additionally, we noticed that 116 out of 321 marker genes were consistently expressed by bacteriocytes in the InS and DeC groups regardless of fluctuations in

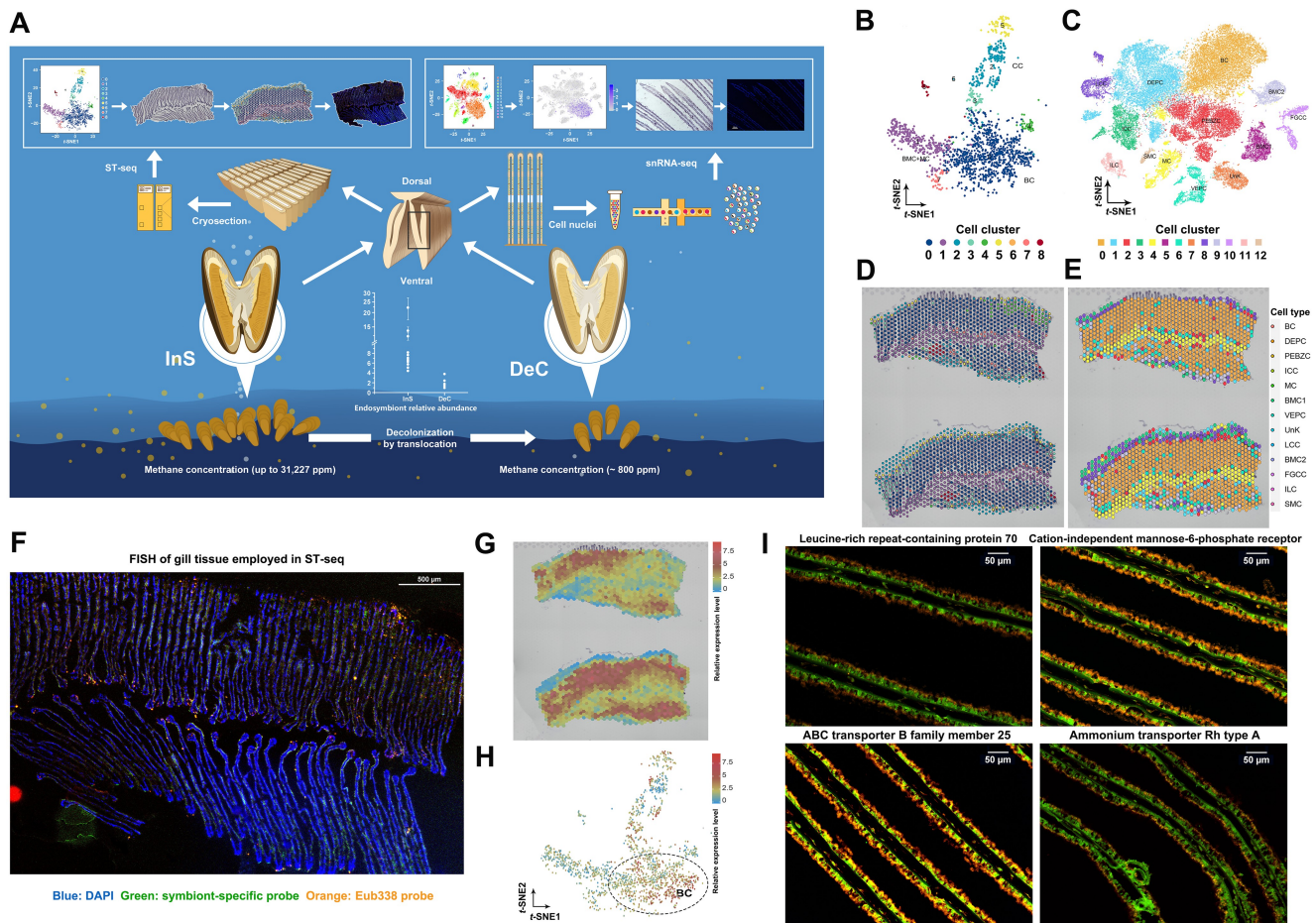


Figure 1 Spatially resolved single-cell transcriptomic atlas of deep-sea mussel gill

A. Experimental workflow and analysis of snRNA-seq and ST-seq data of gill tissue in fully symbiotic (InS group) and partially decolonized (DeC group) deep-sea mussels. A marked decrease in the endosymbiont was observed after the *in situ* translocation assay, as calculated by comparing the relative DNA content of the symbiont methane monooxygenase gene with the internal control host β -actin gene by qRT-PCR ($n = 14$ for the InS group, $n = 5$ for the DeC group). **B.** tSNE projection of the spatial transcriptome clustered by gene expression in the InS group, with color assigned by cell type. **C.** Integrated tSNE projection of single-nucleus transcriptome clustered by gene expression in the InS ($n = 3$) and DeC ($n = 1$) group, with color assigned by cell type. **D.** Projection of cell clusters onto spatial transcriptome barcoded spots in the InS group. Two successive sections were used in the same capture region of the spatial transcriptome as technique replicates. **E.** Spatial transcriptome barcoded spots labeled using snRNA-seq cell type with maximum prediction score. **F.** FISH of endosymbionts with successive gill sections for ST-seq (the InS group). **G.** Expression patterns of bacteriocyte marker genes identified by snRNA-seq are depicted in the spatial transcriptome data. The relative levels, indicated with different colors, of the 219 shared marker genes identified in both snRNA-seq and ST-seq data were projected onto the barcoded spots. **H.** The relative levels of the 219 shared marker genes identified in both snRNA-seq and ST-seq data were projected onto the tSNE plot of the spatial transcriptome. **I.** Double-label FISH of selected marker genes (fluorescein-labeled gene-specific probe, green) from snRNA-seq and ST-seq data and symbionts (DIG-labeled Eub338 eubacteria probe, orange) was conducted to verify the bacteriocyte cluster. snRNA-seq, single-nucleus RNA sequencing; ST-seq, spatial transcriptomics sequencing; FISH, fluorescent *in situ* hybridization; tSNE, t-Distributed Stochastic Neighbor Embedding; DIG, digitoxin; BC, bacteriocyte; DEPC, dorsal end proliferation cell; PEBZC, posterior end budding zone cell; VEPC, ventral end proliferation cell; ICC, intercalary cell; LCC, lateral ciliary cell; FGCC, food groove ciliary cell; SMC, smooth muscle cell; BMC1, basal membrane cell 1; BMC2, basal membrane cell 2; MC, mucus cell; ILC, inter lamina cell; UnK, unknown cell; CC, ciliary cell.

symbiont abundance (Table S3). These included the cation-independent mannose-6-phosphate receptor (Figure 1I), ABC transporter B family member 25 (Figure 1I, Figure S2C), leucine-rich repeat-containing protein 70 (Figure 1I), excitatory amino acid transporter 1, monocarboxylate transporter 13, cathepsin L (*ctsl*), and myoneurin (*mynn*) (Table S3). Of the remaining 205 marker genes, the majority (174) were downregulated in the DeC group in response to symbiont decolonization (Table S3). These included genes for the solute carrier family 40 member 1, ammonium transporter Rh type A (*rhbg-a*) (Figure 1I, Figure S2C), solute carrier family 26 member 10, 24-hydroxycholesterol 7 α -hydroxylase (*cyp39a1*, EC1.14.14.26), sugar phosphate exchanger 2 (*slc37a2*), zinc finger protein 271 (*znf271*), histone-

lysine *N*-methyltransferase *ezh2*, and ETS-related transcription factor (*elf-3*) (Table S3).

Symbiosis-related immune processes in bacteriocytes

How invertebrate hosts, armed with innate immunity alone, forged a symbiotic association with specific chemosynthetic bacteria remains intriguing. We screened for genes and pathways that may participate in the establishment and maintenance of symbiosis in bacteriocyte lineages. Although pattern recognition receptors (PRRs), such as Toll-like receptors (TLRs) and peptidoglycan recognition proteins (PGRPs), could play vital



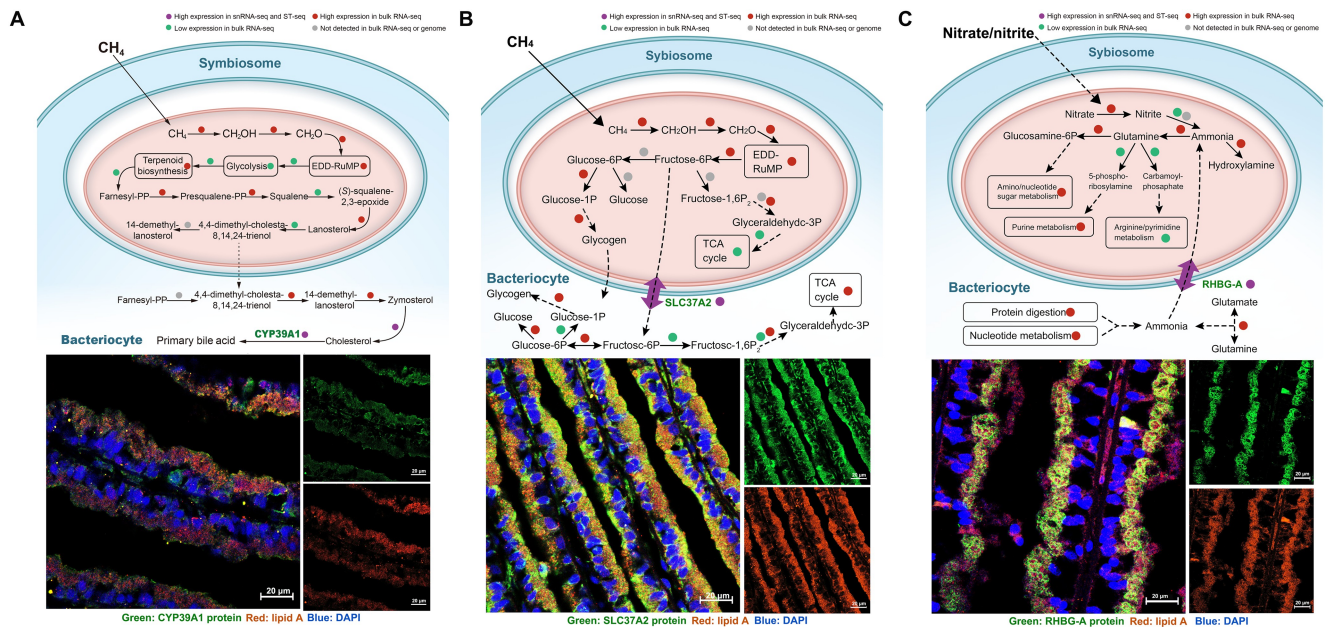


Figure 3 Metabolic interactions between mussel host and endosymbionts

A. Interaction of sterol metabolism between the host and symbionts. The snRNA- and ST-seq data showed the abundant expression of *cyp39a1* in bacteriocytes. CYP39A1 plays a crucial role in host cholesterol metabolism, which relies heavily on sterol intermediates provided by symbionts, as evidenced by genomic and meta-transcriptome data. IF assays also show that most CYP39A1 proteins (in green) are distributed in the apical region of bacteriocyte, which is enriched with endosymbionts (indicated by lipid A signals, in red). **B.** Metabolic interaction of sugar phosphate between the host and symbionts. Although genomic and meta-transcriptome data showed abundant expression of glycogen synthesis genes in both host and symbiont, genes converting fructose-6P to glucose-6P or fructose-1,6P₂ are not detected in the symbiont. Nonetheless, these genes are encoded and actively expressed in mussel gills. In addition, snRNA and ST-seq data showed an abundant expression of the *slc37a2* gene by bacteriocytes. Further IF assays showed the co-location of SLC37A2 proteins (in green) with endosymbionts (in red) inside bacteriocytes. **C.** Interactions of ammonia metabolism between the host and symbionts. Ammonia, a by-product of host metabolism, plays a crucial role in symbiont production of hydroxylamine and glutamine. The mussel host was found to be massively transcribing ammonium transporter genes (up to three different genes) in bacteriocytes, potentially allowing direct import of ammonia for the symbionts. In support of this, IF assays showed the co-location of RHBG-A proteins (in green) with endosymbionts (in red) inside bacteriocytes. IF, immunofluorescence; CYP39A1, 24-hydroxycholesterol 7 α -hydroxylase; SLC37A2, sugar phosphate exchange; RHBG-A, ammonium transporter Rh type A; EDD-RuMP, Entner-Doudoroff pathway and ribulose monophosphate cycle.

expression (Figure 3A, Figure S4A–D; Table S4), evidencing the metabolic dependency of mussel hosts on sterol-derived intermediates of symbionts [33].

Both the mussel host and endosymbionts highly expressed key gluconeogenesis and glycogen biosynthesis genes (e.g., EC2.4.1.18, 2.4.1.21, 2.7.7.27, and 5.4.2.2), as demonstrated by meta-transcriptome data, with these genes being among the top 10% of highly expressed genes in the InS group (Figure 3B, Figures S5 and S6; Table S4), implying robust production of glycogen in mussels and endosymbionts. However, we observed that multiple enzymes in the gluconeogenesis pathway of the endosymbiont were either insufficiently expressed, such as EC5.3.1.9, or completely absent from the genome, including EC3.1.3.11, EC2.7.1.11, and EC4.1.2.13, which is likely to result in overproduction of fructose-6P and shortages of glucose-6P, fructose-1,6P₂, glyceraldehyde-3P, and glucose and reduced glycogen production by the symbiont and host. Conversely, the mussel host expressed all these enzymes; moreover, a sugar phosphate exchanger gene *slc37a2* (responsible for transmembrane transport of fructose-6P, glucose-6P, fructose-1,6P₂, and glyceraldehyde-3P) was found abundantly expressed by bacteriocytes, as shown by snRNA-seq and ST-seq of the InS group, and co-localized with endosymbionts (Figure 3B, Figure S6). The complementation in carbohydrate metabolism strongly

suggests that symbionts can supply fructose-6P directly to the host in exchange for gluconeogenesis intermediates. Consistent with this observation, we also noticed that the endosymbionts were massively transcribing phosphate and carboxylate transporters (including the phosphate-binding protein PstS, maltose 6-phosphate phosphatase, phosphate permease, sodium-dependent dicarboxylate transporter SdcS, and bicarbonate transporter BicA; Table S4), which may facilitate the export of sugar phosphate compounds.

TEM images of bacteriocytes have shown that endosymbionts are contained in separate vacuoles designated as symbiosomes [34]. While the presence of symbiosomes creates a suitable micro-environment for endosymbionts, they also limit access to environmental nutrients such as ammonia. As a by-product of protein digestion and nucleotide metabolism, ammonia is harmful to mussel hosts' cells and tissues but is necessary for endosymbionts. Specifically, we observed higher expression of ammonia consumption-related genes (top 10%) rather than ammonia production-related genes in endosymbionts (Figure 3C, Figure S7; Table S4). Meanwhile, our snRNA- and ST-seq data also revealed that up to three ammonium transporter genes were highly expressed in bacteriocytes (Figures 1G and 3C), which may facilitate the transport of ammonia to symbiosomes. In addition, we noted that the RHBG-A protein, one of the most

abundantly expressed ammonium transporters in bacteriocytes, co-localized with endosymbionts (Figure 3C). At the same time, endosymbionts also abundantly transcribed ammonium transporter genes (such as the ammonia channel gene, top 10% expressed genes in the InS group; Table S4), collectively evidencing the direct flow of ammonia from host to symbiont.

We also found that the expression levels of genes involved in host-symbiont metabolic interactions (i.e., *cyp39a1*, *rhhg-a*, and *slc37a2*) were markedly downregulated in bacteriocytes of the DeC group when the methane supply was limited and when symbiont abundance decreased (Figure S8). Taken together, these findings highlight the possibility that mussel hosts interact with symbionts dynamically based on the environment and symbiont abundance.

Coordinated regulatory networks guide molecular functions in bacteriocytes

While the above results collectively suggest that the host optimizes its immune and metabolic processes to facilitate symbiosis, we questioned whether a coordinated regulatory network

guides these processes. To construct the regulatory network, we conducted weighted gene co-expression network analysis (WGCNA) using snRNA-seq data from the fully-symbiotic (InS group) and decolonized mussels (DeC group). One WGCNA module, Mod05, contained 225 out of 321 marker genes for bacteriocytes (Figure S9A; Tables S5 and S6), suggesting that it represents the network regulating bacteriocyte function. Functional enrichment analysis demonstrated that these 225 genes are involved in the immune and metabolic processes of bacteriocytes (Figure S9B; Table S7).

Next, we screened for highly expressed transcription factors and signal transducers within the regulatory network that may function as hub regulators of bacteriocyte function. Within the regulatory network of bacteriocytes, we identified nine transcription factors, including *znf271*, a zinc finger protein; *elf-3*, an ETS-related transcription factor; *ezh2*, a histone-lysine *N*-methyltransferase; *gata14*, a GATA zinc finger domain-containing protein; *met*, the hepatocyte growth factor receptor; *mynn*, a zinc finger protein myoneurin; *zmiz1*, a zinc finger MIZ domain-containing protein; *helz2*, a helicase with a zinc finger domain; and *GT-3b*, a trihelix transcription factor (Figure 4A). These

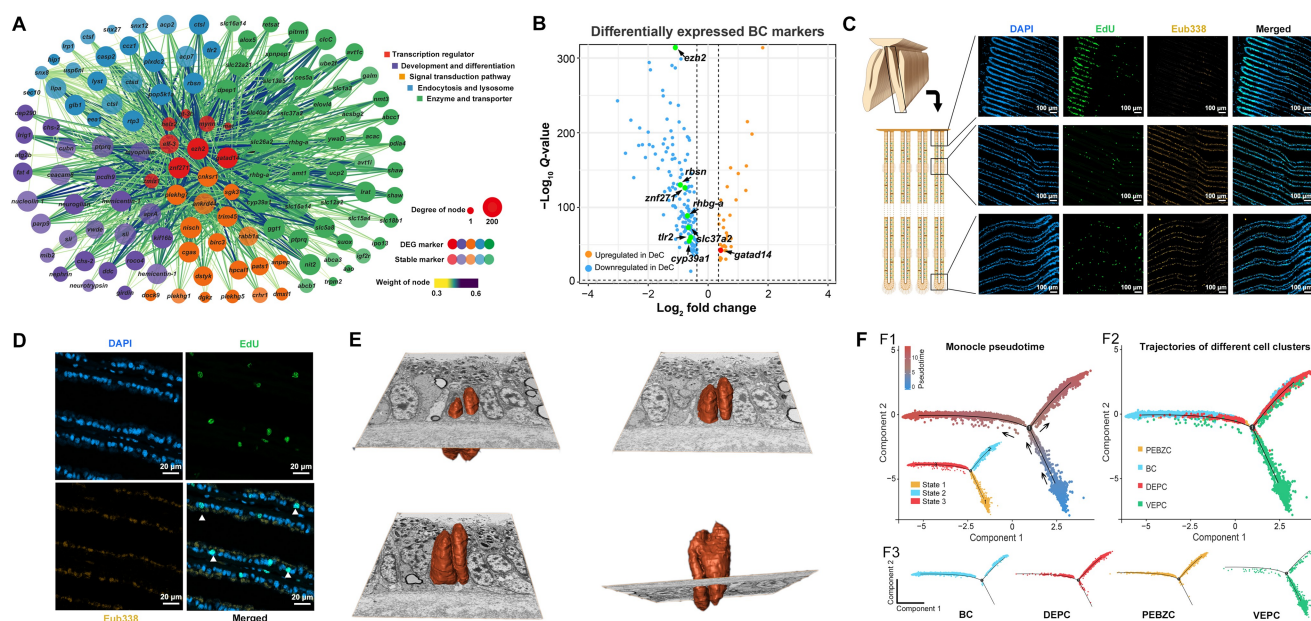


Figure 4 Development trajectory of deep-sea mussel bacteriocyte lineages

A. A genetic regulatory network of bacteriocytes was constructed using WGCNA with snRNA-seq data. Marker genes with similar functions or involved in the same biological processes are labeled with the same color. The WGCNA weight between two given nodes is indicated by different edge colors, and the degree of a given node is indicated by the dot size. B. Expression patterns of bacteriocyte markers in decolonized mussels (DeC group) in comparison with fully symbiotic mussels (InS group). Dozens of marker genes were consistently expressed by bacteriocytes, regardless of variations in symbiont abundance, recognized as stable markers and displayed in translucent colors. The remaining markers, including some hub transcription factors of the genetic regulatory network, were found to be downregulated in the DeC group (in solid colors), accompanied by differential expression of 148 other element markers (the X-axis represents the $\log_2$ -scaled ratio of fold changes between the DeC and InS groups, with dashed vertical lines at ± 0.36 denoting the fold change threshold, and the Y-axis represents the $-\log_{10} Q$ -value of the significant difference test, with a dashed horizontal line at 2 indicating the significance threshold). C. Distribution patterns of proliferation cells in the InS group as revealed by an EdU staining assay. Newly synthesized DNA was labeled with EdU signals (green) while cell nuclei were stained by DAPI (blue), and symbionts were labeled using a Cy3-labeled Eub338 probe (orange). D. Newly proliferated bacteriocytes observed by EdU assay and 3D electron microscopy. EdU signals indicating DNA replication are observed in the nucleus region of some bacteriocytes of the InS group (white triangle). E. Nuclear division in bacteriocytes could also be observed with 3D reconstruction of gill tissues (a total of 91 serial images with a thickness of 100 nm per section). F. Development trajectory of bacteriocyte lineages. Pseudo-time analysis in Monocle shows the successive development process from proliferation cells to bacteriocytes, which could be further divided into three distinct states (F1). Proliferation cells and ciliated cells are distributed mostly in states 1 and 2, while bacteriocytes are mostly distributed in states 2 and 3 (F2, 3). WGCNA, weighted gene co-expression network analysis; EdU, 5-ethynyl 2'-deoxyuridine; *znf271*, zinc finger protein 271; *elf-3*, ETS-related transcription factor 3; *ezh2*, histone-lysine *N*-methyltransferase 2; *gata14*, GATA zinc finger domain-containing protein 14; *met*, hepatocyte growth factor receptor; *mynn*, zinc finger protein myoneurin; *zmiz1*, zinc finger MIZ domain-containing protein 1; *helz2*, 3'-5' exoribonuclease HELZ2; *gt-3b*, trihelix transcription factor GT-3b. Abbreviations for the remaining genes in Figure 4A are listed in Table S6.

transcription factors were highly expressed in bacteriocytes and correlated with the expression of 225 out of 321 bacteriocyte marker genes, including the aforementioned metabolic genes (*cyp39a1*, *slc37a2*, and *rhhg-a*) and immune-related genes (*tlr2*, *rbsn*, *lyst*, and cathepsins) (Figure 4A). The expression of *ezh2*, *znf271*, and *gata14* was also differentially regulated during decolonization, in conjunction with the differential expression of 148 other marker genes (Figure 4B). Moreover, strong interconnectivity (WGCNA weight > 0.3) was observed among these transcription factor genes (Figure 4A), indicating that they may regulate each other synergistically.

Successive trajectory of the development of bacteriocyte lineages

Without culturable samples of post-larval and juvenile deep-sea mussels, the mechanisms underlying the development of bacteriocyte lineages remain elusive. As reported in our previous study [28], there are three types of proliferation cells in the gills of mussel hosts: PEBZCs, DEPCs, and VEPCs. By conducting an EdU-labeling assay, we confirmed widespread proliferative activity along the gill filaments, especially in the dorsal part of the gills, where DEPCs are located (Figure 4C, Figure S8D). Additionally, we observed proliferating signals in some bacteriocytes of adult mussels, as assessed using EdU-labeling and 3D electron microscopy (Figure 4D and E; Video S3). While robust proliferation signals confirmed the continuous birth of bacteriocytes in adult mussel gills, the identification of marker genes for bacteriocytes and proliferation cells provides a unique opportunity to advance our knowledge of bacteriocyte development. Using an expression atlas for proliferation cells and bacteriocyte lineages, we also characterized the bacteriocyte development trajectory with five different models: pseudo-time analysis by Monocle2 (Figure 4F), partition-based graph abstraction (PAGA) analysis, Slingshot, Cytotrace, and RNA velocity analysis (Figure S9C and D). The Monocle analysis showed that gill cells were distributed across three distinct pseudo-time states. Most proliferation cells, as the starting point of differentiation, were in states 1 and 2 (for example, 93.41% of VEPCs were in state 1, while 84.02% of DEPCs and 50.77% of PEBZCs were in state 2, respectively). Conversely, as differentiation endpoints, approximately 94.96% of bacteriocytes were in state 3 (Figure 4F). Similar results were also observed in the other analyses, where bacteriocytes were found in a more differentiated state than DEPCs, PEBZCs, and VEPCs (Figure S9C). These distinct pseudo-time trajectory states collectively confirm developmental heterogeneity within gill cells and position bacteriocytes as one of the most differentiated cells in mussel gills. Particularly, a succession of processes occurred during differentiation (cell type transitions from proliferation cells, especially from DEPCs, to bacteriocytes) and maturation (subtype transitions from state 2 to state 3) of bacteriocytes (Figure 4F, Figure S9C and D).

To further identify the cell differentiation processes in bacteriocyte lineages, we examined the genes abundantly expressed at different pseudo-time states (Table S8). We speculated that progenitor states would encode key marker genes (especially transcription factors) of descendant cells, and that the expression levels of these marker genes would increase as cells

mature, showing maximal expression levels in the fully functional state. Our results indicate that bacteriocytes at state 3 are in a fully functional state, transcribing the majority of marker genes (including hub marker genes) (Table S8). Bacteriocytes in state 2 were functionally immature and showed moderate expression of marker genes. Notably, a large number of genes abundantly expressed at state 3 of proliferation cells (128 out of 625 genes in DEPCs, 121 out of 621 genes in PEBZCs, and 70 out of 101 genes in VEPCs) were bacteriocyte marker genes (Figure S9E). Meanwhile, 33 out of 81 DEPC marker genes were also highly expressed in state 2 PEBZCs (Table S8). These findings suggest that the state-3 proliferation cells may be bacteriocyte progenitors, while DEPCs may be descendants of PEBZCs. This conclusion is also supported by RNA velocity analysis and the Slingshot model, in which a positive velocity from PEBZCs and DEPCs to bacteriocytes was observed in the corresponding cells of adjacent regions (Figure S9C).

Co-option of conserved transcription factors in bacteriocyte development

While the development of symbiotic cells and organs may be a highly organized process, the regulatory networks that guide it are difficult to identify in both model and non-model symbiotic associations. Using our bacteriocyte developmental trajectory, we explored the mechanisms underlying bacteriocyte differentiation and maturation, focusing on genes that are abundantly expressed in the progenitor state. By conducting *t*-Distributed Stochastic Neighbor Embedding (*t*SNE) analysis, we observed an overlap in the *t*SNE distributions between state 2 bacteriocytes and state 3 proliferation cells, and an overlap between state 2 DEPCs and state 2 PEBZCs (Figure 5A), confirming that state 3 DEPCs/PEBZCs are progenitors of bacteriocytes. Specifically, our results indicate that the mRNA transcripts of histone-lysine *N*-methyltransferase *ezh2*, a hub marker gene in the regulatory network of bacteriocytes and a canonical regulator of gene expression and cell differentiation, are abundantly expressed in state 3 of all proliferation cell types (Figure 5B; Table S8). Additionally, the expression of plexin domain-containing protein *plxdc2*, *acp7*, cyclic GMP-AMP synthase (*cgas*), and the serine/threonine-protein kinase *roco4*, which are bacteriocyte marker genes identified in the regulatory network and crucial regulators of cell differentiation and lysosomal function, was also highly upregulated in state 3 of all proliferation cells (Figure S9F; Table S8). In addition, the expression of five other hub transcription factors in the regulatory network of bacteriocytes robustly increased in state 3 DEPCs (*elf-3*, *mynn*, *helz2*, *zmiz1*, and *met*) and two in PEBZCs (*elf-3* and *met*) in comparison with the other cell states (Table S8), suggesting their participation in the transition of proliferation cells into bacteriocytes.

Similarly, a gradual increase in the expression levels of the four aforementioned hub genes (*znf271*, *ezh2*, *zmiz1*, and *helz2*) was also observed during further maturation (from state 2 to state 3) of bacteriocytes (Figure 5C; Table S8). Additionally, 13 transcription factors, which act as activators or repressors, were highly expressed in state 2 bacteriocytes (Figure S9F; Table S8). To address the potential influences of symbionts on the function and development of bacteriocytes, we further searched for

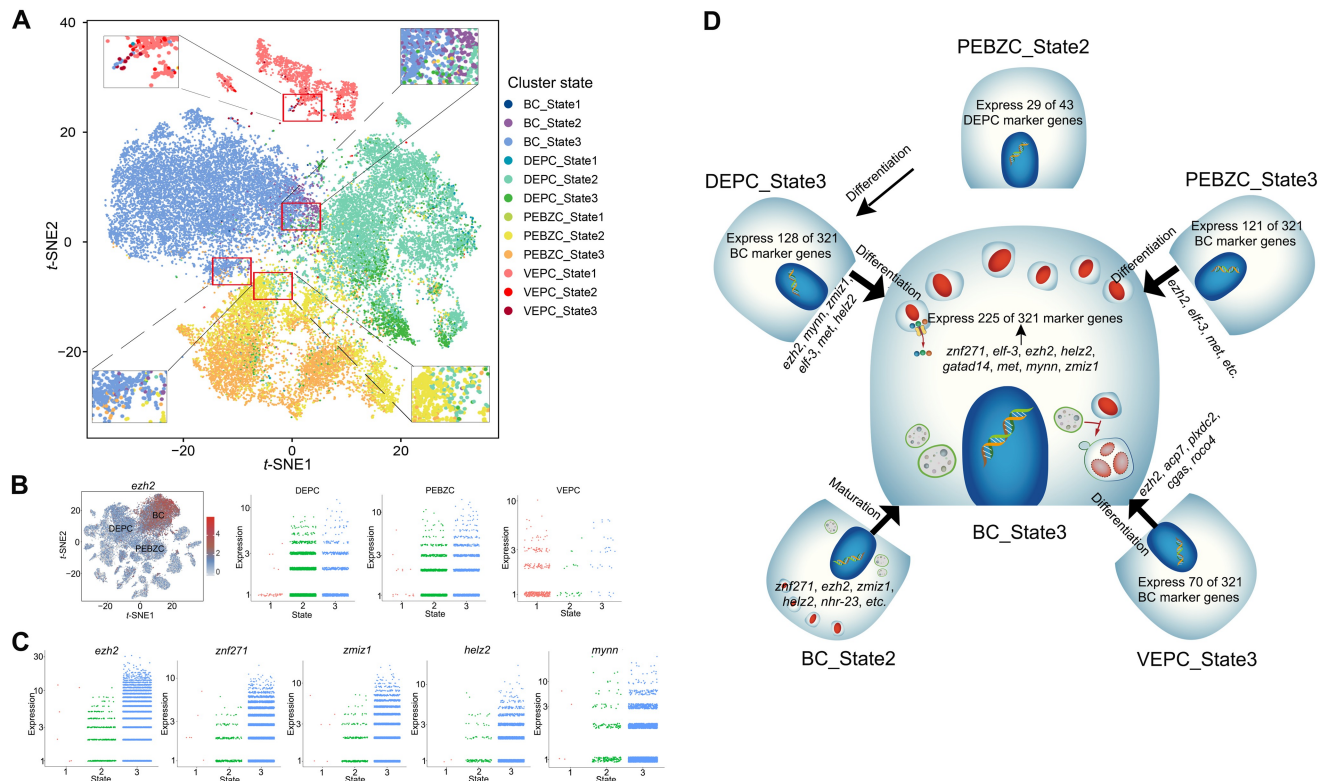


Figure 5 Co-option of conserved transcription factors in the function and development of bacteriocytes

A. Proliferation cells and bacteriocytes in different pseudo-time states were projected on the t-SNE plot of snRNA-seq data. Co-overlaps in the t-SNE distribution of bacteriocytes and state 3 cells of proliferation cells (such as DEPC and PEBZC) and between DEPC and state 2 cells of PEBZC were observed (shown in magnified fields). **B.** Expression patterns of the hub transcription factor *ezh2* in color-coded t-SNE plots (left, color indicating the relative expression level) and different states of proliferation cells (DEPC, PEBZC, and VEPC). **C.** Expression patterns of hub transcription factors in different states of bacteriocytes. **D.** Schematic diagram shows the function and development of bacteriocyte lineages guided by the co-option of conserved transcription factors identified from the genetic regulatory networks. Arrows represent the deduced developmental trajectory (including cell type transition and maturation) in bacteriocytes.

sterol- and sugar phosphate-related genes that were differentially expressed in state 2 of bacteriocytes and after the decolonization assay, as sterols and sugar phosphates are two main nutrients derived from symbionts. Particularly, we observed that the nuclear hormone receptor family member *nhr-23* was abundantly expressed in state 2 (Figure S9F; Table S8). In addition, seven genes involved in the sterol-related signaling pathway (including *cyp39a1* and six nuclear receptor genes) were also differentially expressed in bacteriocytes of decolonized mussels (Table S3). Nuclear receptor genes, including *nhr-23*, are known to play a crucial role in the sterol-related signaling pathway by binding to steroidal ligands and regulating cell function and development.

While the aforementioned hub transcription factors and signaling transducers may be trigger molecules promoting bacteriocyte development, we asked whether they were evolutionarily novel genes that diverged from non-symbiotic ancestors or conserved across evolution. Our phylogenetic analysis of protein sequences obtained from homologues of these genes in Lophotrochozoa indicates that these genes are highly conserved in mollusks and are in accordance with the general host phylogeny (Figure S10). We propose that bacteriocyte development occurred through the co-option of conserved genes rather than the emergence of evolutionarily novel genes (Figure 5D).

Discussion

In addition to representing an evolutionary novelty, symbiotic cells and organs provide a suitable niche for symbiosis and endow holobionts with unique adaptations to their surrounding environments [6]. In this study, we integrated data obtained from single-nucleus omics, spatial omics, phagocytosis assays, EdU labeling assays, and 3D electron microscopy to uncover the regulatory network that controls the function and development of bacteriocytes in deep-sea mussels. Our findings suggest that these processes are guided by an ancestral intrinsic toolkit through the co-option of conserved transcription factors and modulation by endosymbionts. Our results suggest how the two partners interact closely and cooperatively shape the function and development of symbiotic cells, which is crucial for understanding the evolution of chemosymbiosis and the adaptations of holobionts in habitats such as the deep sea [35,36].

The sophisticated immune network controls the “ins” and “outs” of endosymbionts

The establishment and maintenance of endosymbiosis with exogenous bacteria are challenging for multicellular organisms

because of the host immune system, which requires adjustments that entail engulfing and proliferating symbionts within the host cell [37]. Although it is well recognized that some PRRs have greatly expanded in the deep-sea mussel genome and could act as immune receptors for symbiont [19], only a few canonical PRRs were found to be abundantly expressed in bacteriocytes, either in our snRNA-seq data or ST-seq data. Instead, several *lrr* genes were found to be abundantly expressed in bacteriocytes, which have been suggested as potential intracellular receptors for methanotrophic endosymbionts in previous studies [31]. While the aforementioned canonical PRRs might participate in other processes, such as early symbiotic acquisition in juveniles and defense against pathogenic bacteria, our findings suggest that only a few PRRs, including some non-canonical ones, are needed to maintain symbiosis in normal bacteriocytes of adult mussels. Nevertheless, bacteriocytes can still phagocytize exogenous bacteria in a nonselective manner, even when fully colonized by endosymbionts, as observed in our phagocytosis assay and in some recent studies [38]. The phagocytic ability of gill cells may be conserved in most mollusks, including non-symbiotic ones [38–40], which directly facilitates the entry of endosymbionts into deep-sea mussels and may contribute significantly to the establishment of symbiosis in other mollusks. To facilitate phagocytosis and subsequent lysosome-mediated digestion of bacteria, a mass of phagosome- and lysosome-related genes was also found to be abundantly expressed by bacteriocytes. For example, *rbsn* and *eez1*, as two marker genes of bacteriocytes, could be recruited to early endosomes and therefore play a crucial role in the formation of the phagocytic cup; the *rab* family and *snx27* are essential for intracellular trafficking, especially between phagosomes and intracellular compartments [41]. Similarly, *lst* and *sh3gl3* could play a necessary role in lysosome maturation, while *ctsd/f/l* and *acp2/7* are required for lysosome-mediated biodegradation. The massive expression of the aforementioned genes collectively underpins the unique function and cell structure of bacteriocytes, as observed by phagocytosis assay and TEM analysis. However, the phagocytosis of exogenous bacteria also raises questions about whether and how mussels discriminate between symbionts and non-symbionts, and between symbionts in good and bad states, within phagosomes or symbiosomes, since it is often observed that bacteriocytes digest some bacteria resembling symbionts via lysosomes. Recently, it has been reported that the digestion of symbionts is regulated by mTORC1, particularly under conditions of reduced nutrient supply (either due to endosymbiont death or methane removal) [42–44]. At the same time, we rarely observed proliferating endosymbionts or free symbionts engulfed in bacteriocytes of adult mussels, as evidenced by EdU labeling assay and 3D electron microscopy. We therefore speculated that lysosome-mediated symbiont digestion is a secondary option to obtain nutrition under normal circumstances, but an efficient way to control the symbiont population and obtain nutrition under abrupt stresses or other emergencies. In support of this speculation, genes involved in phagosome and lysosome maturation were found massively downregulated in bacteriocytes of decolonized mussels. Our findings also support that symbiotic associations in deep-sea mollusks are nutrition-driven and may

involve metabolic interactions between mussel hosts and symbionts [11].

The close interactions between the metabolism of mussels and endosymbionts

Since the first characterization of chemosymbiosis, it has long been debated how invertebrate hosts, including deep-sea mussels, acquire nutrients from and supply essential substances to the symbionts. Besides the lysosome-mediated digestion of symbionts, mussel hosts are suggested to acquire nutrients through the uptake of secreted symbiont metabolites [38]. However, identifying the exact metabolites and the mechanisms of their transport between symbionts and bacteriocytes has remained unclear and difficult to characterize [19,20,33,45,46]. Genomic information shows that methanotrophic endosymbionts can provide sterol intermediates to the mussel hosts. Geier et al. recently identified several specialized metabolites at the host-microbe interface using metaFISH and AP-MALDI-MSI analyses [46]. With the help of state-of-the-art single-nucleus and spatial sequencing, we here provide evidence that deep-sea mussels have significantly reshaped bacteriocyte metabolism to maximize symbiotic profits for both partners. For example, bacteriocytes encode dozens of genes involved in the biosynthesis and transport of carbohydrates, lipids, amino acids, and vitamins, which improve acquisition of nutrients from symbionts. Our findings also reveal that bacteriocytes express the sugar-phosphate exchanger gene *scl37a2*, which co-localizes with symbionts and may facilitate the retrieval of fructose-6P directly from methanotrophic symbionts. As a member of the SLC37 family, SLC37A2 is typically associated with the endoplasmic reticulum (ER) and functions as a Pi-linked antiporter for fructose-6P or glucose-6P [47]. While it is widely acknowledged that the ER plays a role in the formation and development of symbiosomes [48], our results provide a vivid example of how ER-associated proteins like SLC37A2 specifically contribute to symbiosome function. In bacteriocytes, fructose-6P is further converted into glucose-6P, fructose-1,6P₂, and glyceraldehyde-3P, which could support the gluconeogenesis and the tricarboxylic acid (TCA) cycles in both mussel hosts and symbionts. In addition, we suggest that mussel hosts supply ammonia to symbionts directly via ammonium transporters *rhhg-a*. Although ammonium is a preferred nitrogen source for endosymbionts, it is toxic to animal hosts and must be excreted. While ammonium transporters typically belong to three superfamilies (Amt/Mep/Rh), the AMT superfamily is most commonly associated with ecto- and endomycorrhizal symbiosis in plants [49]. In contrast, this study provides the first evidence of how RHBG-A, a member of the Rh superfamily, facilitates ammonia transport in symbiotic relationships. Similarly, it has been shown that the thiotrophic symbionts of *Riftia* can release succinate directly to the host after CO₂ fixation; the sea anemone *Aiptasia* can also provide ammonia directly to symbiotic algae and control their growth [50,51]. Moreover, a recent study on vestimentiferan tubeworms showed that thiotrophic symbionts massively transcribe ammonia transporter genes to absorb and detoxify ammonia from the host [29]. The direct transport of fructose-6P

and ammonia in mussel holobionts, therefore, could provide a shortcut for both mussel hosts and methanotrophic symbionts to obtain essential nutrients, which may greatly increase the efficiency and profits of symbiosis to adapt to the harsh environment of the deep sea and serve as a supplementary way to control the symbiont population. In support of this speculation, it is well recognized that a single bacteriocyte can harbor up to thousands of methanotrophic symbionts, while the free-living ones are rarely detected in either bottom seawater or deep-sea carbonates [52].

The co-option of conserved genes in guiding the development of bacteriocytes

As a specialized and stable niche for endosymbiosis, bacteriocytes are intriguing and crucial in understanding the evolution of symbiosis [2,6]. While previous studies demonstrated that the proliferation cells in the growth zones of gills can differentiate into bacteriocytes, our single-nucleus and spatial transcriptome data, for the first time as far as we know, traced the successive development trajectory of bacteriocytes. We provide evidence that proliferation cells and bacteriocytes can be divided into different states and that only a subset of proliferation cells are the progenitors of bacteriocytes. The differentiation and maturation of bacteriocytes may be guided by a set of mollusk-conserved transcription factors (including *ezh2*, *elf-3*, *znf271*, and *met*). Besides, the regulatory networks responsible for the development of bacteriocytes are likely to also control the immune and metabolic processes, two of the most noticeable features of bacteriocyte function. All these transcription factors are evolutionarily conserved across mollusks and are in agreement with the general host phylogeny. Additionally, although less investigated in mollusks, homologues of the majority of these genes, such as *ezh2*, *elf-3*, *met*, and *mynn*, are all known as crucial regulators of cell proliferation and differentiation [53–56]. Among these genes, *ezh2* is of particular interest as it is ubiquitously expressed by all gill cell clusters (especially for bacteriocytes) and is drastically upregulated in state 3 of all proliferation cells. The expression pattern of *ezh2* suggests a leading role of histone methylation in controlling the function and development of bacteriocytes, which also supports the conclusion that the formation of symbiotic cells in deep-sea mussels is due to the co-option of conserved rather than the emergence of evolutionarily novel genes [57]. Similar phenomena have also been observed in the deep-sea scaly-foot snail *Chrysomallon squamiferum*, where the formation of a biomineralized skeleton is driven by an ancestral intrinsic toolkit conserved across mollusks [58].

The possible role of symbionts in the function and development of bacteriocytes

Notably, we also found that the function and development process of bacteriocytes could respond to the symbiont losses induced by methane reduction. The global decline of bacteriocyte marker genes and increase of cells in an immature state in decolonized mussels demonstrate the phenotypic plasticity of symbiotic cells and suggest the participation of symbionts in

regulating the function and development of molluscan bacteriocytes. Symbiont participation in the development of symbiotic cells and organs has attracted increasing attention in recent years but has remained largely unclear at the mechanistic level [59,60]. Here, we found that the *cyp39a1* gene and a crucial nuclear receptor gene (*nhr-23*) were differentially expressed during the maturation of bacteriocytes. In addition, up to seven genes involved in the sterol-related signaling pathway (including *cyp39a1* and six nuclear receptor genes) were also differentially expressed in the bacteriocytes of decolonized mussels, highlighting the possible participation of symbiont sterol metabolism in regulating the function and development of mussel bacteriocytes. In support of this hypothesis, we previously observed the suppression of sterol/steroid biosynthesis in symbiont-depleted deep-sea mussels after long-term atmospheric cultivation without methane supply [44]. Furthermore, other studies have also shown that symbiotic *Wolbachia* modulate host development and reproduction via the steroid-nuclear receptor signaling pathway [61]. Sterol metabolism in symbionts may therefore be a common mechanism in the development of symbiotic cells and organs since several animal hosts rely on their symbionts for sterol intermediates [62–64]. Besides sterol metabolism, it is noteworthy that glucose and ammonia metabolism could also serve as intracellular signals of symbiosis-related processes (e.g., via mTORC1-mediated symbiont digestion) [65–68]. Although the regulatory role of symbionts in the function and development of symbiotic cells remains to be fully elucidated, the symbiont-mediated phenotypic plasticity in molluscan bacteriocytes could benefit the host to respond rapidly and globally to the fluctuating environment of the deep sea, especially when the methane supply changes. In addition, the coordinated regulation of endocytosis and digestion, as well as the metabolic interactions with symbionts, may also help the mollusk host to reshape the symbiotic associations according to the local environment, which therefore further assists in their presence across a wide range of deep-sea habitats [69–71]. The dynamic plasticity in the function and development of molluscan bacteriocytes, especially the global repression of marker genes and increase in the number of cells in an immature state in decolonized mussels, also supports the idea that maturation is not a single-ended process but a dynamic continuum of adaptive states set by genes and the environment [72].

There are also some limitations to the present study. For example, although spatial transcriptome analysis was conducted, the relationship between gene expression and cell location could not be addressed due to the limited resolution of ST-seq data (55 μm in diameter per spot with a 100 μm center-to-center distance between spots). The spot size in ST-seq may also aggregate signals from multiple cells, thereby compromising single-cell resolution, despite its integration with snRNA-seq data. In addition, some of the hypotheses raised need to be verified in future studies using mussel samples in various symbiotic states, including aposymbiotic, recolonized, post-larval, and juvenile mussels, and employing new experimental paradigms (such as high-resolution ST-seq and scRNA-seq). Nevertheless, by combining single-nucleus and spatial transcriptome data, a phagocytosis assay, an EdU assay, and 3D electron microscopy data, we successfully established a pipeline to reveal the molecular functions and development of bacteriocyte lineages in non-

model deep-sea mussels. Our results indicate that the function and development of bacteriocytes are highly coordinated through the co-option of conserved genes and dynamically modulated by symbionts. In addition, it was surprising to find that both endosymbionts and bacteriocytes persisted even after the two-year transplantation to a low methane site, where the methane concentration is 39-fold lower than in the seepage region, a level similar to that in regions with massive piles of dead mussels. Meanwhile, the immune and metabolic processes within mussel bacteriocytes could be dynamically modulated to sustain long-lasting symbiotic associations, thereby enhancing the resilience of deep-sea mussels to methane reduction. Collectively, the intimate symbiotic associations and robust plasticity of symbiotic cells have likely made deep-sea mussels one of the most successful organisms in the deep sea.

Materials and methods

Experimental design

The goal of the present study was to address the molecular function and developmental trajectory of bacteriocytes in the deep-sea mussel *G. platifrons*. To achieve this goal, we employed methanotrophic *G. platifrons* collected from cold seeps as our model, and used spatial transcriptomics, single-nucleus transcriptomics, and meta-transcriptomics to characterize both the bacteriocytes and endosymbionts. We also employed an *in situ* transplantation assay to generate decolonized mussels and to elucidate the dynamic responses of symbiotic cells in response to symbiont depletion. We used EdU-labeling, phagocytosis assays, FISH assays, IF assays, TEM, 3D electron microscopy, and phylogenetic analysis to verify hypotheses obtained from sequencing data.

In situ transplantation and animal collection

G. platifrons mussels were collected from the cold seeps (22°06'N, 119°17'E) of the South China Sea during the cruises of 2018 (cruise KEXUE-1807) and 2020 (cruise KEXUE-2002). Permits for the cruises and sample collections were granted by the Institute of Oceanology, Chinese Academy of Sciences. Mussel samples were collected on dives 181, 206, 231, and 241 by remotely operated vehicle (ROV) FAXIAN. The mussel fauna lives at 1120 m beneath the surface at a temperature of approximately 3.35°C, salinity of approximately 35.54 psu (practical salinity unit), dissolved oxygen of 2.98–3.17 mg/l, and methane concentrations of up to 31,227 ppm in the seepage region [73]. To protect the specimens from temperature and pressure fluctuations during sampling, all mussel samples were collected using our self-designed isothermal isobaric sampler and a self-designed manually controlled macrofauna *in situ* sampling device as described previously [43]. The self-designed isothermal isobaric sampler consists of two parts. The inner part is a titanium lumen that can protect the samples from depressurization (filled with seawater and therefore can keep the mussel alive during sampling). The outer part of our sampler is an acrylic cabin with replaceable ice bags that can prevent heat exchange with the

surrounding seawater during sample recovery. After retrieving and depressurizing the isothermal isobaric sampler, the mussels were immediately dissected into two halves to remove excess seawater and quickly frozen in an isopentane bath (Catalog No. M813375, Macklin, Shanghai, China) with liquid nitrogen for snRNA-seq and ST-seq. The macrofauna *in situ* sampling device was installed on the ROV platform and manually controlled by the ROV pilot. Mussel samples can be placed into the *in situ* sampling device for immediate treatment with RNAsafer stabilizer reagent or other reagents to minimize expression alterations of samples during recovery. Briefly, a total of 20 mussels living in close proximity to active fluid seeps were collected as representatives of normal mussels (designated as the InS group). Among them, seven mussels were collected using an isothermal isobaric sampler in 2018, seven were collected using a multipurpose *in situ* sampling device in 2020 after being treated with RNAsafer stabilizer reagent (Catalog No. R0424, Omega Bio-Tek, Norcross, GA) *in situ*, and the remaining six were collected using a multipurpose *in situ* sampling device after being treated with 4% paraformaldehyde *in situ*. Three mussels from the InS group were used for snRNA-seq (Chromium platform of 10x Genomics, Pleasanton, CA). One of these three mussels was also subjected to spatial transcriptome sequencing (ST-seq on the Visium platform of 10x Genomics). All 14 mussels collected by the isothermal isobaric sampler or after RNA-stabilizing treatment were also subjected to meta-transcriptomic sequencing. Mussels collected after paraformaldehyde treatment were subjected to ISH, IF, and TEM imaging. During the 2018 cruise, we also conducted an *in situ* transplantation assay, translocating dozens of mussels from the seepage region to an authigenic carbonate region with a low concentration of CH₄ (at approximately 100 m from the seepage, the methane concentration decreased to 800 ppm). Among these transplanted mussels, five were retrieved 604 days later during the 2020 cruise and designated as the DeC group. For these retrieved mussels, one was collected using an isothermal isobaric sampler and then employed for snRNA-seq, ST-seq, and meta-transcriptome. Four were collected with the aforementioned multipurpose *in situ* sampling device (treated with RNAsafer stabilizer reagent) and subjected to meta-transcriptome.

snRNA-seq

Four individual mussels of similar size (approximately 80 mm in length) from the InS and DeC groups were used for snRNA-seq. Considering the difficulties involved in the isolation and *in vitro* culture of gill cells, snRNA-seq analysis of *G. platifrons* gill tissue was conducted using nuclei isolated from fresh frozen samples collected in an isothermal isobaric manner to minimize the potential effects on gene expression caused by mussel sampling and single cell preparation, and to maintain consistency with our previous study. The isolated nuclei were assayed following the 10x Genomics single-cell protocol by generating single-nucleus gel bead-in-emulsion (GEM) on a GemCode single-cell instrument (<https://www.10xgenomics.com/support>). Full-length cDNA synthesis was performed using Chromium Next GEM single cell 3' reagent kit (v3.1) and then sequenced on an Illumina HiSeq X Ten platform (Gene Denovo Biotechnology, Guangzhou, China).

Cryosectioning and ST-seq

To help determine the taxonomy of all gill cells, including bacteriocytes, we also conducted ST-seq with a cross-sectioned middle region of the gill (ventral view) from the InS group. The gill tissue for 10x Visium ST-seq was first incubated with pre-cooled methanol (4°C) for 15 min to reduce potential damage from the cryostat blade during cryosectioning. The tissue was then embedded in O.C.T. compound (Catalog No. 4583, Sakura, Torrance, CA) and cross-sectioned at 10 µm through the middle part of the gill using a cryostat (Leica CM1950, Heidelberg, Germany). Gill sections were then stained with 4',6-diamidino-2-phenylindole (DAPI, Catalog No. P36935, Thermo Fisher Scientific, Waltham, MA) and hematoxylin and eosin staining (HE, Catalog No. E607317, Sangon Biotech, Shanghai, China), and the morphology of the gill-tissue sections was assessed under light microscopy (Nikon ECLIPSE Ni, Tokyo, Japan). Successive sections obtained from the embedded tissue blocks were placed on Visium spatial tissue optimization slides within the capture area for the tissue optimization assay. For the sequencing assay, the gill sections were first mounted on Visium spatial gene expression slides and permeabilized according to previously described parameters (permeabilization time of 9 min) after HE staining and imaging using a light microscope. After reverse transcription and library preparation, all samples were sequenced on an Illumina HiSeq X Ten platform (Gene Denovo Biotechnology). The remaining sections of the same tissue were further subjected to ISH assays to aid cell type annotation.

Cell clustering, cell type annotation, and analysis of differentially expressed genes

Cell clustering of *G. platifrons* gill tissue was first performed using snRNA-seq-derived data. After quality control, all raw reads were mapped onto a *G. platifrons* reference genome. The genome was first reported by Sun et al. [19] and updated recently with Hi-C and high-depth PacBio long-read sequencing by us (GenBank accession NO. JAOEFJ000000000). Unique molecular identifiers (UMIs) in each sequenced read were counted and corrected for sequencing errors. Using valid barcodes that were identified based on the EmptyDrops method [74], the gene matrices of all cells were produced and imported into Seurat (v3.1.1) for cell clustering using a graph-based clustering approach [75].

Single-cell clustering of *G. platifrons* gill tissue was also performed using ST-seq-derived data, with the advantage of cell type annotation. A slide image obtained before permeabilization was first imported into Space Ranger software (<https://www.10xgenomics.com/support/software/space-ranger/>) for fiducial and Visium barcoded spot alignment. After decoding correlations between tissue, capture spots, and barcodes, a splice-aware alignment of sequencing reads to the *G. platifrons* genome was performed using STAR in the Space Ranger software package.

An integrated analysis of data derived from snRNA-seq, ST-seq, and *in situ* hybridization assay was then used for cell type annotation in the InS group. Briefly, anchor-based integration was first used to integrate ST-seq data with snRNA-

seq data using the FindIntegrationAnchors command in Seurat (v3.2). All cell-type labels in snRNA-seq were then transferred to spatial data using the TransferData command. Cell type prediction scores indicating the similarity between ST-seq spots and snRNA-seq cell clusters were calculated simultaneously, and only spot-cluster pairs with the highest scores were considered for further cell-type annotation. For annotation and cell-type verification, cell types in gill tissue were inspected using HE-stained images of gill cryosections and tissue plots colored by the clustering of ST-seq data. Several cell markers obtained using the FindAllMarkers function in Seurat were also cloned and subjected to *in situ* hybridization using gill tissues from the InS group.

To further investigate the biological function of all identified cell clusters, differentially expressed genes (DEGs) were surveyed in both snRNA-seq and ST-seq data. For snRNA-seq, expression values of all identified genes in a given cluster were compared against those of the other cell clusters using the Wilcoxon rank sum test as suggested recently [76]. Genes expressed mainly in the target cluster (more than 25% of cells were designated as a target cluster), and showing at least 1.28-fold upregulated expression levels and $P < 0.01$, were considered DEGs per cell cluster. According to previous studies [77–79], model-based analysis of single-cell transcriptomics (MAST) in the R package was used to determine DEGs in a single cell cluster using the same criteria as those used for snRNA-seq. All genes in ST-seq data were further analyzed for spatially-specific DEGs using markvariogram in the Seurat R package and mark-segregation hypothesis testing in the trendsceek R package. Genes with “r.metric.5” parameter values of < 0.8 and $P < 0.01$ were designated as spatially DEGs.

Meta-transcriptome sequencing

To explore the expression atlas of symbionts in the InS and DeC groups, we performed meta-transcriptome sequencing using 14 samples from the InS group and five samples from the DeC group as described previously. Total RNA extraction, removal of eukaryotic and prokaryotic rRNA, and synthesis of cDNA libraries using bulk-seq were conducted as described previously [43]. cDNA libraries from the InS and DeC groups were finally sequenced on an Illumina HiSeq 2500 platform with paired-end reads performed by Novogene (Tianjin, China). After quality control, the filtered reads were aligned against both the endosymbiont and *G. platifrons* genomes using HISAT (v2.0.4) [80]. The genome of methanotrophic symbionts was first reported by Takishita et al. [33], and we updated it with high-depth PacBio long-read sequencing [raw data deposited in National Center for Biotechnology Information (NCBI) with accession No. PRJNA 893022] [81]. After mapping the raw reads to the combined pan-genome of the host and symbiont, the reads originating from either the host or the symbiont were extracted and treated separately for expression analysis. The expression levels of host and symbiont genes were calculated using HTSeq (v0.6.1) [82], while significant differences between groups were determined using DESeq2 (v1.10.1) [83]. Additionally, the top 10% of genes with the most abundant mRNA transcripts in the InS group were designated as abundantly expressed genes to survey the biological

processes that actively occurred under normal conditions *in situ*. For the PCoA analysis, the Bray-Curtis distance between samples was first calculated and subjected to subsequent analysis and visualization.

GO/KEGG analysis, cell trajectory construction, and WGCNA of bacteriocytes

GO annotation of *G. platifrons* genes was obtained using Blast2GO software (v5.2) [84] and employed for GO enrichment analysis using homemade scripts. The number of DEGs for every GO term was first calculated, and the significantly enriched GO terms were determined using a hypergeometric test ($FDR < 0.05$). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was conducted using in-house scripts with KEGG annotations obtained from the KEGG database (<http://www.genome.jp/kegg/pathway.html>).

Single cell trajectory analysis was conducted using highly variable features (genes) identified by Seurat and Monocle (v2.6.4; <https://github.com/cole-trapnell-lab/monocle-release>) in a reversed graph embedding algorithm, and confirmed by Slingshot (<https://github.com/kstreet13/slingshot>), Cytotrace (<https://cytoTRACE.stanford.edu/>), PAGA analysis by Scanpy (v1.6.0; <https://scanpy.readthedocs.io/en/stable/>), and RNA velocity analysis by scVelo (<https://scvelo.readthedocs.io/en/stable/>). The gene expression matrix of proliferation cells and bacteriocytes was used in the analysis and visualized using the orderCells function of Monocle ($\sigma = 0.001$, $\lambda = \text{NULL}$, $\text{param. gamma} = 10$, $\text{tol} = 0.001$). For the PAGA analysis, the connectivity of each cell cluster was calculated based on the partition-based graph abstraction algorithm. After cell embedding with ForceAtlas2, the pseudo-time value of each cell was then calculated with the DPT algorithm. For the RNA velocity analysis, loom files were generated using velocity (v0.17.3) and were subsequently merged using the combine function from the loompy Python module. An AnnData object was then created using the scVelo Python module. After identifying the highly variable genes, the velocities (directionalities) were computed based on the stochastic model and projected onto the UMAP embeddings generated from Seurat. Slingshot and Cytotrace analyses were conducted in R with default parameters and using snRNA-seq data generated from Seurat. WGCNA was conducted using the WGCNA package (v1.47; <https://cran.r-project.org/web/packages/WGCNA/index.html>) in R with gene expression values of $\text{power} = 7$ and $\text{minModuleSize} = 50$, and visualized using Cytoscape (v3.8.2; <https://cytoscape.org/>).

EdU-labeling, phagocytosis, ISH assays, IF assays, 3D electron microscopy, quantitative real-time PCR, and phylogenetic analysis

The EdU-labeling assay was carried out *in situ* using mussels collected from the seepage region. Briefly, mussels were incubated with 5-ethynyl 2'-deoxyuridine (EdU, final concentration of

40 μM ; Catalog No. E10415, Thermo Fisher Scientific) in a self-designed, manually controlled macrofauna *in situ* experiment device for approximately 18 h, and then retrieved using an isothermal isobaric sampler. The click-iT plus EdU imaging kit (Catalog No. C10337, Thermo Fisher Scientific) was then used to visualize the EdU signal.

Phagocytosis assay was conducted using primary gill cells obtained from the freshly collected mussels using previously described methods with modifications [39,85]. Because endosymbiotic methanotrophs are nonculturable, *Vibrio alginolyticus*, an environmental bacterium isolated from cold seeps, was used in this assay and labeled with fluorescein isothiocyanate (FITC; Catalog No. 104157, Sigma, St. Louis, MO). To collect primary gill cells, gill tissue was first treated with 1% trypsin (diluted in sterilized seawater) at 4°C for 30 min, and then centrifuged at 300 g (4°C, 5 min) and 800 g (4°C, 5 min). Cell pellets were resuspended in modified L15 medium (Catalog No. 11415064, Gibco, Carlsbad, CA; supplemented with 0.54 g/l KCl, 0.6 g/l CaCl_2 , 1 g/l MgSO_4 , 3.9 g/l MgCl_2 , and 20.2 g/l NaCl) to a final concentration of 1×10^6 cells/ml and incubated with the same volume of FITC-labeled *V. splendidus* (1×10^8 cells/ml) for 30 min at 4°C in the dark. Primary cells were then washed three times with modified L15 medium to remove extracellular bacteria. Cells were stained with DAPI and DiI perchlorate, and imaged using a laser scanning confocal microscope (Zeiss LSM710, Jena, Germany).

FISH and double-label FISH assays were conducted as described previously [24,28]. For cell-type verification, cell markers were cloned using gene-specific primers to synthesize FISH probes. FISH analysis of symbionts was performed using a cyanine 3 (Cy3)-labeled Eub338 eubacteria probe (5'-GCTGCTC CCGTAGGAGT-3') or an FITC-labeled pmoB (methanotroph-specific gene) probe (5'-CGAGATATTATCCTCGCCTG-3') or a digoxigenin (DIG)-labeled pmoB probe. For the double-label FISH assay of bacteriocytes, Gill sections were first hybridized overnight with fluorescein-labeled cell marker probes and DIG-labeled methanotroph-specific rRNA probe. After incubations with anti-fluorescein-peroxidase (POD) (Catalog No. 11426346910, Roche, Basel, Switzerland) or anti-DIG-POD (Catalog No. 11207733910, Roche), the fluorescent signals of cell marker probes and methanotroph probes were developed by Alexa Fluor 488 tyramide reagent (Catalog No. B40932, Thermo Fisher Scientific) and Alexa Fluor 594 tyramide reagent (Catalog No. B40935, Thermo Fisher Scientific) according to the manufacturer's protocol, respectively. After mounting with DAPI, gill tissues from both FISH and double-label FISH assays were visualized and imaged using a laser scanning confocal microscope (Zeiss LSM710).

For the IF assay, unique peptide fragments of 24-hydroxycholesterol 7 α -hydroxylase CYP39A1, sugar phosphate exchanger SLC37A2, and ammonium transporter *RHBG-A* were first synthesized by Sangon Biotech (Shanghai, China) and employed as antigens to produce rabbit polyclonal antibodies. After antigen affinity purification and specificity verification by enzyme-linked immunosorbent assay (ELISA) and Western blotting, the antibodies were then used in an IF assay using paraffin-embedded gill tissue (5 μm) according to methods described previously [24], with the assistance of ServiceBio (Wuhan, China). Specifically, an anti-lipid A antibody (Catalog No. ab8467,

Abcam, Cambridge, MA) was used to indicate endosymbionts. After incubation with FITC- and Cy3-labeled secondary antibodies (Servicebio, Wuhan, China), gill sections were visualized and imaged using a laser scanning confocal microscope (Zeiss LSM710).

For the 3D electron microscopy assay, gill tissues collected from adult mussels (obtained isothermally) were fixed with paraformaldehyde-glutaraldehyde, treated with 1.0% osmium tetroxide (OsO_4), and infiltrated with acrylic resin. The embedded tissues were then serially sectioned at 100 nm using an ultramicrotome (EM UC7, Leica, Vienna, Austria) and mounted onto a silicon wafer before imaging. The scanning electron microscope micrographs of the mounted sections were captured using Helios NanoLab 600i FIB-SEM (FEI, Hillsboro, OR). The obtained serial section images were aligned in Amira (v2019.3; <https://www.thermofisher.com/hk/en/home/electron-microscopy/products/software-em-3d-vis/amira-software.html>) for 3D reconstruction of bacteriocytes and cell proliferation analysis.

Estimation of the symbiont abundance in both the InS and DeC groups was conducted using a quantitative real-time PCR assay according to the method described previously [86]. Briefly, genomic DNA from the gill tissue was extracted using the E.Z.N.A. Mollusc DNA kit (Catalog No. D3373, Omega Bio-Tek). Symbiont abundance was then calculated by comparing the relative DNA content of the symbiont methane monooxygenase gene with that of the internal control host genes (β -actin) using quantitative real-time PCR. The primers used in the assay were as described in our previous study [86]. All data were normalized using the $2^{-\Delta\Delta C_t}$ method and subjected to one-way analysis of variance (ANOVA) in SPSS (v20; <https://www.ibm.com/spss>) to detect significant differences between samples.

For phylogenetic analysis of hub transcription factors, homologous proteins were first obtained using the National Center for Biotechnology Information NCBI BLASTp and aligned using SeaView [87]. Maximum likelihood phylogenetic trees of these proteins were then constructed using MEGA software (v11) in the Jones-Taylor-Thornton model with a bootstrap of 100.

Data availability

All sequencing data have been deposited in the NCBI (BioProject: PRJNA838712), which are publicly accessible at <https://www.ncbi.nlm.nih.gov/bioproject/>. The data have also been deposited in the BioProject database at the National Genomics Data Center (NGDC), China National Center for Bioinformation (CNCB) (BioProject: PRJCA036932), which are publicly accessible at <https://ngdc.cncb.ac.cn/bioproject/>. Raw images for 3D electron microscopy have been deposited at Figshare (https://figshare.com/articles/figure/Raw_image_for_3D_electron_microscopy_of_mussel_gill/23575734).

CRedit author statement

Hao Chen: Conceptualization, Writing – original draft, Funding acquisition, Project administration. **Mengna Li:** Methodology, Investigation, Writing – original draft. **Zhaoshan Zhong:** Methodology, Investigation, Resources. **Inge Seim:** Writing –

review & editing. **Minxiao Wang:** Resources, Funding acquisition, Writing – review & editing. **Chao Lian:** Resources, Methodology, Investigation. **Lianhong Zhuo:** Investigation, Validation, Visualization. **Xinjiang Wan:** Methodology, Validation. **Hao Wang:** Methodology, Validation. **Guanghui Han:** Methodology, Validation. **Li Zhou:** Methodology, Investigation. **Huan Zhang:** Methodology, Software. **Lei Cao:** Methodology, Investigation. **Chaolun Li:** Supervision, Funding acquisition, Writing – review & editing. All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests. The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest, or non-financial interest in the subject matter or materials discussed in this manuscript.

Supplementary material

Supplementary material is available at *Genomics, Proteomics & Bioinformatics* online (<https://doi.org/10.1093/gpbjnl/qzaf109>).

Acknowledgments

The present study was supported by the National Natural Science Foundation of China (Grant Nos. 42030407, 42106134, 42327805, and 42221005), the National Key Research and Development Program of China (Grant No. 2024YFC2816100), the Marine S&T Fund of Shandong Province for Pilot National Laboratory for Marine Science and Technology (Qingdao), China (Grant No. 2022QNLM030004), and the Key Research Program of Frontier Sciences, China (Grant No. ZDBS-LY-DQC032). We thank the crew members of the R/V *Kexue* for their assistance in sample collection, and all the laboratory staff for continuous technical advice and helpful discussions. We also thank Prof. Qiang Lin from the South China Sea Institute of Oceanology, Chinese Academy of Sciences, for insightful comments and suggestions that improved the manuscript. We also express our sincere thanks to Dr. Haiyan Yu from Shandong University for the kind help in confocal microscopy imaging.

ORCID

0000-0001-6697-0809 (Hao Chen)
 0009-0005-7866-1495 (Mengna Li)
 0000-0003-1643-5407 (Zhaoshan Zhong)
 0000-0001-8594-7217 (Inge Seim)
 0000-0002-6567-3295 (Minxiao Wang)
 0009-0009-5696-737X (Chao Lian)
 0009-0008-7609-9138 (Lianhong Zhuo)
 0009-0003-7011-8397 (Xinjiang Wan)
 0000-0002-9435-5501 (Hao Wang)
 0009-0002-5040-3601 (Guanghui Han)
 0000-0001-6270-6314 (Li Zhou)

0000-0001-7269-8004 (Huan Zhang)
 0009-0003-8615-8617 (Lei Cao)
 0000-0003-1890-0587 (Chaolun Li)

References

- [1] Douglas AE. Symbiosis as a general principle in eukaryotic evolution. *Cold Spring Harbor Perspect Biol* 2014; 6:a016113.
- [2] Gilbert SF, Bosch TCG, Ledón-Rettig C. Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nat Rev Genet* 2015;16:611–22.
- [3] Nyholm SV, McFall-Ngai MJ. A lasting symbiosis: how the *Hawaiian bobtail* squid finds and keeps its bioluminescent bacterial partner. *Nat Rev Microbiol* 2021;19:666–79.
- [4] Smith TE, Moran NA. Coordination of host and symbiont gene expression reveals a metabolic tug-of-war between aphids and *Buchnera*. *Proc Natl Acad Sci U S A* 2020; 117:2113–21.
- [5] Rosset SL, Oakley CA, Ferrier-Pagès C, Suggett DJ, Weis VM, Davy SK. The molecular language of the cnidarian–dinoflagellate symbiosis. *Trends Microbiol* 2021;29:320–33.
- [6] Douglas AE. Housing microbial symbionts: evolutionary origins and diversification of symbiotic organs in animals. *Philos Trans R Soc Lond B Biol Sci* 2020;375:20190603.
- [7] McFall-Ngai MJ, Ruby EG. Symbiont recognition and subsequent morphogenesis as early events in an animal–bacterial mutualism. *Science* 1991;254:1491–4.
- [8] Braendle C, Miura T, Bickel R, Shingleton AW, Kambhampati S, Stern DL. Developmental origin and evolution of bacteriocytes in the aphid–*Buchnera* symbiosis. *PLoS Biol* 2003;1:e21.
- [9] Matsuura Y, Kikuchi Y, Miura T, Fukatsu T. Ultrabithorax is essential for bacteriocyte development. *Proc Natl Acad Sci U S A* 2015;112:9376–81.
- [10] Van Dover CL, German CR, Speer KG, Parson LM, Vrijenhoek RC. Evolution and biogeography of deep-sea vent and seep invertebrates. *Science* 2002;295:1253–7.
- [11] Sogin EM, Leisch N, Dubilier N. Chemosynthetic symbioses. *Curr Biol* 2020;30:R1137–42.
- [12] Dubilier N, Bergin C, Lott C. Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat Rev Microbiol* 2008;6:725–40.
- [13] Rodrigues CF, Laming SR, Gaudron SM, Oliver G, Le Bris N, Duperron S. A sad tale: has the small mussel *Idas argenteus* lost its symbionts. *Biol J Linn Soc* 2015;114:389–405.
- [14] Fontanez KM, Cavanaugh CM. Evidence for horizontal transmission from multilocus phylogeny of deep-sea mussel (Mytilidae) symbionts. *Environ Microbiol* 2014; 16:3608–21.
- [15] Miyazaki JI, de Oliveira Martins L, Fujita Y, Matsumoto H, Fujiwara Y. Evolutionary process of deep-sea *Bathymodiolus* mussels. *PLoS One* 2010;5:e10363.
- [16] Duperron S, Quiles A, Szafranski KM, Léger N, Shillito B. Estimating symbiont abundances and gill surface areas in specimens of the hydrothermal vent mussel *Bathymodiolus puteoserpentis* maintained in pressure vessels. *Front Mar Sci* 2016;3:16.
- [17] Sun Y, Wang M, Cao L, Seim I, Zhou L, Chen J, et al. Mosaic environment-driven evolution of the deep-sea mussel *Gigantidas platifrons* bacterial endosymbiont. *Microbiome* 2023;11:253.
- [18] Ansorge R, Romano S, Sayavedra L, Porras MÁG, Kupczok A, Tegetmeyer HE, et al. Functional diversity enables multiple symbiont strains to coexist in deep-sea mussels. *Nat Microbiol* 2019;4:2487–97.
- [19] Sun J, Zhang Y, Xu T, Zhang Y, Mu H, Zhang Y, et al. Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nat Ecol Evol* 2017;1:121.
- [20] Ponnudurai R, Kleiner M, Sayavedra L, Petersen JM, Moche M, Otto A, et al. Metabolic and physiological interdependencies in the *Bathymodiolus azoricus* symbiosis. *ISME J* 2017;11: 463–77.
- [21] Wang H, Zhang H, Wang M, Chen H, Lian C, Li C. Comparative transcriptomic analysis illuminates the host–symbiont interactions in the deep-sea mussel *Bathymodiolus platifrons*. *Deep Sea Res 1 Oceanogr Res Pap* 2019;151:103082.
- [22] Wentrup C, Wendeberg A, Schimak M, Borowski C, Dubilier N. Forever competent: deep-sea bivalves are colonized by their chemosynthetic symbionts throughout their lifetime. *Environ Microbiol* 2014;16:3699–713.
- [23] Franke M, Geier B, Hammel JU, Dubilier N, Leisch N. Coming together — symbiont acquisition and early development in deep-sea bathymodioline mussels. *Proc Biol Sci* 2021;288:20211044.
- [24] Wang H, Zhang H, Zhong Z, Sun Y, Wang M, Chen H, et al. Molecular analyses of the gill symbiosis of the bathymodiolin mussel *Gigantidas platifrons*. *iScience* 2021;24:101894.
- [25] Levy S, Elek A, Grau-Bové X, Menéndez-Bravo S, Iglesias M, Tanay A, et al. A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell* 2021;184:2973–87.e18.
- [26] Alacid E, Richards TA. A cell–cell atlas approach for understanding symbiotic interactions between microbes. *Curr Opin Microbiol* 2021;64:47–59.
- [27] Hu M, Zheng X, Fan CM, Zheng Y. Lineage dynamics of the endosymbiotic cell type in the soft coral *Xenia*. *Nature* 2020;582:534–8.
- [28] Wang H, He K, Zhang H, Zhang Q, Cao L, Li J, et al. Deciphering deep-sea chemosynthetic symbiosis by single-nucleus RNA-sequencing. *Elife* 2024;12:RP88294.
- [29] Wang H, Xiao H, Feng B, Lan Y, Fung CW, Zhang H, et al. Single-cell RNA-seq reveals distinct metabolic “microniches” and close host–symbiont interactions in deep-sea chemosynthetic tubeworm. *Sci Adv* 2024; 10:eadh3053.
- [30] Chu H, Mazmanian SK. Innate immune recognition of the microbiota promotes host–microbial symbiosis. *Nat Immunol* 2013;14:668–75.
- [31] Chen H, Wang M, Zhang H, Wang H, Lv Z, Zhou L, et al. An LRR-domain containing protein identified in *Bathymodiolus platifrons* serves as intracellular recognition receptor for the endosymbiotic methane-oxidation bacteria. *Fish Shellfish Immunol* 2019;93:354–60.
- [32] Grabovec IP, Smolskaya SV, Baranovsky AV, Zhabinskii VN, Dichenko YV, Shabunya PS, et al. Ligand-binding properties

- and catalytic activity of the purified human 24-hydroxycholesterol 7 α -hydroxylase, CYP39A1. *J Steroid Biochem Mol Biol* 2019;193:105416.
- [33] Takishita K, Takaki Y, Chikaraishi Y, Ikuta T, Ozawa G, Yoshida T, et al. Genomic evidence that methanotrophic endosymbionts likely provide deep-sea *Bathymodiolus* mussels with a sterol intermediate in cholesterol biosynthesis. *Genome Biol Evol* 2017;9:1148–60.
- [34] Hinde R, Trautman DA. Symbiosomes. In: Seckbach J, editor. *Symbiosis: mechanisms and model systems*. Dordrecht: Springer; 2002, p.207–20.
- [35] Laming SR, Gaudron SM, Duperron S. Lifecycle ecology of deep-sea chemosymbiotic mussels: a review. *Front Mar Sci* 2018;5:282.
- [36] Weaver PPE, Billett DSM, Van Dover CL. Environmental risks of deep-sea mining. In: Salomon M, Markus T, editors. *Handbook on marine environment protection*. Cham: Springer; 2018:215–45.
- [37] Gross R, Vavre F, Heddi A, Hurst GDD, Zchori-Fein E, Bourtzis K. Immunity and symbiosis. *Mol Microbiol* 2009; 73:751–9.
- [38] Tame A, Maruyama T, Yoshida T. Phagocytosis of exogenous bacteria by gill epithelial cells in the deep-sea symbiotic mussel *Bathymodiolus japonicus*. *R Soc Open Sci* 2022; 9:211384.
- [39] Li Y, Song X, Wang W, Wang L, Yi Q, Jiang S, et al. The hematopoiesis in gill and its role in the immune response of Pacific oyster *Crassostrea gigas* against secondary challenge with *Vibrio splendidus*. *Dev Comp Immunol* 2017; 71:59–69.
- [40] Pichon D, Cudennec B, Huchette S, Djediat C, Renault T, Paillard C, et al. Characterization of abalone *Haliotis tuberculata*-*Vibrio harveyi* interactions in gill primary cultures. *Cytotechnology* 2013;65:759–72.
- [41] Scott CC, Vacca F, Gruenberg J. Endosome maturation, transport and functions. *Semin. Cell Dev Biol* 2014;31:2–10.
- [42] Tame A, Maruyama T, Ikuta T, Chikaraishi Y, Ogawa NO, Tsuchiya M, et al. mTORC1 regulates phagosome digestion of symbiotic bacteria for intracellular nutritional symbiosis in a deep-sea mussel. *Sci Adv* 2023;9:eadg8364.
- [43] Chen H, Wang M, Li M, Lian C, Zhou L, Zhang X, et al. A glimpse of deep-sea adaptation in chemosynthetic holobionts: depressurization causes DNA fragmentation and cell death of methanotrophic endosymbionts rather than their deep-sea *Bathymodiolinae* host. *Mol Ecol* 2021; 30:2298–312.
- [44] Li M, Chen H, Wang M, Zhong Z, Lian C, Zhou L, et al. Phenotypic plasticity of symbiotic organ highlight deep-sea mussel as model species in monitoring fluid extinction of deep-sea methane hydrate. *Sci Total Environ* 2025; 958:178048.
- [45] Ponnudurai R, Sayavedra L, Kleiner M, Heiden SE, Thurmer A, Felbeck H, et al. Genome sequence of the sulfur-oxidizing *Bathymodiolus thermophilus* gill endosymbiont. *Stand Genomic Sci* 2017;12:50.
- [46] Geier B, Sogin EM, Michellod D, Janda M, Kompauer M, Spengler B, et al. Spatial metabolomics of *in situ* host-microbe interactions at the micrometre scale. *Nat Microbiol* 2020;5:498–510.
- [47] Chou JY, Mansfield BC. Chapter Ten — the SLC37 family of sugar-phosphate/phosphate exchangers. In: Bevensee MO, editor. *Current Topics in Membranes*, vol 73. London: Academic Press; 2014, p.357–82.
- [48] Almeida C. A potential third-order role of the host endoplasmic reticulum as a contact site in interkingdom microbial endosymbiosis and viral infection. *Env Microbiol Rep* 2021;13:255–71.
- [49] Williamson G, Harris T, Bizior A, Hoskisson PA, Pritchard L, Javelle A. Biological ammonium transporters: evolution and diversification. *FEBS J* 2024;291:3786–810.
- [50] Cui G, Konciute MK, Ling L, Esau L, Raina JB, Han B, et al. Molecular insights into the Darwin paradox of coral reefs from the sea anemone *Aiptasia*. *Sci Adv* 2023;9:eadf7108.
- [51] Felbeck H, Jarchow J. Carbon release from purified chemoautotrophic bacterial symbionts of the hydrothermal vent tubeworm *Riftia pachyptila*. *Physiol Zool* 1998;71:294–302.
- [52] Duperron S, Lorion J, Samadi S, Gros O, Gaill F. Symbioses between deep-sea mussels (Mytilidae: Bathymodiolinae) and chemosynthetic bacteria: diversity, function and evolution. *C R Biol* 2009;332:298–310.
- [53] Cao R, Zhang Y. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* 2004;14:155–64.
- [54] Brembeck FH, Opitz OG, Libermann TA, Rustgi AK. Dual function of the epithelial specific ets transcription factor, ELF3, in modulating differentiation. *Oncogene* 2000; 19:1941–9.
- [55] Nakamura T, Sakai K, Nakamura T, Matsumoto K. Hepatocyte growth factor twenty years on: Much more than a growth factor. *J Gastroen Hepatol* 2011;26:188–202.
- [56] Yang S, Ning G, Hou Y, Cao Y, Xu J, Wu J, et al. Myoneurin regulates BMP signaling by competing with Ppm1a for Smad binding. *iScience* 2022;25:104495.
- [57] Jacob F. Evolution and tinkering. *Science* 1977;196:1161–6.
- [58] Sun J, Chen C, Miyamoto N, Li R, Sigwart JD, Xu T, et al. The scaly-foot snail genome and implications for the origins of biomineralised armour. *Nat Commun* 2020;11:1657.
- [59] Altura MA, Stabb E, Goldman W, Apicella M, McFall-Ngai MJ. Attenuation of host NO production by MAMPs potentiates development of the host in the squid-vibrio symbiosis. *Cell Microbiol* 2011;13:527–37.
- [60] Foster JS, Apicella MA, McFall-Ngai MJ. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* 2000;226:242–54.
- [61] Negri I. Wolbachia as an “infectious” extrinsic factor manipulating host signaling pathways. *Front Endocrinol (Lausanne)* 2012;2:115.
- [62] Bien T, Hambleton EA, Dreisewerd K, Soltwisch J. Molecular insights into symbiosis — mapping sterols in a marine flatworm-algae-system using high spatial resolution MALDI-2-MS imaging with ion mobility separation. *Anal Bioanal Chem* 2021;413:2767–77.
- [63] Kok LT, Norris DM, Chu HM. Sterol metabolism as a basis for a mutualistic symbiosis. *Nature* 1970;225:661–2.
- [64] Hambleton EA, Jones VAS, Maegle I, Kvaskoff D, Sachsenheimer T, Guse A. Sterol transfer by atypical

- cholesterol-binding NPC2 proteins in coral-algal symbiosis. *Elife* 2019;8:e43923.
- [65] Hardie DG. AMP-activated protein kinase: maintaining energy homeostasis at the cellular and whole-body levels. *Annu Rev Nutr* 2014;34:31–55.
- [66] Leprévier G, Rotblat B. How does mTOR sense glucose starvation? AMPK is the usual suspect. *Cell Death Discov* 2020; 6:27.
- [67] Cui G, Liew YJ, Li Y, Kharbatia N, Zahran NI, Emwas AH, et al. Host-dependent nitrogen recycling as a mechanism of symbiont control in Aiptasia. *PLoS Genet* 2019;15:e1008189.
- [68] Xiang T, Lehnert E, Jinkerson RE, Clowes S, Kim RG, DeNofrio JC, et al. Symbiont population control by host-symbiont metabolic interaction in Symbiodiniaceae-cnidarian associations. *Nat Commun* 2020;11:108.
- [69] Beinart RA, Sanders JG, Faure B, Sylva SP, Lee RW, Becker EL, et al. Evidence for the role of endosymbionts in regional-scale habitat partitioning by hydrothermal vent symbioses. *Proc Natl Acad Sci U S A* 2012;109:E3241–50.
- [70] Breusing C, Genetti M, Russell SL, Corbett-Detig RB, Beinart RA. Horizontal transmission enables flexible associations with locally adapted symbiont strains in deep-sea hydrothermal vent symbioses. *Proc Natl Acad Sci U S A* 2022;119: e2115608119.
- [71] Ücker M, Ansoorge R, Sato Y, Sayavedra L, Breusing C, Dubilier N. Deep-sea mussels from a hybrid zone on the Mid-Atlantic Ridge host genetically indistinguishable symbionts. *ISME J* 2021;15:3076–83.
- [72] Alvarez-Dominguez JR, Melton DA. Cell maturation: hallmarks, triggers, and manipulation. *Cell* 2022;185:235–49.
- [73] Cao L, Lian C, Zhang X, Zhang H, Wang H, Zhou L, et al. *In situ* detection of the fine scale heterogeneity of active cold seep environment of the Formosa Ridge, the South China Sea. *J Mar Syst* 2021;218:103530.
- [74] Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, participants in the 1st Human Cell Atlas Jamboree, et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol* 2019;20:63.
- [75] Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 2015;33:495–502.
- [76] Pullin JM, McCarthy DJ. A comparison of marker gene selection methods for single-cell RNA sequencing data. *Genome Biol* 2024;25:56.
- [77] Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* 2015;16:278.
- [78] Chen Y, Zhang L, Shi X, Han J, Chen J, Zhang X, et al. Characterization of the nucleus pulposus progenitor cells via spatial transcriptomics. *Adv Sci* 2024;11:2303752.
- [79] Ji K, Chen L, Wang X, Wen B, Yang F, Deng W, et al. Integrating single-cell RNA sequencing with spatial transcriptomics reveals an immune landscape of human myometrium during labour. *Clin Transl Med* 2023;13:e1234.
- [80] Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 2015; 12:357–60.
- [81] Sun Y, Wang M, Zhong Z, Chen H, Wang H, Zhou L, et al. Adaptation to hydrogen sulfide-rich environments: strategies for active detoxification in deep-sea symbiotic mussels, *Gigantidas platifrons*. *Sci Total Environ* 2022;804:150054.
- [82] Anders S, Pyl PT, Huber W. HTSeq — a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31:166–9.
- [83] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- [84] Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005;21:3674–6.
- [85] Chen H, Zhou Z, Wang H, Wang L, Wang W, Liu R, et al. An invertebrate-specific and immune-responsive microRNA augments oyster haemocyte phagocytosis by targeting Cgkx2. *Sci Rep* 2016;6:29591.
- [86] Sun Y, Wang M, Li L, Zhou L, Wang X, Zheng P, et al. Molecular identification of methane monooxygenase and quantitative analysis of methanotrophic endosymbionts under laboratory maintenance in *Bathymodiolus platifrons* from the South China Sea. *PeerJ* 2017;5:e3565.
- [87] Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 2010;27:221–4.