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

Hao Wang[1,2,3#], Kai He[4#], Huan Zhang[1#], Quanyong Zhang[5#], Lei Cao[1], Jing Li[1,6], Zhaoshan Zhong[1], Hao Chen[1], Li Zhou[1], Chao Lian[1], Minxiao Wang[1], Kai Chen[5], Pei-Yuan Qian[2,3*], and Chaolun Li[1,6,7*]

[1]Center of Deep-Sea Research, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China

[2]Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Nansha, Guangzhou, China

[3]Department of Ocean Science, Hong Kong University of Science and Technology, Hong Kong S.A.R, China.

[4]Key Laboratory of Conservation and Application in Biodiversity of South China, School of Life Sciences, Guangzhou University, Guangzhou, Guangdong 510006, China

[5]State Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, China

[6]South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China

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

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

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

#These authors contributed equally to this work

The authors declare no conflict of interest.

Abstract:

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

Introduction:

Mutualistic interactions between multicellular animals and their microbiota play a fundamental role in animals' adaptation, ecology, and evolution [1,2]. By associating with symbionts, host animals benefit from the metabolic capabilities of their symbionts and gain fitness advantages that allow them to thrive in habitats they could not live in on their own[1,3]. Prime examples of such symbioses are bathymodioline mussels and gammaproteobacteria endosymbionts[4]. The bathymodioline mussels occur worldwide at deep-sea chemosynthetic ecosystems, such as cold-seeps, hydrothermal vents, and whale falls[4]. The host mussel acquired sulfur-oxidising (SOX) and/or methane-oxidising (MOX) symbiont through horizontal transmission at their early life stage[3]. Their symbionts, which are hosted in a specialised gill epithelial cell, namely, the bacteriocytes[5,6], utilise the chemical energy from the reduced chemical compounds such as CH_4 and H_2S released from cold seeps or hydrothermal vents to fix carbon and turn into carbon source for the host mussel[7,8]. The ecological success of the bathymodioline symbioses is apparent: the bathymodiolin mussels are often among the most dominant species in the deep-sea chemosynthetic ecosystems[9,10]. Thus, it is critical to know how the bathymodioline mussels interact with the symbionts and maintain the stability and efficiency of the symbiosis.

The gill of bathymodiolin mussels has adapted to deep-sea chemosynthetic life at the levels of molecule, cell and tissue[11,12]. Compared to shallow water mussels[13], the gill filaments of bathymodiolin mussels have enlarged surfaces allowing them to hold more symbionts per filament (Supplementary Fig. S1). This adaption requires not only novel cellular and molecular mechanisms to maintain and support the enlarged gill filament structure but also a strong ciliary ventilation system to pump vent/seep fluid to fuel the symbiotic bacteria. Previous studies based on whole-genome sequencing and bulk RNA-seq projects have shown that genes in the categories of nutrient transporters, lysosomal proteins, and immune receptors are either expanded in the host mussel's genome or up-regulated in the gill[12, 14-17], suggesting that these genes involved in host-symbiont interaction. Though providing deep molecular insights, these studies mainly used homogenised tissues that average genes' expression levels amongst different cell types and eliminate cell and gene spatial distribution information[18-21]. In addition,

44 the broad expression and function of potentially 'symbiosis-related' proteins also greatly
45 limited data interpretation. Therefore, a systemic atlas of gill cell types and the descriptions of
46 cell-type-specific gene expressional profiles are warranted to a better understanding of the
47 host-symbiont interaction and environmental adaptation mechanisms of the bathymodioline
48 symbiosis.

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

Results and discussion

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

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

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

To minimise the dropout effect in snRNAseq, we utilised a reciprocal PCA (RPCA) strategy to project cells in the three samples onto the same space based on conserved expressed genes among them[25]. This strategy maximises the number of cells per cluster regardless of the organism's state and therefore maximises genes per cluster[26]. We revealed 14 cell clusters and their corresponding marker genes (Figs. 1B- D) with Seurat[27,28]. Because there are few available canonical marker genes per cell type for *G. platifrons* and molluscs, we carefully characterised each cell cluster by 1) examining the cluster marker genes' functions, 2) identifying the expression pattern of cluster marker gene using whole-mount *in situ* hybridisation (WISH) and/or double-fluorescent *in situ* hybridisation (FISH) analyses, 3) conducting scanning electron microscopy (SEM) and transmission electron microscopy (TEM) analyses, and 4) evaluating the stability per cluster using a bootstrapping algorithm. We then successfully identified 13 of the 14 cell clusters, which could be generally divided into four groups, including 1) the supportive cells, 2) the ciliary cells, 3) the proliferation cells, and 4) the bacteriocytes. One cell cluster remained ambiguous. Notably, a bootstrapping strategy suggested that all clusters were highly supported when using all three samples (Supplementary Fig. S2A). When using each sample

individually, most clusters were also well supported except of the three ciliary cells (Supplementary Fig. S2 B-D). This could be due to that these cells were derived from the same precursor cells, and at least partially because they were represented by fewer numbers of nuclei and thus fewer available genes, especially in the food groove ciliary cell (Supplementary Table S1). The spatial and functional annotation of the cell clusters revealed that the cell clusters worked in close cooperation to support the symbiont and maintain a high-efficiency chemosynthetic system.

2. Cell atlas of *G. platifrons gill*

Supportive cells: We have identified four supportive cell types, namely, inter lamina cells basal membrane cells (BMC)1, BMC2, and mucus cells (Figs 2A and 2B).

The inter lamina cells, marked by high expression of fibrillar-forming collagens, were the cells located between the two layers of basal membranes. These cells were previously identified as amoeboid cells[13], and their function has not been explored. The inter lamina cells were densely distributed around the food groove and at the rim of the gill filament. While on the middle part of the gill filament, the inter lamina cells distributed in parallel rows. It might help to connect the two sheets of basal membranes and maintain the spatial integrity of the gill filament. In addition, we observed the enrichment of ribosomal proteins involved in protein synthesis (Supplementary Table S2), indicating a high metabolic rate[29] of inter lamina cells. This implies that the inter lamina cells may process the nutrients acquired from the symbiont.

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

119 stabilising the epithelial-derived cells, such as bacteriocytes and intercalary cells, on the surface
120 of the basal membrane.
121 Mucus cells are specialised secretory cells with intracellular mucus vacuoles[13]. In shallow-
122 water filter-feeding bivalves, mucus cells secret mucus, which cooperates with the ciliary
123 ventilation system to capture, process, and transport food particles to their mouths, known as
124 filter feeding mechanism[31-34]. Deep-sea mussels may not necessarily retain this function since
125 they do not normally acquire food resources such as phytoplankton and planktonic bacteria but
126 obtain most of nutrient through their chemosynthetic symbiont. Thus, it was hypothesised that
127 mucus cells are involved in other biological functions such as immune responses to pathogens[18].
128 Herein, genes encoding proteins with microbe-binding functionalities were enriched in mucus
129 cells (Fig. 2B and Supplementary Table S2), such as C-type lectins (Bpl_scaf_58694-0.22 and
130 Bpl_scaf_38133-0.8), C1q domain-containing protein (Bpl_scaf_55216-2.11), and Fibrinogen C
131 domain-containing protein (Bpl_scaf_18519-2.15)[35,36]. The expression of similar immune genes
132 was upregulated in a shallow water mussel *Mytilus galloprovincialis* when challenged by a
133 pathogenic bacteria[37]. Our WISH and double-FISH analyses showed that mucus cells were
134 embedded within the outer rim cilia and scattered on the gill lamella alongside the
135 bacteriocytes (Figs. 2G and 2K). These data collectively suggest that mucus cells may help
136 mussel maintain the immune homeostasis of gill. Interestingly, WISH and double-FISH analyses
137 showed that mucus cells were also distributed alongside the gill lamella's food groove and the
138 inner edge, where the density of bacteriocytes is low (Fig. 2G). Because the food groove is the
139 main entry of food practical to the labial palps and mouth[38], this distribution pattern implies
140 that mucus cells may be also involved in capturing planktonic bacteria and sending them to the
141 mouth.
142 **Ciliary and smooth muscle cells:** A remarkable feature of bathymodioline mussel's gill is its
143 ciliary ventilation system, which constantly agitates the water and provides the symbiont with
144 the necessary gas[39]. We identified four types of ciliary cells (Figs. 3A and 3B): apical ridge ciliary
145 cells (ARCCs), food groove ciliary cells (FGCCs), lateral ciliary cells (LCCs), and intercalary cells
146 (ICs), as well as a newly identified type of smooth muscle cells. All ciliary cells were marked by
147 canonical cilium genes, such as genes encoding flagellar proteins, ciliary motor kinesin proteins,

148 ciliary dynein proteins, and ciliary microtubules and were clearly distinguishable by specifically
149 expressed genes (Supplementary Table S3 and Supplementary Fig. S3). ARCCs were
150 characterized by expression of Bpl_scaf_20631-1.16, which encodes homeobox Dlx6a-like
151 protein (Fig 3D and 3F) which is a marker of the apical ectodermal ridge[40]. FGCCs expressed
152 genes encoding primary cilia development regulator Tubby-related protein 3[41]
153 (Bpl_scaf_24834-2.3) and primary ciliary cell structural protein TOG array regulator of axonemal
154 microtubules protein 1[42,43] (Bpl_scaf_55620-1.3), suggesting the FGCCs are the sensory ciliary
155 cells that gather information from the surrounding environment. Both ARCCs and FGCCs were
156 located around the ventral tip of the gill filament. LCCs were distributed as two parallel rows
157 along the gill lamella's outer rim and ciliary disks' outer rim (Figs. 3G and 3H) and had highly
158 expressed genes involved in cilium structure, cilium movement and ATP synthases (Fig. 3B and
159 Supplementary Table S3), indicating that LCCs may have a strong ability to beat their cilia.
160 Interestingly, we identified a group of previously unreported smooth muscle cells (SMCs) co-
161 localised with the LCCs as showed by WISH (Fig. 3I and Supplementary Fig. S4). SMCs strongly
162 expressed several Low-density lipoprotein receptors (LDL receptor) and LDLR-associated
163 proteins[44] (Fig. 3B). In addition, these cells also expressed the angiotensin-converting enzyme-
164 like protein[45] (Bpl_scaf-29477-5.10) and the 'molecular spring' titin-like protein[46]
165 (Bpl_scaf_56354-7.8) (Supplementary Table S3). The expression of these genes could be
166 commonly found in human vascular smooth muscle cells[47-49]. Collectively, we suspect that SMCs
167 are involved in lateral cilium movement and the gill slice contraction.
168 ICs are the specialised ciliary cells surrounding the bacteriocytes (Figs. 3B and 3J). WISH
169 analysis showed that the expressions of IC markers had apparent spatial variations (Fig. 3J). The
170 expression of ICs marker rootletin[50-52] (Bpl_scarf_22362-6.9) was considerably higher in the ICs
171 close to the frontal edge of the gill filament, and the expression gradually decreased along with
172 the direction of inter lamina water flow (Fig. 3K and Supplementary Fig. S5), implying that ICs
173 ventilate the water flow and the mucus through the gill filaments. Furthermore, compared with
174 the other three types of ciliary cells, the ICs expressed several genes encoding transcription
175 factors involved in determining cell fate (Fig. 3B and Supplementary Table S3), such as
176 transcription factors Sox 8 (Bpl_scarf_40595-5.30) and Wnt pathway cell polarity regulator

177 secreted frizzled-related protein 5 (Bpl_scaf_57424-6.1)[53], suggesting that the ICs might also
178 play regulatory roles[54-56].
179 **Proliferation cells**: It has long been known that the bathymodioline mussel gill has three
180 types of proliferation cells that are conserved throughout all filibranchia bivalves: the budding
181 zone at the posterior end of the gill where new gill filaments are continuously formed, and the
182 dorsal and ventral ends of each gill filaments[57,58]. These "cambial-zone"-like cell populations
183 could continuously proliferate throughout the whole life span of the mussel[58]. Our snRNA-seq
184 data recognised three types of proliferation cells, which is consistent with previous findings (Fig.
185 4A, Fig. 4B and Supplementary Table S4). The gill posterior end budding zone cells (PEBZCs) are
186 located on the first few freshly proliferated filaments of the posterior tip of gill (Figs. 4C and 4D).
187 The PBEZCs marker gradually disappeared at around the 11th -12th role of the gill filament,
188 suggesting the maturation of the gill filaments, which is similar to the developmental pattern
189 reported in another deep-sea mussel *Bathymodiolus azoricus* [59]. The dorsal end proliferation
190 cells (DEPCs), which expressed the hallmarks genes of muscular tissue, as well as cell
191 proliferation and differentiation regulators, were the proliferation cells in connective tissue at
192 the dorsal end of the gill slice (Fig. 4A, 4E). The ventral end proliferation cells (VEPCs) were two
193 symmetrical triangle-shaped cell clusters of small symbiont-free cells (Fig. 4F). In VEPCs, genes
194 encoding ribosomal proteins, chromatin proteins, RNA and DNA binding proteins, and cell
195 proliferation markers were all up-regulated (Fig. 4B and Supplementary Table S4), indicating
196 that VEPCs are meristem-like cells[59,60].
197 It has been hypothesised that new proliferation cells will be colonised by symbionts, serving
198 as the vital mechanism of bacteriocyte recruitment[59]. We then determined which proliferation
199 cell type gives rise to the mature cells, especially the bacteriocytes, which has little available
200 information regarding their precursor and transmission mode in bathymodioline mussels [59,61].
201 Slingshot trajectory analysis suggested that the PEBZCs might be the origin of all gill epithelial
202 cells, including the other two proliferation cells (VEPC and DEPC) and bacteriocytes
203 (Supplementary Fig. S6). The sole exception was BMC2, which may be derived from VEPC
204 rather than PEBZC. This result is consistent with previous studies which suggested new gill
205 filaments of the filibranch mussels are formed in the gill's posterior budding zones[59]. The

206 colonisation by the symbiont might play a crucial role in determining the fate of the
207 bacteriocytes. Noticeably, in *G. platifrons*, only the pillar-shaped first row of gill filament,
208 comprised of small meristem-like cells, was symbiont-free (Figs. 4D and D'), whereas all the
209 other gill filaments were colonised by symbionts. This pattern of symbiosis establishment is
210 different from that of *B. azoricus*, in which PEBZCs are symbiont free and are gradually
211 colonised by symbionts released from the bacteriocytes on the adjacent mature gill filaments
212 after maturation[59]. On mature gill filaments, the DEPCs and VEPCs are seemingly the sources of
213 new cells that sustain the growth of the gill filament from both dorsal and ventral directions,
214 respectively[57].
215
216 **3. Bacteriocytes and host-symbiont interaction:**
217 We conducted the whole-mount FISH using a bacterial 16S rRNA probe for symbiont to
218 determine the spatial distribution of bacteriocytes, and their positions relative to the other cell
219 types on the gill filament (Figs. 5A, B). Bacteriocytes covered the majority of the surface of the
220 gill filament, except the ventral tip, the ciliary disk, and the frontal edge (lateral ciliary). The
221 bacteriocytes were surrounded by intercalary cells with microvilli and cilia on the surface (Fig.
222 5C). Noticeably, the bacteriocytes' size and FISH signal intensity, which reflected the symbiont
223 number in each bacteriocyte, gradually decreased from the heavily ciliated frontal edge to the
224 inner edge, demonstrating that the symbiont abundance in each bacteriocyte was affected by
225 the environmental gradient[11].
226 The gene expression profile of bacteriocytes aligned well with the morphological analysis,
227 which suggested that the bacteriocytes have structural, metabolic, and regulatory adaptions to
228 cope with the symbiont (Figs. 5D, 5F and 5I). It has been hypothesised that the bacteriocytes
229 extract nutrients from the symbiont through two pathways: lysing the symbiont through the
230 lysosome (the 'farming' pathway) or directly utilising the nutrient produced by the symbiont
231 (the 'milking' pathway)[62,63]. Our TEM observations clearly detected intracellular vacuole and
232 lysosome system of the bacteriocyte that harbour, transport, and digest the symbionts (Fig. 5D),
233 which is consistent with previous studies[18,64]. The snRNA data showed bacteriocytes expressed
234 cellular membrane synthesis enzyme (phosphoethanolamine N-methyltransferase,

235 Bpl_scaf_15282-3.10) and a series of lysosomal proteins such as lysosomal proteases
236 (cathepsins; Bpl_scaf_61711-5.12, Bpl_scaf_46838-6.40 and Bpl_scaf_59648-4.5; lysosomal
237 alpha-glucosidase-like), protease regulators 56 (TMPRSS15 proteins; Bpl_scaf_52188-1.15 and
238 Bpl_scaf_15410-0.9), and lysosomal traffic regulator proteins (rabenosyn-5, Bpl_scaf_54816-0.3
239 and Bpl_scaf_52809-1.6, lysosomal-trafficking regulator-like isoform X1). Among the proteases,
240 cathepsins were thought to be evolutionary conserved molecular tools that host utilized to
241 control the residence of their symbiont microbes[65]. They were also highly expressed in
242 symbiotic tissue of other deep-sea chemosynthetic animals, such as vesicomyid clams and
243 vestimentiferan tubeworms[66-68]. Bacteriocytes also expressed genes encoding cellular vesicle
244 transports (kinesins, Bpl_scaf_14819-0.11, Bpl_scaf_54265-1.13, and Bpl_scaf_4784-1.40)[53],
245 potential amino acid transporter[57] (Bpl_scaf_36159-5.5), and genes involved in intracellular
246 vesicle transport, such as the FYVE and coiled-coil domain-containing protein 1[58]
247 (Bpl_scaf_33726-5.7). In addition to form and mobilise early endosomes, the protein products
248 of these genes could transport symbiont-secreted nutrient vesicles to the host (Fig. 5F and
249 Supplementary Table S5), supporting the "milking" pathway.
250 The symbiont of *G. platifrons* belongs to type I methanotrophy, of which the core metabolic
251 function is linked with the development of intracytoplasmic membranes leading to a high
252 lipid/biomass content[69,70]. Recent lipid biomarker analyses showed that the gill of *G. platifrons*
253 contains a high amount of bacterial lipids, which are directly utilised by the host to synthesise
254 most of its lipid contents[66]. Downstream of the bacteriocyte's metabolic cascade, genes
255 encoding proteins that may be involved in fatty acid/lipid metabolism, such as perilipin2
256 (Bpl_scaf_27158-3.8), which is the critical protein to form intracellular lipid droplets[71], and a
257 variety of fatty acid metabolism enzymes (acetyl carboxylase 2, Bpl_scaf_55250-5.11[72]; fatty
258 acid desaturase 1-like isoform X1, Bpl_scaf_35916-0.6[73,74]; long-chain fatty acid-ligase ACSBG2-
259 like isoform X1, Bpl_scaf_28862-1.5 [75,76]), were up-regulated, suggesting that the fatty acid
260 could be a major form of nutrients passing from the symbiont to the host mussel.
261 Additionally, bacteriocytes expressed several solute carriers, including sodium/ascorbate
262 cotransporter (solute carrier family 23 members 1-like, Bpl_scaf_32311-1.19),
263 sodium/potassium/calcium exchanger (sodium potassium calcium exchanger 3-like,

264 Bpl_scaf_14503-1.28), sodium/chloride ion cotransporter (solute carrier family 12 member 3-
265 like isoform X1, Bpl_scaf_44604-3.7), ferrous iron transporter (solute carrier family 40 member
266 1-like, Bpl_scaf_63447-0.12) and zinc transporter (zinc transporter ZIP14-like, Bpl_scaf_44428-
267 5.3). The solute carriers are a large family of ATP-dependent transporters that shuttle a variety
268 of small molecules across the cellular membrane[77]. In several model symbiotic systems, solute
269 carriers play a vital role in host-symbiont interaction by either providing the symbiont with
270 substrates[78,79] or transporting symbiont-produced nutrients to the host [80-82]. Previous studies
271 demonstrated that solute carrier genes were expanded in deep-sea chemosynthetic animals'
272 genome[83] and highly expressed in symbiotic organs[84] including bathymodioline mussel's gill[12].
273 In bathymodioline mussels, previous bulk RNA-seq studies detected up-regulated expression of
274 a large variety of solute carriers in the gill, which is consistent with the present study,
275 suggesting that solute carriers may play crucial roles in shuttling nutrients in and out of
276 bacteriocytes and in maintaining the suitable intracellular micro-environment for the
277 symbiont[85,86].
278
279 **4. Cell-type-specific response to environmental stresses:**
280 To examine cell-type-specific acclimatisation to environmental changes the expression
281 profiles of differentially expressed genes (DEGs) were compared between the three states
282 (Fanmao, starved and reconstituted animals) of samples collected in our *in situ* transplantation
283 experiment (full lists of DEGs are shown in Supplementary Table S6-S18). Notably, for each
284 state three mussels were processed but pooled for nuclei extraction before snRNA sequencing
285 (see Method for detail). We tested our hypothesis here that at the starvation site, a relatively
286 low concentration of methane shall upset symbiont metabolism and thus substantially affect
287 symbiont-hosting bacteriocytes by assessing the transcriptional changes per cell type. We
288 calculated the distances between the three sets of samples on the UMAP plot (Fig. 6A)
289 estimating the centroid coordinates per sample per cell type, and calculating the Euclid distance
290 between the centroid coordinates of each pair of states per cell type. The impact of starvation
291 was variable across cell types, as reflected by the cross-state distances (Fig. 6B, green bar).
292 Starvation resulted in the most significant transcriptional changes in bacteriocytes reflected by

293 a large Fanmao-vs-starvation distance, followed by VEPC and inter lamina cells. On the other
294 hand, starvation had little impact on the transcriptions of ciliary cells and most supportive cells
295 such as the food groove ciliary cell, BMC2, and mucus cells. Fig 6B also shows that after
296 reconstitution (moving the mussels back to the methane-rich site Fanmao for 3 days), the
297 expressional profile of bacteriocytes rapidly changed back, reflected by a large starvation-vs-
298 reconstitution distance and much smaller Fanmao-vs-reconstitution distance. This result
299 coincided with and was supported by our pseudo-time analysis for bacteriocytes (Fig. 6C),
300 showing that bacteriocytes in the reconstitution are in intermediate and transitional states
301 between Fanmao and starvation.

302 For the bacteriocytes population, we conducted cell trajectory pseudo-timing and detailed
303 DEG analysis to interpret the mechanism of host-symbiont interactions. The branched heatmap
304 showed distinct gene expression patterns amongst the three treatments (Fig. 6D and
305 supplementary Fig. S7). The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway
306 enrichment analysis of the bacteriocytes' DEGs provides an overall view of the pathways
307 enriched in each environmental state (Supplementary Fig. S8 A-C). The genes encoding
308 ribosomal proteins are highly expressed in the bacteriocytes under the methane-rich 'Fanmao'
309 state (Supplementary Fig. S8A), suggesting an active protein synthesis and cellular metabolism
310 [87]. Moreover, organic ion transporters (A BCB P-glycol, Bpl_scaf_18613-16.10; canalicular
311 multispecific organic anion transporter 2, Bpl_scaf_52110-4.43; sodium- and chloride-
312 dependent glycine transporter 1-like, Bpl_scaf_13642-5.8), which may be involved in
313 transporting symbiont-produced nutrients, are also highly expressed.

314 In starved *G. platifrons*, negative regulators of cell proliferation, such as autocrine
315 proliferation repressors and receptor-type tyrosine phosphatase beta, were up-regulated in
316 bacteriocytes, suggesting repression of cell growth. We also observed the enrichment of genes
317 in the apoptosis pathway (Supplementary Fig. S8B), which is attributed to the upregulation of
318 caspases (Bpl_scaf_25165-3.8 and Bpl_scaf_24225-0.12) and cathepsins (Bpl_scaf_64706-0.16,
319 Bpl_scaf_61711-5.12) which can trigger caspases-dependent cell death, suggesting cellular
320 stress condition. Correspondingly, a gene encoding baculoviral IAP repeat-containing protein
321 (Bpl_scaf_6172-0.3), which can bind caspase and inhibit apoptosis, was up-regulated in

322 bacteriocytes (Fig. 6E). In *Drosophila*, the baculoviral IAP proteins are important in animal's
323 response to cellular stress and promote cell survival[88,89]. Similarly, the gene encoding MAP3K14
324 (Bpl_scaf_60908-4.14) and potential E3 ubiquitin ligase RNF213 (Bpl_scaf_25983-1.26,
325 Bpl_scaf_1894-1.45, Supplementary Fig. S9), which were apoptosis suppressor or regulator,
326 were up-regulated[90]. These results suggest that starved bacteriocytes were carrying out cell-
327 type-specific adjustments to cope with stresses. The E3 ubiquitin ligases could also work as
328 intracellular immune sensors of bacterial lipopolysaccharides[91]. Thus, the encoded protein may
329 be able to active downstream immunological toolkit to digest the symbiont population for
330 nutrients.
331 As mentioned above, bacteriocytes obtain nutrients from endosymbionts. KEGG pathway
332 analyses suggested that phagosome, lysosome-related pathways were upregulated in the
333 starvation state, indicating that bacteriocytes were actively digesting the endosymbionts in the
334 starved *G. platifrons*. Interestingly, protein synthesis activity was more activated in inter lamina
335 cells and VEPCs as supported by the up-regulation of genes encoding ribosomal proteins in both
336 cell populations (Supplementary Tables S6 and S17). This was contrary to the situation in
337 bacteriocytes and may be a consequence of the high activity of "farming" in bacteriocytes.
338 We anticipated that after moving the starved *G. platifrons* back to Fanmao site,
339 bacteriocytes and their endosymbionts would "reconstitute" leading to the partly restored
340 "farming" and "milking" pathways. This was confirmed by higher expression of fatty acid
341 metabolic genes in the mussels in the reconstitution state than those in the starvation state,
342 such as long-chain-fatty-acid-ligase ACSBG2-like (Bpl_scaf_28862-1.5), elongation of very long-
343 chain fatty acids 7-like (Bpl_scaf_5959-1.9), and fatty acid desaturase 1-like (Bpl_scaf_35916-
344 0.6). These findings were also consistent with the result of KEGG analyses. The mitochondrial
345 trifunctional enzyme (Bpl_scaf_28862-1.5) was also highly expressed, suggesting a high level of
346 energy-producing activity. The KEGG analyses showed that glutamatergic synthase was
347 enriched in the reconstitution state in comparison with those in both the Fanmao and
348 starvation states. In the insect aphid–Buchnera endosymbiosis model, the host-produced
349 glutamate could be transported and directly utilised by the symbiont[92,93]. Similar glutamate-
350 based host–symbiont metabolic interaction mechanisms were also proposed in the deep-sea

351 mussel *Bathymodiolus thermophilus* and *B. azoricus*[63,94]. As the pseudotime analyses showed
352 above (Fig. 6C), reconstitution was the intermediate state between Fanmao and starvation.
353 Metabolic and gene regulatory functions of bacteriocytes were still different from that in
354 Fanmao. The KEGG analyses suggested regulatory (Rap1 signalling, retrograde endocannabinoid
355 signalling, chemokine signalling, glucagon signalling, thyroid hormone synthesis, mRNA
356 surveillance pathway, etc.) and metabolic pathways were enriched in the reconstitution state
357 (fatty acid metabolism, salivary secretion, aldosterone synthesis and section, insulin secretion
358 and pancreatic section) (Supplementary Fig. S8C). For example, we detected up-regulation of
359 the genes encoding carbonic anhydrases (Bpl_scaf_33596-7.3 and Bpl_scaf_48274-0.3),
360 electroneutral sodium bicarbonate exchanger 1 (Bpl_scaf_61230-5.11), and globin-like proteins
361 (Bpl_scaf_50392-1.14 and Bpl_scaf_24370-0.5), which could provide the symbiont with carbon
362 dioxide and oxygen necessary for the symbiont's chemosynthetic metabolism[83-94].
363
364 **5. Summary and outlook:**
365 Using the deep-sea mussel *G. platifrons* as a model organism, we demonstrated the power
366 of integrating snRNA-seq and WISH data in unravelling the mechanism behind animal-microbe
367 symbiosis. The robustness of this strategy showed stable and highly distinguishable expression
368 patterns of each cell type regardless of the different environmental states. We successfully
369 profiled the specific roles of different types of cells, including the previously unknown cell types,
370 in maintaining the structure and function of the gill. The supportive cells that located in
371 between (inter lamina cells) and on the basal membrane's surface (BMC1 and BMC2) helped
372 maintain the anatomical structure of the basal membrane. Ciliary and smooth muscle cells
373 involved in gill slice contraction and cilium beat allowing the gathering of material and
374 information from the surrounding environment. Proliferation cells (PBEZCs, DEPCs and VEPCs)
375 gave rise to new cells, including bacteriocytes which obtained nutrients from endosymbionts
376 using intracellular vacuoles and lysosomes.
377 The snRNA analyses also revealed that different cell types collaborated to support
378 bacteriocytes' functionality. The PBEZCs gave rise to new bacteriocytes that allowed symbiont
379 colonisation. Bacteriocytes attached to the basal membrane and were stabilised by supportive

380 cells. Mucus cells co-localise with bacteriocytes and help to maintain immune homeostasis. The
381 beating ciliary cells controlled the water flow, providing bacteriocytes with necessary inorganic
382 substances from the environment. This new information on cell-cell interaction certainly
383 advanced our overall understanding of how endosymbiotic microbes and host cells
384 communicate and collaborate, which cannot be easily achieved through other methods.
385 Moreover, the analysis of snRNA data from bacteriocytes has provided insight into the
386 molecular mechanisms employed by the host to maitain and regulate the symbiosis. Notably,
387 the bacteriocytes-enriched transcripts involved in harbouring, digesting symbionts, and
388 transporting nutrients produced by the symbionts were identify and characterised. Our *in situ*
389 transplant experiments by moving mussels between methane-rich and methane-limited sites
390 also provided clues of cell-type specific responses to environmental change. Under a methane-
391 limited environment, the staved mussels more actively consumed endosymbionts through the
392 "farming" pathway. After being moved back to the methane-rich environment, mussels
393 produced more glutamates which sustained the regrowth of symbionts. These preliminary
394 findings showed that the deep-sea mussels were able to control their endosymbionts using a
395 set of genes in response to environmental change. Due to the limitation of remotely operated
396 vehicle (ROV) dives and sampling capacity, we had to have pooled samples of each state for
397 nuclei extraction and sequencing. Thus, the cells per cluster could be considered as technical
398 than biological replicates. Although this sampling process strategy has been broadly used in
399 snRNA/scRNA sequencing[95], we recognise the possible violation of assumptions in *p*-value
400 calculation for DEGs between the three states.
401 Overall, the single-cell spatial and functional atlas developed in the present work will
402 decipher some common principles of symbiosis and environmental adaption mechanisms of
403 animals. The workflow developed in the present study could provide insightful references for
404 researchers focusing on the mechanistic study of the biological adaptation of biologically and
405 ecologically important non-model animals.
406
407 **Acknowledgements:** This study was supported by the Science and Technology Innovation
408 Project of Laoshan Laboratory (Project Number No. LSKJ202203104), the National Natural

Science Foundation of China (Grant No. 42030407), Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou)(HJ202101, SMSEGL20SC01), Major Project of Basic and Applied Basic Research of Guangdong Province (2019B030302004), the Research Grants Council of Hong Kong (C6026-19G-A and 16101822) and the Guangdong Natural Science Funds for Distinguished Young Scholar (2022B1515020033). We appreciate all the assistance provided by the crew of *R/V* 'Kexue' and the operation team of ROV 'Faxian'.

Author contributions: W.H., K.C., P.Y.Q. and L.C.L designed the study. Z.H., L. Cao, Z.Z.S., L. Chao and W.M.X conducted the *in situ* transplant experiments and collected the samples; Z.Q.Y. and W.H. constructed the snRNA-seq library. H.K. performed the bioinformatics analysis. W.H., Z.H., L. Cao, J.L., H.C. and L.Z. performed the experiments. This paper was written by W.H., H.K., K.C., P.Y.Q. and L.C.L. All authors revised the manuscript. All authors read, approved, and contributed to the final manuscript.

Competing interests: The authors declare no competing interests.

425 ## Figure legends

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

437

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

447

448 **Figure 3:** Ciliary cell populations of *G. platifrons* gill. A: UMAP representation of the four ciliary
449 cell populations and potential smooth muscle cell population. B: Expression profiles of the cell
450 markers that are specific or enriched in the ciliary cell populations. The sizes of the circles
451 represent the percentages of cells in those clusters that expressed a specific gene. The genes
452 shown in red were validated by WISH or double FISH. C-E, G and I-J: WISH characterisation of
453 the selected representative cell population markers. F and H: SEM analysis of the ciliary cells of

454 *G. platifrons* gill. K: Schematic of the water flow agitated by different ciliary cells. The marker
455 genes confirmed by WISH in the current study are indicated in red.
456
457 **Figure 4:** Proliferation cell populations of *G. platifrons* gill. A: UMAP representation of the three
458 proliferation cell populations. B: Expression profiles of the cell markers that are specific or
459 enriched in the supportive cell populations. The sizes of the circles represent the percentages of
460 cells in those clusters that expressed a specific gene. Genes shown in red were validated by
461 WISH or double FISH. C and D: Photographic and schematic analyses of the spatial position of
462 the three Proliferation cell populations. E, F and J: FISH and WISH characterisation of the
463 selected population markers. The marker genes confirmed by ISH or WISH in the current study
464 are indicated in red. Scale bar: 50 µm.
465
466 **Figure 5:** Characterisation of the bacteriocytes of *G. platifrons*. A: Schematic of the overall
467 structure of *G. platifrons* gill filaments. B: Whole-mount FISH analyses of the overall distribution
468 of bacteriocytes on the *G. platifrons* gill filament. C: SEM analysis of the bacteriocytes. D: TEM
469 analysis of a bacteriocyte. E: UMAP representation of *G. platifrons* bacteriocytes. F: Expression
470 profiles of the cell markers that are specific or enriched in the bacteriocytes. The sizes of the
471 circles represent the percentages of cells in those clusters that expressed a specific gene.
472 Double FISH validated the genes shown in red. I: Schematic of the host-symbiont interaction
473 based on the single-cell transcriptome of *G. platifrons* bacteriocytes. The marker genes
474 confirmed by ISH in the current study are indicated in red.
475
476 **Figure 6:** Analysis of cell population-specific DEGs. A: UMAP representation of the impact of
477 deep-sea *in situ* transplant treatments on the gene expression pattern of each cell population.
478 The cells from different treatments were labelled with different colours. The dashed line
479 encircled bacteriocyte populations have a considerably altered expression profile. B: Histogram
480 of cross-state distances between the centroids of the Fanmao, Starvation and Reconstitution
481 groups per cell type on UMAP. The black dashed lines indicate the bacteriocyte populations
482 whose expression profile was remarkably altered. C: Visualization of bacteriocytes onto the

483 pseudo time map using monocle. The black lines indicate the main path of the pseudotime
484 ordering of the cells. D: Bifurcation of selected gene expression along two branches in response
485 to environmental perturbation. Genes are clustered hierarchically into two groups, illustrating
486 up- (cluster 1) and down- (cluster2) regulated genes in the starvation state compared with
487 Fanmao. Genes in red colour were discussed in Section 4. The heat map showing the gene
488 expression profiles of all bacteriocytes' DEGs is shown in supplementary Fig. S7; E: Proposed
489 model for the molecular mechanisms of host-symbiont interactions in response to
490 environmental changes.
491
492

493 ## Materials and methods:

494 **Deep-sea *in situ* transplant experiment and sample collection:** *In situ* transplant experiment
495 was conducted at the 'F-site' cold seep during the R/V 'Kexue' 2020 South China Sea cold seep
496 cruise. The overall design of the *in situ* transplant experiment and environmental states are
497 shown in Figure 1. First, mussels in the methane-rich 'Fanmao' site (meaning 'prosperous' site,
498 22° 06′ 55.425″ N,119° 17′ 08.287″ E, depth 1117m) were scoped into three nylon bags
499 with approximately 10 mussels in each bag. Then, two bags of mussels were transplanted to the
500 low methane Starvation site (Figure 1, 22° 07′ 00″ N, 119° 17′ 07.02″ E, depth 1147.42m).
501 After 11 days of transplantation, one bag of mussels in the Starvation site was moved back to
502 the 'Fanmao' site. On the 14th day of the transplant, three bags of mussels: one bag from the
503 Fanmao site (designated as the 'Fanmao' sample), one bag from the starvation site (designated
504 as the 'Starvation' sample) and one bag of mussels which were first transplanted to the
505 Starvation site for 11 days and moved back to the Fanmao site for 3 days (designated as the
506 'Reconstitution' sample) were all retrieved by ROV Fanxian. The mussels were kept in a
507 hydraulic pressure-sealed biobox during the ascending of the ROV. The biobox is made of heat-
508 insulated material, which will prohibit heat exchange with warm surface water. The samples
509 were immediately processed once onboard R/V Kexue. For snRNA-seq, the posterior end tip of
510 the mussel's gill was dissected (Figure 1C), snap-frozen with liquid nitrogen and then stored at
511 −80 °C until use. For WISH, ISH and FISH, the gills of the mussels were first fixed with 4%

512 paraformaldehyde (PFA, prepared with autoclaved 0.22 µM membrane filtered *in situ* seawater)
513 at 4 °C overnight. Then, the gill was washed with ice-cold 1× PBS three times, dehydrated and
514 stored in 100% methanol at −20 °C.
515
516 **Single-nucleus RNA-sequencing of *G. platifrons*:** The gill nucleus was extracted using the Nuclei
517 PURE prep nuclei isolation kit (Sigma-Aldrich). For each sample, posterior end tips from 3
518 individual mussels were randomly selected and pooled together. The posterior tips were
519 homogenised in 10 mL of ice-cold lysis solution (Nuclei PURE lysis buffer with 0.1% Triton X100
520 and 1 mM DTT) on ice. The cell nuclei were then separated by sucrose gradient centrifugation
521 according to the manufacturer's protocol. The cell nuclei pellets were washed by re-suspending
522 in ice-cold DPBS–BSA solution (1× DPBS, 0.04% nuclease-free BSA, 0.01% RNase inhibitor;
523 Takara). The nucleus was spun down by centrifugation at 500×g for 5 min at 4 °C. This step was
524 repeated to remove the contaminants from the cell plasma. Finally, the nucleus was re-
525 suspended in the DPBS–BSA solution. The concentration of the cell nucleus was counted by Cell
526 Countess II. Single-nucleus RNA-seq libraries were then constructed with the BD Rhapsody™
527 single-cell analysis system using the BD Rhapsody WTA Amplification Kit according to the
528 manufacturer's protocol. The libraries were then sequenced with the Illumina HiSeq 4000
529 platform. The clean reads of three datasets were submitted to the National Centre for
530 Biotechnology Information Sequence Read Archive database (Bioproject: PRJNA779258).
531
532 **Bioinformatics:** Three raw datasets (Fanmao, Starvation and Reconstitution) were processed
533 individually by following the BD single-cell genomics analysis setup user guide (Doc ID: 47383
534 Rev. 8.0) and BD single-cell genomics bioinformatics handbook (Doc ID: 54169 Rev. 7.0). In brief,
535 a *G. platifrons* genome reference v1.0 (the genome could be downloaded at Dryad Digital
536 Repository http://dx.doi.org/10.5061/dryad.h9942)[15] was indexed and built using STAR 2.5.2b.
537 Gene names were prepared according to the BD-single cell guideline. The sequencing data were
538 then mapped to the indexed genome using the BD Rhapsody single-nucleus pipeline v1.9 with
539 default parameters to generate an expression matrix per dataset.

540 Each data matrix was converted to a SingleCellExperiment data format, and empty barcodes
541 were removed using the emptyDrops function of DropletUtils v3.14[96]. Then, we converted and
542 processed the data in Seurat v3[28]. We removed cells that had <100 or > 2500 genes and <100
543 or > 6000 unique molecular identifiers (UMIs). We also removed genes that had <10 UMIs in
544 each data matrix. Then, we log-normalised the data and used DoubletFinder v2.0[97] to remove
545 potential doublet, assuming a 7.5% doublet formation rate. The numbers of retained nuclei
546 were 9717, 21614 and 28928 for Fanmao, Starvation and Reconstitution data, respectively
547 (Supplementary Table S19). We used the top 3000 highly variable genes for principal
548 component analysis (PCA) and a reciprocal PCA approach to integrating the three datasets[98].
549 The first 40 principal components (PCs) were used for uniform manifold approximation and
550 projection (UMAP) dimensional reduction and following clustering (Supplementary Fig. S10).
551 We used an empirical parameter, a resolution of 0.2, in the FindClusters function. The average
552 expression of each gene per cluster were also provided as supplementary data S1. Based on the
553 WISH results (see below), we manually combined the clusters with similar gene expressions and
554 WISH patterns. Eventually, we assigned all the cells to 14 reliable clusters (Supplementary Fig.
555 S11). We adopted a bootstrap sampling strategy to evaluate the stability for each cluster[99]. We
556 ran bootstrapping with 100 iterations using cells in the three samples together and in each
557 sample individually. We determined marker genes per cluster using the FindAllMarkers function
558 based on the Wilcoxon rank sum test. The marker genes were used for KEGG enrichment using
559 ClusterProfiler v4.2[100].
560
561 **Cell trajectory analysis:** We conducted slingshot trajectory analyses using Slingshot[101] v2.2.1
562 for a subset of clusters ('intercalary cells,' 'dorsal-end proliferation cells,' 'ventral-end
563 proliferation cells,' 'mucus cells,' 'basal membrane cells 1,' 'basal membrane cells 2' and
564 'bacteriocytes') to explore the developmental trajectory of cells, assuming that all these cells
565 were developed from the same precursor ('posterior-end budding zone'). We performed PCA
566 for the subsampled data set and conducted dimensional reduction using phateR v1.0.7 for
567 slingshot analyses. We used the Potential of Heat-diffusion for Affinity-based Trajectory
568 Embedding (PHATE) because it could better reveal developmental branches than other tools[102].

569 We also examined the effect of deep-sea transplant experiments on shaping gene expression
570 patterns by comparing the expression levels amongst the three different states of a given
571 cluster. This comparison was done for the bacteriocytes. The biased number of cells per state
572 could affect the results of the dimensional reduction and calculation of marker genes, and the
573 sequenced nucleus per state was unbalanced; therefore, we first downsampled the cells per
574 cluster per state to about the same number. Thereafter, we performed PCA and dimensional
575 reduction using UMAP and PHATE, calculated marker genes and conducted slingshot trajectory
576 and KEGG enrichment analyses as mentioned above for each cell type.
577
578 **Phylogenetic estimation:** Phylogenetic estimation was conducted for E3 ubiquitin ligase
579 RNF213 genes. We downloaded RNF213 genes from GenBank for representative vertebrate and
580 invertebrate species across the tree of life. The amino acid sequences were aligned with the *G.*
581 *platifrons* sequences annotated as E3 ubiquitin ligase (Bpl_scaf_25983-1.26 and Bpl_scaf_1894-
582 1.45) using MAFFT v7.450. The alignment was used to estimate a maximum-likelihood best
583 gene tree and to calculate bootstrap values on each node using RAxML v8.2.
584
585 ***G. platifrons* gill fixation and storage:** The gill tissues of *G. platifrons* were dissected within
586 minutes after the ROV, and samples were retrieved on board R/V Kexue. The gill tissues were
587 briefly washed with ice-cold filtered and autoclaved *in situ* seawater (FAISW) and then fixed in 4%
588 PFA prepared in FAISW at 4 °C overnight. The gill tissues were washed three times with ice-cold
589 PBST, dehydrated in 100% methanol and stored at -20 °C until use.
590
591 **Whole-mount *in situ* hybridisation:** For WISH and double FISH analyses, the DNA fragments
592 (~1000bp) of the targeted genes were first PCR amplified with gene-specific primers (GSPs)
593 pairs using *G. platifrons* gill cDNA as template (the sequences of targeted genes and gene-
594 specific primers were provided in Supplementary Data S2). The amplified fragments were
595 ligated into the pMD18-simpleT vector (Takara) and transformed into *E. coli*. Individual colonies
596 were picked up, and their plasmids were sequenced to confirm the inserts. The templates for in
597 vitro mRNA transcription were amplified using T7 forward GSP (sense probe control) or Sp6

598 reversed GSPs (antisense probe) combined with either forward or reversed gene-specific primer.
599 Labelled probes and control probes were generated using digoxigenin (DIG)-12-UTP (Roche) or
600 fluorescein-12-UTP (Roche) according to the protocol described by Thisse[103] with Sp6 and T7
601 RNA polymerase, respectively.
602 For WISH, the connective region at the end of the W-shaped gill filament was cut off, and each
603 gill slice was carefully peeled off with fine-tip tweezers. We dissected gill tissues from 5
604 individual mussels and pooled all the gill slices together. The gill slices were then rehydrated in
605 75%, 50% and 25% methanol-PBST (1×PBS with 0.1% Tween 20) for 15 min each, followed by
606 3×5 min PBST washes. The gill slices were then permeabilised with 2 µg/mL proteinase K in
607 PBST for 30 min at 37 °C. Post-digestion fixation was conducted by fixing the gill slices with 4%
608 PFA in PBST for 30 min at room temperature (RT). After 3×5 min PBST wash to remove the
609 residual fixative, the gill slices were pre-hybridised with a hybridisation mix (HM; containing 50%
610 formamide, 5×saline–sodium citrate (SSC), 0.1% Tween 20, 10 µg/mL heparin, 500 µg/mL yeast
611 tRNA) for 1 h at 65 °C. For each hybridisation, 5-10 gill slices were added to 400 µL of fresh HM
612 containing ~0.5 ng/µL DIG-labelled probe. Hybridisation was conducted in a 55 °C shaking water
613 bath overnight. Post-hybridisation washes were performed according to the following steps:
614 the gill slices were first washed 3×15 min with hybridisation washing buffer (50% formamide,
615 5×SSC, 0.1% Tween 20), followed by 3×15 min 2×SSC with 0.1% Tween 20 and 3×15 min 0.2×SSC
616 with 0.1% Tween 20. The washings were also conducted in a shaking water batch, and all the
617 washing buffers were pre-heated to 55 °C. The samples were washed three times with PBST
618 and then blocked in blocking buffer (2.5% sheep serum, 2% BSA in PBST) for 1 h at RT. Each
619 sample was incubated with 1:10,000 diluted anti-DIG-AP antibody (Roche) at 4 °C overnight.
620 The samples were incubated with the antibody, then washed for 6×15 min with PBST, followed
621 by washing in 3×15 min alkaline Tris buffer (100 mM Tris-HCl, pH 9.5; 100 mM NaCl; 50 mM
622 $MgCl_2$). The samples were incubated in nitro blue tetrazolium/5-Bromo-4-chloro-3-indolyl
623 phosphate staining solution (Sangon). After the desired expression pattern was revealed, the
624 staining reaction was stopped by 3×15 min PBST–EDTA wash (PBST, 1 mM ETDA). The gill slices
625 were cleared by incubating in 100% glycerol overnight at 4 °C and then mounted on glass slides.
626 The results of control hybridisations (with sense probes) were provided in Supplementary Fig.

627 S12-S15. WISH analyses were repeated with another batch of *G. platifrons* gill slices samples
628 collected during the *R/V* 'Kexue' 2017 South China Sea cold seep cruise to confirm the
629 consistency in expression patterns.
630
631 **Paraffin embedding and double fluorescent *in situ* hybridisation:** The methanal-dehydrated
632 gill slices were incubated in 100% ethanol, a 1:1 mixture of 100% ethanol and xylene and Xylene
633 twice for 1 h each at RT. The samples were embedded by incubating in Paraplast Plus (Sigma-
634 Aldrich) for 2 h at 65 °C and then cooled down to RT. Sections with 5 µM thickness were cut
635 using a microtome (Leica).
636 For double FISH, sections were dewaxed by incubating in xylene twice, a 1:1 mixture of 100%
637 ethanol and xylene, 100% ethanol twice, 95% ethanol, 85% ethanol and 75% ethanol for 15 min
638 each at RT. The sections were washed with PBST three times for 10 min each and then
639 permeabilised by 2 µg/mL proteinase K (NEB) in PBST for 15 min at RT. Post-digestion fixation
640 was conducted by incubating the sections in 4% PFA in PBST for 30 min at RT. The sections were
641 washed three times with PBST for 15 min each. Pre-hybridisation was conducted by incubating
642 the sections in HM for 1 h at 55 °C. Then, the *in situ* hybridisation was performed by incubating
643 the sections in ~0.5 ng/µL fluorescein (Roche)-labelled probed prepared in fresh HM overnight
644 at 55 °C. The sections were washed three times with 2×SSC for 15 min each at 55 °C, cooled
645 down to room temperature and washed three times with PBST.
646 A second-round hybridisation was conducted on the DIG-labelled oligonucleotides to label the
647 symbiont. The slices were hybridised for 1 h at 46 °C with 100 ng DIG-labelled *G. platifrons*
648 symbiont-specific probe in FISH buffer (0.9 M NaCl, 0.02 M Tris-HCl, 0.01% sodium dodecyl
649 sulphate [SDS] and 30% formamide). The slices were then washed with FISH washing buffer (0.1
650 M NaCl, 0.02 M Tris HCl, 0.01% SDS and 5 mM EDTA) three times at 5 min each at 48 °C. The
651 slices were washed with PBST three times and then blocked with blocking buffer (2.5% sheep
652 serum and 2% BSA in sheep serum) for 1 h at room temperature.
653 The slices were then incubated with 1:1000 diluted anti-fluorescein–peroxidase (POD) (Roche)
654 overnight at 4 °C, then washed six times with PBST for 15 min each and three times with TNT
655 buffer (100 mM Tris-HCl, pH 7.5; 100 mM NaCl; 0.1% Tween 20) for 15 min each. Afterwards,

656 the fluorescent signal of the *G. platifrons* gene expression pattern was developed by the TSA
657 fluorescein kit (Akoya Biosciences) according to the manufacturer's protocol. The slices were
658 washed three times, and the remaining POD activity was quenched by incubation in 1%
659 hydrogen peroxide solution for 1 h at room temperature. Then, the slices were washed three
660 times with PBS, blocked with blocking buffer for 30 min at room temperature, incubated with
661 1:2500 diluted anti-DIG–POD (Roche) for 2 h at RT and washed with PBST for six times and TNT
662 for three times. The symbiont FISH signal was developed using the TSA Cy3 kit (Akoya
663 Biosciences). Finally, the slices were washed with PBST, stained with DAPI, and mounted with
664 ProLong Diamond Antifade Mountant (Thermo Fisher).
665
666 **Microscopy imaging:** All the WISH samples and whole-mount 16S FISH images were observed
667 and imaged with a Nikon Eclipse Ni microscope with a DS-Ri2 camera. The double FISH slides
668 were imaged with a Zeiss LSM710 confocal microscope.
669
670 **Electron microscopy analysis:** The gill slices of the *G. platifrons* were dissected and fixed in
671 electron microscopy fixative (2.5% glutaraldehyde and 2% PFA) at 4 °C. For SEM analysis, the
672 samples were dehydrated in a graded ethanol series and then dried at the critical point. The
673 samples were then coated with gold (sputter/carbon Thread, EM ACE200) and observed under
674 a scanning electron microscope (VEGA3, Tescan). For TEM analysis, the samples were rinsed
675 with double distilled water, post-fixed with 1% osmium tetroxide and then washed with double
676 distilled water. The samples were then rinsed, dehydrated and embedded in Ep812 resin. Ultra-
677 thin sections were obtained with an ultramicrotome (70 nm thickness, Reichert-Jung Ultracut E).
678 The sections were then double-stained with lead citrate and uranyl acetate. The cells were
679 observed under a transmission electron microscope (JEM1200, Jeol) operated under 100 kV.

680 **References**

681 1 Kremer, N. *et al.* Initial symbiont contact orchestrates host-organ-wide transcriptional
682 changes that prime tissue colonization. *Cell Host & Microbe.* **14**, 183-194 (2013).
683 https://doi.org:10.1016/j.chom.2013.07.006
684 2 Bang, C. *et al.* Metaorganisms in extreme environments: do microbes play a role in
685 organismal adaptation? *Zoology (Jena)* (2018). https://doi.org:10.1016/j.zool.2018.02.004
686 3 Franke, M., Geier, B., Hammel, J. U., Dubilier, N. & Leisch, N. Coming together-symbiont
687 acquisition and early development in deep-sea bathymodioline mussels. *Proc. R. Soc.* **288**,
688 20211044 (2021). https://doi.org:10.1098/rspb.2021.1044
689 4 Dubilier, N., Bergin, C. & Lott, C. Symbiotic diversity in marine animals: the art of
690 harnessing chemosynthesis. *Nat. Rev. Microbiol.* **6**, 725 (2008).
691 https://doi.org:10.1038/nrmicro1992
692 5 Sogin, E. M., Kleiner, M., Borowski, C., Gruber-Vodicka, H. R. & Dubilier, N. Life in the dark:
693 phylogenetic and physiological diversity of chemosynthetic symbioses. *Annu. Rev.*
694 *Microbiol.* **75**, 695-718 (2021). https://doi.org:10.1146/annurev-micro-051021-123130
695 6 Xu, T., Feng, D., Tao, J. & Qiu, J.-W. A new species of deep-sea mussel (Bivalvia: Mytilidae:
696 *Gigantidas*) from the South China Sea: Morphology, phylogenetic position, and gill-
697 associated microbes. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **146**, 79-90 (2019).
698 https://doi.org:https://doi.org/10.1016/j.dsr.2019.03.001
699 7 DeChaine, E. G. & Cavanaugh, C. M. Symbioses of methanotrophs and deep-sea mussels
700 (Mytilidae: Bathymodiolinae). *Prog. Mol. Subcell. Biol.* **41**, 227-249 (2006).
701 https://doi.org:10.1007/3-540-28221-1_11
702 8 Fujiwara, Y. *et al.* Phylogenetic characterization of endosymbionts in three hydrothermal
703 vent mussels influence on host distributions. *Mar. Ecol. Prog. Ser.* **208**, 147-155 (2000).
704 https://doi:10.3354/meps208147
705 9 Kiel, S. The vent and seep biota : aspects from microbes to ecosystems. (Springer, 2010).
706 https://doi.org/10.1007/978-90-481-9572-5
707 10 Vrijenhoek, R. C. Genetic diversity and connectivity of deep-sea hydrothermal vent
708 metapopulations. *Mol. Ecol.* **19**, 4391-4411 (2010). https://doi.org:10.1111/j.1365-
709 294X.2010.04789.x
710 11 Halary, S., Riou, V., Gaill, F., Boudier, T. & Duperron, S. 3D FISH for the quantification of
711 methane- and sulphur-oxidizing endosymbionts in bacteriocytes of the hydrothermal vent
712 mussel *Bathymodiolus azoricus*. *ISME J* **2**, 284-292 (2008).
713 https://doi.org:10.1038/ismej.2008.3
714 12 Zheng, P. *et al.* Insights into deep-sea adaptations and host-symbiont interactions: A
715 comparative transcriptome study on Bathymodiolus mussels and their coastal relatives.
716 *Mol. Ecol.* **26**, 5133-5148 (2017). https://doi.org:10.1111/mec.14160
717 13 Fiala-Médioni, A., Métivier, C., Herry, A. & Le Pennec, M. *Ultrastructure of the gill of the*
718 *hydrothermal-vent mytilid Bathymodiolus sp.*. *Mar. Biol.* **92**, 65–72 (1986).
719 https://doi.org/10.1007/BF00392747
720 14 Wong, Y. H. *et al.* High-throughput transcriptome sequencing of the cold seep mussel
721 Bathymodiolus platifrons. *Sci. Rep.* **5**, 16597 (2015). https://doi.org:10.1038/srep16597

722 15 Sun, J. *et al.* Adaptation to deep-sea chemosynthetic environments as revealed by mussel
723 genomes. *Nat. Ecol. Evol.* **1**, 121 (2017). https://doi.org:10.1038/s41559-017-0121
724 16 Bettencourt, R. *et al.* An Insightful Model to Study Innate Immunity and Stress Response
725 in Deep‐Sea Vent Animals: Profiling the Mussel *Bathymodiolus azoricus*. *Organismal and*
726 *Molecular Malacology*, **8**, (2017). https://doi.org:10.5772/68034.
727 17 Barros, I. *et al.* Post-capture immune gene expression studies in the deep-sea
728 hydrothermal vent mussel *Bathymodiolus azoricus* acclimatized to atmospheric pressure.
729 *Fish Shellfish Immunol* **42**, 159-170 (2015). https://doi.org:10.1016/j.fsi.2014.10.018
730 18 Wang, H. *et al.* Molecular analyses of the gill symbiosis of the bathymodiolin mussel
731 *Gigantidas platifrons*. *iScience* **24**, 101894 (2021).
732 https://doi.org:10.1016/j.isci.2020.101894
733 19 Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging. Spatially
734 resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, 6090 (2015).
735 https://doi.org:10.1126/science.aaa6090
736 20 Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and
737 bioinformatics pipelines. *Exp. Mol. Med.* **50**, 1-14 (2018). https://doi.org:10.1038/s12276-
738 018-0071-8
739 21 Saliba, A. E., Westermann, A. J., Gorski, S. A. & Vogel, J. Single-cell RNA-seq: advances and
740 future challenges. *Nucleic Acids Res* **42**, 8845-8860 (2014).
741 https://doi.org:10.1093/nar/gku555
742 22 Chen, X., Teichmann, S. A. & Meyer, K. B. From tissues to cell types and back: single-cell
743 gene expression analysis of tissue architecture. *Annu. Rev. Biomed. Data Sci.* **1**, 29-51
744 (2018). https://doi.org:10.1146/annurev-biodatasci-080917-013452
745 23 Wu, H., Kirita, Y., Donnelly, E. L. & Humphreys, B. D. Advantages of single-nucleus over
746 single-cell RNA sequencing of adult kidney: rare cell types and novel cell states revealed in
747 fibrosis. *J. Am. Soc. Nephrol.* **30**, 23-32 (2019). https://doi.org:10.1681/ASN.2018090912
748 24 Feng, D. *et al.* Using Bathymodiolus tissue stable carbon, nitrogen and sulfur isotopes to
749 infer biogeochemical process at a cold seep in the South China Sea. *Deep Sea Res. Part I*
750 *Oceanogr. Res.* **104**, 52-59 (2015).
751 https://doi.org:https://doi.org/10.1016/j.dsr.2015.06.011
752 25 Kharchenko, P. V., Silberstein, L. & Scadden, D. T. Bayesian approach to single-cell
753 differential expression analysis. *Nat. Methods.* **11**, 740-742 (2014).
754 https://doi.org:10.1038/nmeth.2967
755 26 Elyanow, R., Dumitrascu, B., Engelhardt, B. E. & Raphael, B. J. netNMF-sc: leveraging gene-
756 gene interactions for imputation and dimensionality reduction in single-cell expression
757 analysis. *Genome. Res.* **30**, 195-204 (2020). https://doi.org:10.1101/gr.251603.119
758 27 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587
759 (2021). https://doi.org:10.1016/j.cell.2021.04.048
760 28 Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902 (2019).
761 https://doi.org:10.1016/j.cell.2019.05.031
762 29 Petibon, C., Malik Ghulam, M., Catala, M. & Abou Elela, S. Regulation of ribosomal protein
763 genes: An ordered anarchy. *WIREs RNA* **12**, e1632 (2021).
764 https://doi.org:https://doi.org/10.1002/wrna.1632

30 Welcker, D. *et al.* Hemicentin-1 is an essential extracellular matrix component of the dermal–epidermal and myotendinous junctions. *Sci. Rep.* **11**, 17926 (2021). https://doi.org:10.1038/s41598-021-96824-4

31 Dufour, S. C. & Beninger, P. G. A functional interpretation of cilia and mucocyte distributions on the abfrontal surface of bivalve gills. *Mar. Biol.* **138**, 295-309 (2001). https://doi.org:10.1007/s002270000466

32 Dufour, S. C. Gill anatomy and the evolution of symbiosis in the bivalve family Thyasiridae. *Biol. Bull.* **208**, 200-212 (2005). https://doi.org:10.2307/3593152

33 Gómez-Mendikute, A., Elizondo, M., Venier, P. & Cajaraville, M. P. Characterization of mussel gill cells in vivo and in vitro. *Cell Tissue Res.* **321**, 131-140 (2005). https://doi.org:10.1007/s00441-005-1093-9

34 Beninger, P. G. & Dufour, S. Mucocyte distribution and relationship to particle transport on the pseudolamellibranch gill of *Crassostrea virginica* (Bivalvia:Ostreidae). *Mar. Ecol. Progr. Ser.* **137**, 133-138 (1996). https://doi.org:10.3354/meps137133

35 Gerdol, M. *et al.* The C1q domain containing proteins of the Mediterranean mussel Mytilus galloprovincialis: A widespread and diverse family of immune-related molecules. *Dev. Comp. Immunol.* **35**, 635-643 (2011). https://doi.org:https://doi.org/10.1016/j.dci.2011.01.018

36 Wang, W., Song, X., Wang, L. & Song, L. Pathogen-derived carbohydrate recognition in molluscs immune defense. *Int. J. Mol. Sci.* **19**, 721 (2018). https://doi.org:10.3390/ijms19030721

37 Saco, A., Rey-Campos, M., Novoa, B. & Figueras, A. Transcriptomic response of mussel gills after a *Vibrio splendidus* infection demonstrates their role in the immune response. *Front. Immunol.* **11**, 615580 (2020). https://doi.org:10.3389/fimmu.2020.615580

38 Richoux, N. B. & Thompson, R. J. Regulation of particle transport within the ventral groove of the mussel (*Mytilus edulis*) gill in response to environmental conditions. *J. Exp. Mar. Biol. Ecol.* **260**, 199-215 (2001). https://doi.org:10.1016/s0022-0981(01)00254-4

39 Riisgård, H. U., Egede, P. P. & Barreiro Saavedra, I. Feeding behaviour of the mussel, *Mytilus edulis*: New observations, with a minireview of current knowledge. *J. Mar. Biol.* **2011**, 312459 (2011). https://doi.org:10.1155/2011/312459

40 Heude, É., Shaikho, S. & Ekker, M. The dlx5a/dlx6a genes play essential roles in the early development of zebrafish median fin and pectoral structures. *PLOS ONE* **9**, e98505 (2014). https://doi.org:10.1371/journal.pone.0098505

41 Han, S. *et al.* TULP3 is required for localization of membrane-associated proteins ARL13B and INPP5E to primary cilia. *Biochem. Biophys. Res. Commun.* **509**, 227-234 (2019). https://doi.org/10.1016/j.bbrc.2018.12.109

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

43 Das, A., Dickinson, D. J., Wood, C. C., Goldstein, B. & Slep, K. C. Crescerin uses a TOG domain array to regulate microtubules in the primary cilium. *Mol. Biol. Cell.* **26**, 4248-4264 (2015). https://doi.org:10.1091/mbc.E15-08-0603

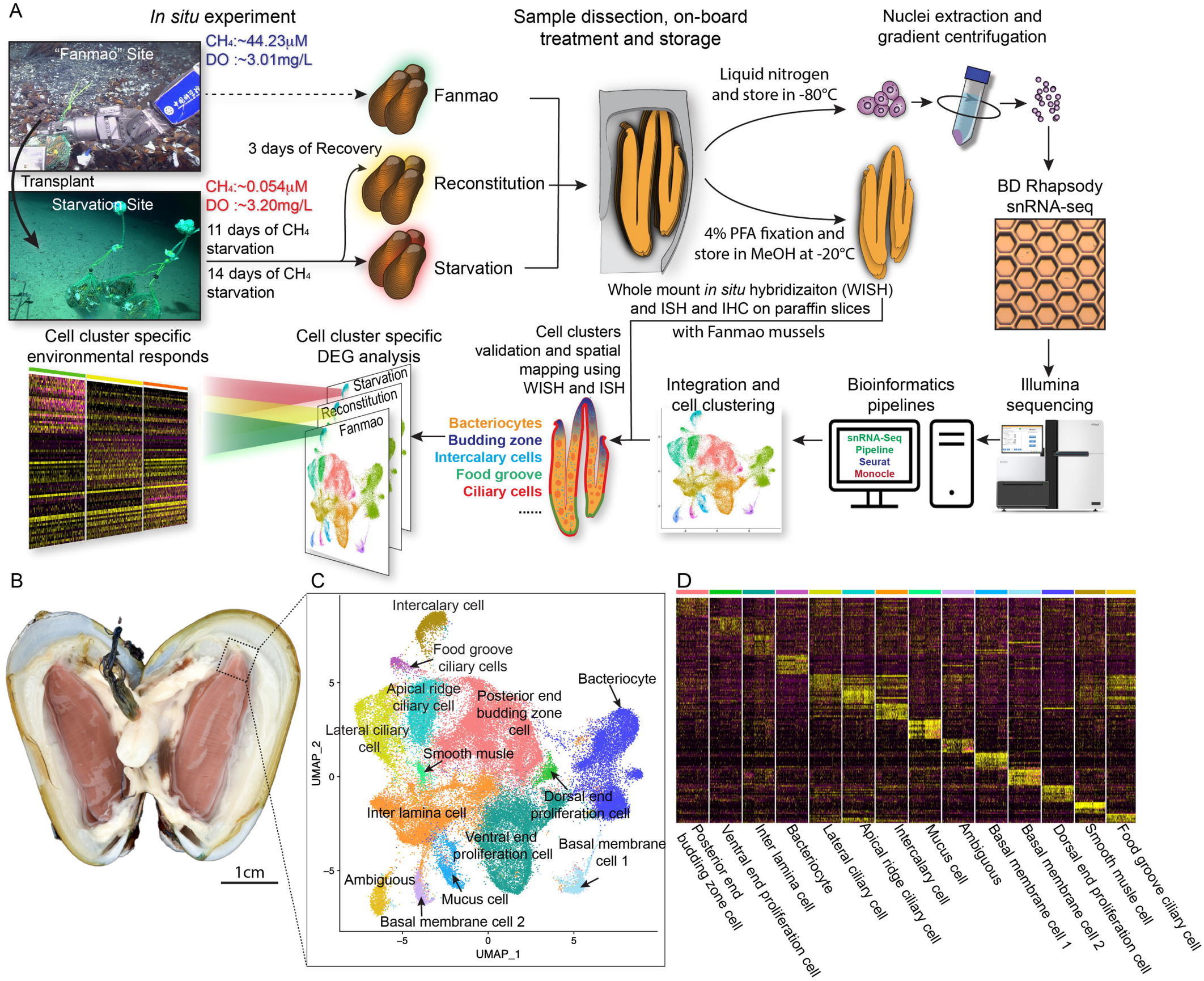
44 Llorente-Cortes, V., Martinez-Gonzalez, J. & Badimon, L. LDL receptor-related protein mediates uptake of aggregated LDL in human vascular smooth muscle cells. *Arterioscler. Thromb. Vasc. Biol.* **20**, 1572-1579 (2000). https://doi.org:10.1161/01.atv.20.6.1572

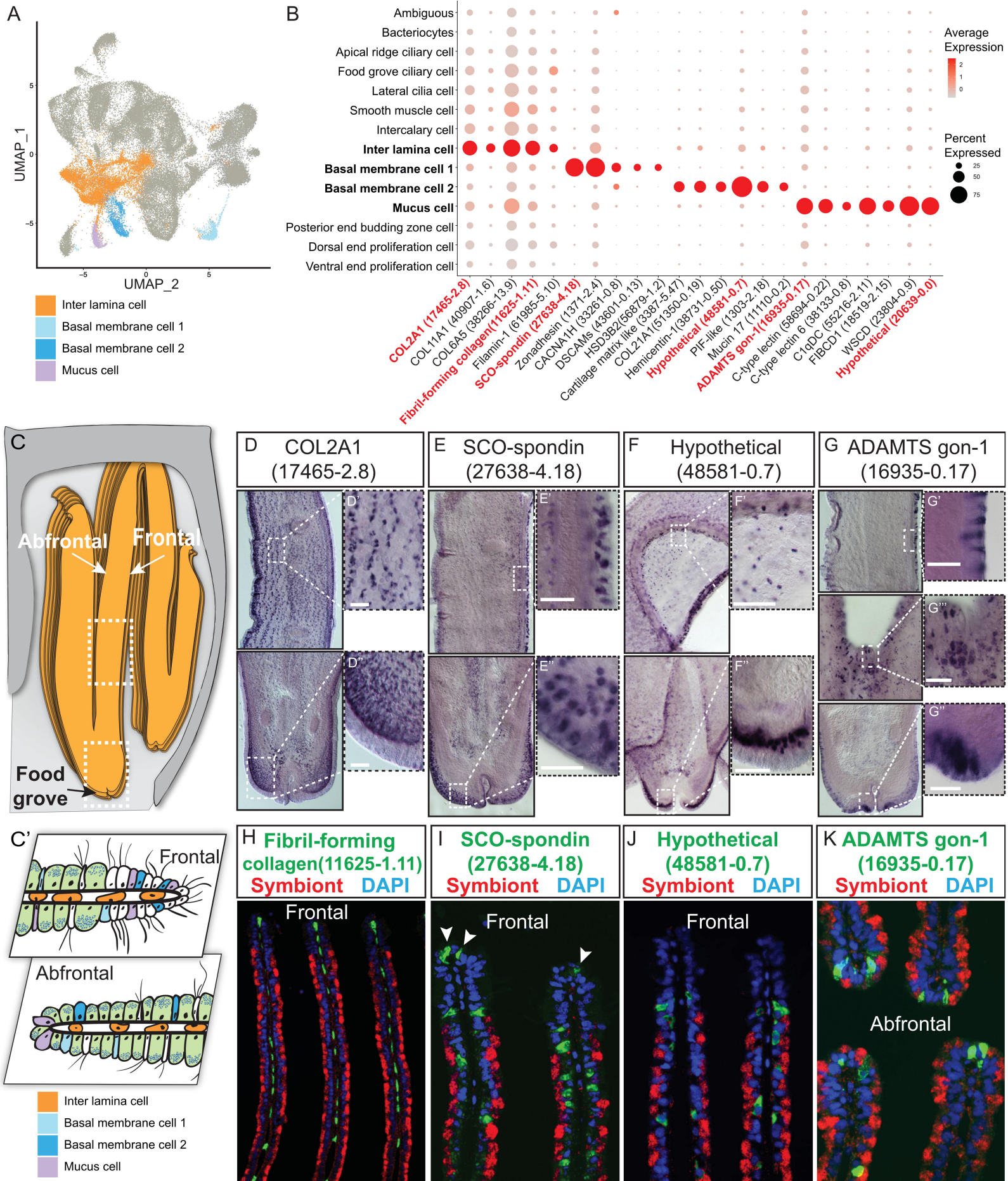
809 45 Chen, X. *et al.* Angiotensin-converting enzyme in smooth muscle cells promotes
810 atherosclerosis-brief report. *Arterioscler. Thromb. Vasc. Biol.* **36**, 1085-1089 (2016).
811 https://doi.org:10.1161/ATVBAHA.115.307038
812 46 Linke, W. A. & Grützner, A. Pulling single molecules of titin by AFM—recent advances and
813 physiological implications. *Pflug. Arch. Eur.* **456**, 101-115 (2008).
814 https://doi.org:10.1007/s00424-007-0389-x
815 47 St Paul, A., Corbett, C. B., Okune, R. & Autieri, M. V. Angiotensin II, Hypercholesterolemia,
816 and vascular smooth muscle cells: A perfect trio for vascular pathology. *Int. J. Mol. Sci.* **21**
817 4525 (2020). https://doi.org:10.3390/ijms21124525
818 48 Ytrehus, K., Ludvigsen, S., Mancusi, C., Gerdts, E. & de Simone, G. Heart angiotensin-
819 converting enzyme and angiotensin-converting enzyme 2 gene expression associated with
820 male sex and salt-sensitive hypertension in the Dahl Rat. *Front. physiol.* **12** (2021).
821 https://doi.org:10.3389/fphys.2021.663819
822 49 Keller, T. C., 3rd *et al.* Role of titin in nonmuscle and smooth muscle cells. *Adv. Exp. Med.*
823 *Biol.* **481**, 265-277 (2000). https://doi.org:10.1007/978-1-4615-4267-4_16
824 50 Chen, J. V. *et al.* Rootletin organizes the ciliary rootlet to achieve neuron sensory function
825 in *Drosophila*. *J. Cell. Biol.* **211**, 435-453 (2015). https://doi.org:10.1083/jcb.201502032
826 51 Styczynska-Soczka, K. & Jarman, A. P. The *Drosophila* homologue of rootletin is required
827 for mechanosensory function and ciliary rootlet formation in chordotonal sensory
828 neurons. *Cilia* **4**, 9 (2015). https://doi.org:10.1186/s13630-015-0018-9
829 52 Mohan, S., Timbers, T. A., Kennedy, J., Blacque, O. E. & Leroux, M. R. Striated rootlet and
830 nonfilamentous forms of rootletin maintain ciliary function. *Curr. Biol.* **23**, 2016-2022
831 (2013). https://doi.org:10.1016/j.cub.2013.08.033
832 53 Jones, S. E. & Jomary, C. Secreted Frizzled-related proteins: searching for relationships
833 and patterns. *Bioessays* **24**, 811-820 (2002). https://doi.org:10.1002/bies.10136
834 54 Gillis, W. Q., Bowerman, B. A. & Schneider, S. Q. The evolution of protostome GATA
835 factors: molecular phylogenetics, synteny, and intron/exon structure reveal orthologous
836 relationships. *BMC Evol. Biol.* **8**, 112-112 (2008). https://doi.org:10.1186/1471-2148-8-
837 112
838 55 Zhang, Z., Liu, L., Twumasi-Boateng, K., Block, D. H. S. & Shapira, M. FOS-1 functions as a
839 transcriptional activator downstream of the *C. elegans* JNK homolog KGB-1. *Cell. Signal.*
840 **30**, 1-8 (2017). https://doi.org:https://doi.org/10.1016/j.cellsig.2016.11.010
841 56 Phochanukul, N. & Russell, S. No backbone but lots of Sox: Invertebrate Sox genes. *Int. J.*
842 *Biochem. Cell Biol.* **42**, 453-464 (2010).
843 https://doi.org:https://doi.org/10.1016/j.biocel.2009.06.013
844 57 Cannuel, R., Beninger, P. G., McCombie, H. & Boudry, P. Gill Development and its
845 functional and evolutionary implications in the blue mussel *Mytilus edulis* (Bivalvia:
846 Mytilidae). *Biol. Bull.* **217**, 173-188 (2009). https://doi.org:10.1086/BBLv217n2p173
847 58 Leibson, N. L. & Movchan, O. T. Cambial zones in gills of Bivalvia. *Mar. Biol.* **31**, 175-180
848 (1975). https://doi.org:10.1007/BF00391629
849 59 Wentrup, C., Wendeberg, A., Schimak, M., Borowski, C. & Dubilier, N. Forever competent:
850 deep-sea bivalves are colonized by their chemosynthetic symbionts throughout their
851 lifetime. *Environ. Microbiol.* **16**, 3699-3713 (2014). https://doi.org:10.1111/1462-
852 2920.12597

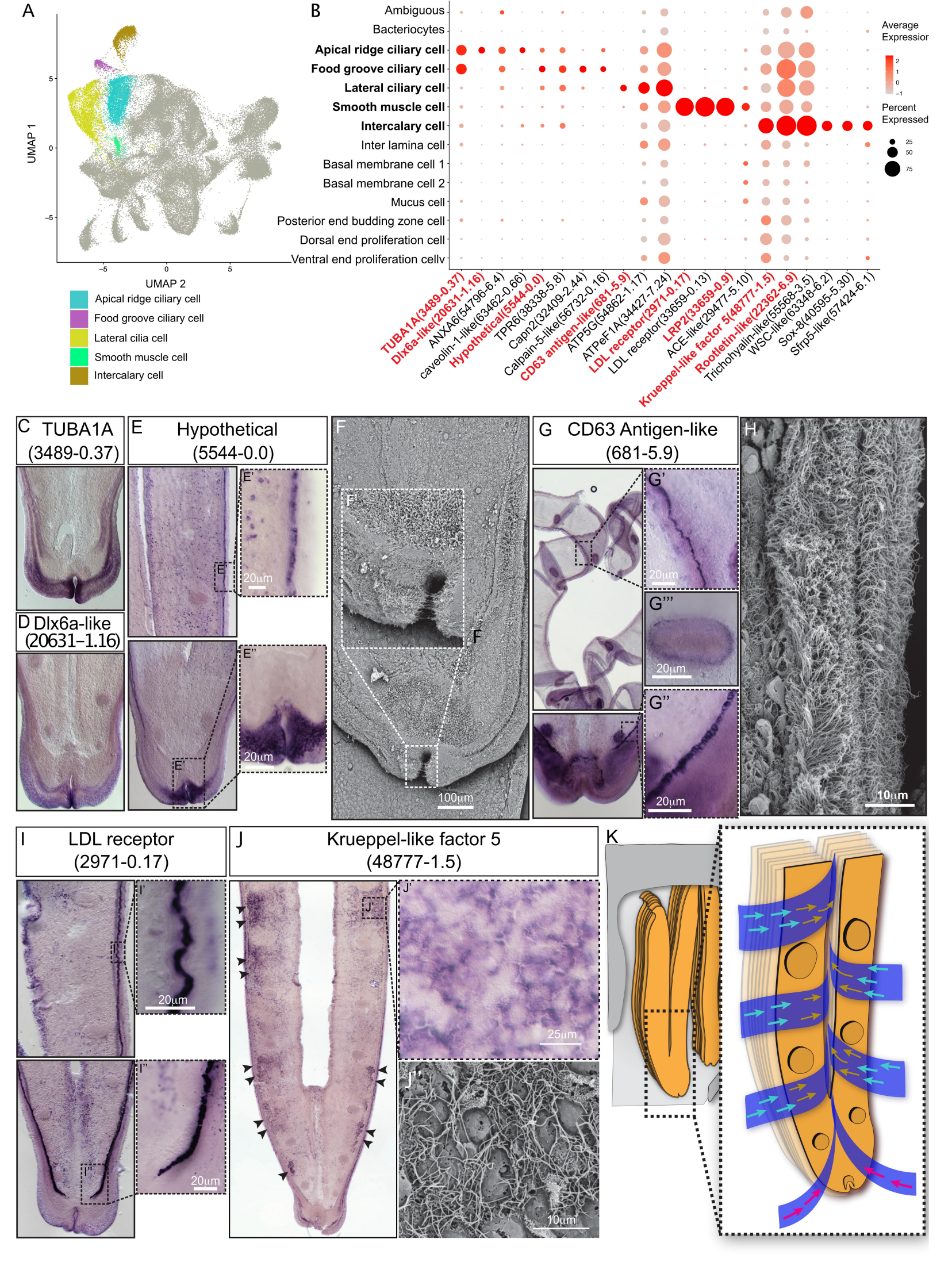
853 60 Mohieldin, A. M. *et al.* Proteomic identification reveals the role of ciliary extracellular-like
854 vesicle in cardiovascular function. *Adv. Sci.* **7**, 1903140 (2020).
855 https://doi.org:https://doi.org/10.1002/advs.201903140
856 61 Neumann, D. & Kappes, H. On the growth of bivalve gills initiated from a lobule-producing
857 budding zone. *Biol. Bull.* **205**, 73-82 (2003). https://doi.org:10.2307/1543447
858 62 Streams, M. E., Fisher, C. R. & Fiala-Médioni, A. Methanotrophic symbiont location and
859 fate of carbon incorporated from methane in a hydrocarbon seep mussel. *Mar. Biol.* **129**,
860 465-476 (1997). https://doi.org:10.1007/s002270050187
861 63 Ponnudurai, R. *et al.* Metabolic and physiological interdependencies in the *Bathymodiolus*
862 *azoricus* symbiosis. *ISME J* **11**, 463-477 (2017). https://doi.org:10.1038/ismej.2016.124
863 64 Barry, J. P. *et al.* Methane‐based symbiosis in a mussel, *Bathymodiolus platifrons*, from
864 cold seeps in Sagami Bay, Japan. *Invertebr. Biol.* **121**, 47-54 (2002).
865 https://doi.org:doi:10.1111/j.1744-7410.2002.tb00128.x
866 65 Renoz, F., Noel, C., Errachid, A., Foray, V. & Hance, T. Infection dynamic of symbiotic
867 bacteria in the pea *Aphid acyrthosiphon pisum* gut and host immune response at the early
868 steps in the infection process. *Plos One* **10,** 0122099 (2015).
869 https://doi.org:ARTNe012209910.1371/journal.pone.
870 66 Guan, H. *et al.* Lipid Biomarker Patterns Reflect Nutritional Strategies of Seep-Dwelling
871 Bivalves From the South China Sea. *Front. Mar. Sci.* **9,** 3398 (2022).
872 https://doi.org:10.3389/fmars.2022.831286
873 67 Li, Y. *et al.* Genomic adaptations to chemosymbiosis in the deep-sea seep-dwelling
874 tubeworm *Lamellibrachia luymesi*. *BMC Biol.* **17**, 91 (2019).
875 https://doi.org:10.1186/s12915-019-0713-x
876 68 Sun, Y. *et al.* Genomic signatures supporting the symbiosis and formation of chitinous
877 tube in the deep-sea tubeworm *Paraescarpia echinospica. Mol. Biol. Evol.* **38**, 4116-4134
878 (2021). https://doi.org:10.1093/molbev/msab203
879 69 Takishita, K. *et al.* Genomic evidence that methanotrophic endosymbionts likely provide
880 deep-sea bathymodiolus mussels with a sterol intermediate in cholesterol biosynthesis.
881 *Genome Biol. Evol.* **9**, 1148-1160 (2017). https://doi.org:10.1093/gbe/evx082
882 70 Demidenko, A., Akberdin, I. R., Allemann, M., Allen, E. E. & Kalyuzhnaya, M. G. Fatty acid
883 biosynthesis pathways in *Methylomicrobium buryatense* 5G(B1). *Front. Microbiol.* **7**, 2167
884 (2016). https://doi.org:10.3389/fmicb.2016.02167
885 71 Brasaemle, D. L. *et al.* Adipose differentiation-related protein is an ubiquitously expressed
886 lipid storage droplet-associated protein. *J. Lipid Res.* **38**, 2249-2263 (1997).
887 https://doi.org/10.1016/S0022-2275(20)34939-7
888 72 Cheng, D. *et al.* Expression, purification, and characterization of human and rat acetyl
889 coenzyme A carboxylase (ACC) isozymes. *Protein Expr. Purif.* **51**, 11-21 (2007).
890 https://doi.org:10.1016/j.pep.2006.06.005
891 73 Monroig, Ó. & Kabeya, N. Desaturases and elongases involved in polyunsaturated fatty
892 acid biosynthesis in aquatic invertebrates: a comprehensive review. *Fish. Sci.* **84**, 911-928
893 (2018). https://doi.org:10.1007/s12562-018-1254-x
894 74 Kabeya, N. et al. Unique fatty acid desaturase capacities uncovered in Hediste diversicolor
895 illustrate the roles of aquatic invertebrates in trophic upgrading. *Philos. Trans. R. Soc.* **375**,
896 20190654 (2020). https://doi.org:10.1098/rstb.2019.0654

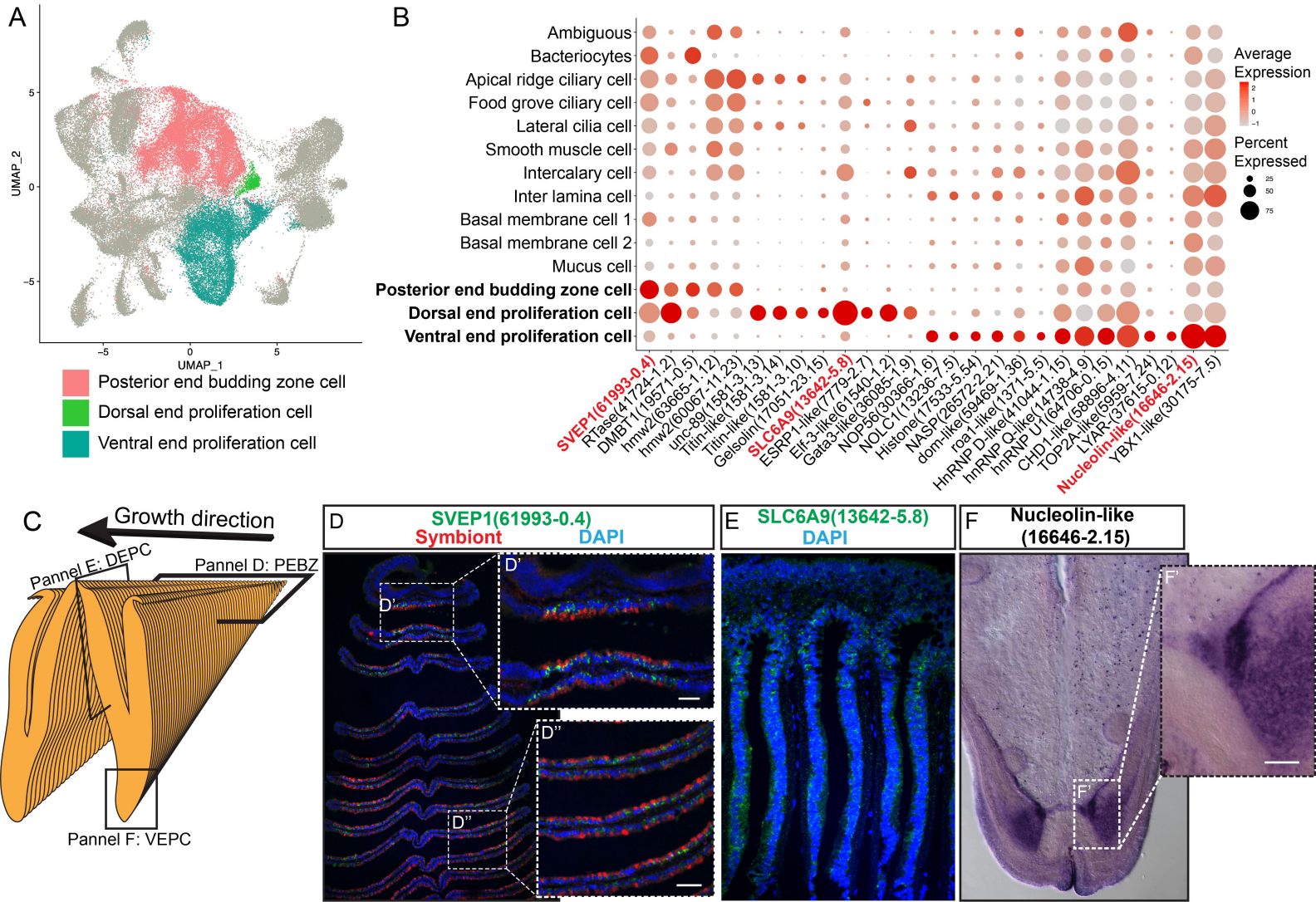
897 75 Soupene, E. & Kuypers, F. A. Mammalian long-chain acyl-CoA synthetases. *Exp. Biol. Med.*
898 *(Maywood).* **233**, 507-521 (2008). https://doi.org:10.3181/0710-MR-287
899 76 Alves-Bezerra, M. *et al.* Long-chain acyl-CoA synthetase 2 knockdown leads to decreased
900 fatty acid oxidation in fat body and reduced reproductive capacity in the insect *Rhodnius*
901 *prolixus*. *Biochim. Biophys. Acta.* **1861**, 650-662 (2016).
902 https://doi.org:10.1016/j.bbalip.2016.04.007
903 77 Hoglund, P. J., Nordstrom, K. J. V., Schioth, H. B. & Fredriksson, R. The solute carrier
904 families have a remarkably long evolutionary history with the majority of the human
905 families present before divergence of bilaterian species. *Mol. Biol. Evol.* **28**, 1531-1541
906 (2011). https://doi.org:10.1093/molbev/msq350
907 78 Mohamed, A. R. *et al.* Dual RNA-sequencing analyses of a coral and its native symbiont
908 during the establishment of symbiosis. *Mol. Ecol.* **29**, 3921-3937 (2020).
909 https://doi.org:10.1111/mec.15612
910 79 Bertucci, A., Foret, S., Ball, E. E. & Miller, D. J. Transcriptomic differences between day and
911 night in *Acropora millepora* provide new insights into metabolite exchange and light-
912 enhanced calcification in corals. *Mol. Ecol.* **24**, 4489-4504 (2015).
913 https://doi.org:10.1111/mec.13328
914 80 Hamada, M. *et al.* Metabolic co-dependence drives the evolutionarily ancient Hydra-
915 Chlorella symbiosis. *Elife* **7,** e35122 (2018). https://doi.org:10.7554/eLife.35122
916 81 Duncan, R. P., Feng, H., Nguyen, D. M. & Wilson, A. C. Gene family expansions in aphids
917 maintained by endosymbiotic and nonsymbiotic traits. *Genome Biol. Evol.* **8**, 753-764
918 (2016). https://doi.org:10.1093/gbe/evw020
919 82 Feng, H. L. *et al.* Trading amino acids at the aphid-Buchnera symbiotic interface. *P. Natl.*
920 *Acad. Sci. USA.* **116**, 16003-16011 (2019). https://doi.org:10.1073/pnas.1906223116
921 83 Ip, J. C. *et al.* Host-endosymbiont genome integration in a deep-sea chemosymbiotic clam.
922 *Mol. Biol. Evol.* **38**, 502-518 (2021). https://doi.org:10.1093/molbev/msaa241
923 84 Hongo, Y. *et al.* Expression of genes involved in the uptake of inorganic carbon in the gill
924 of a deep-sea vesicomyid clam harboring intracellular thioautotrophic bacteria. *Gene* **585**,
925 228-240 (2016). https://doi.org:https://doi.org/10.1016/j.gene.2016.03.033
926 85 Sotiriou, S. *et al.* Ascorbic-acid transporter Slc23a1 is essential for vitamin C transport into
927 the brain and for perinatal survival. *Nat. Med.* **8**, 514-517 (2002).
928 https://doi.org:10.1038/0502-514
929 86 Aydemir, T. B. et al. Zinc transporter ZIP14 functions in hepatic zinc, iron and glucose
930 homeostasis during the innate immune response (endotoxemia). *PLoS One* **7**, e48679
931 (2012). https://doi.org:10.1371/journal.pone.0048679
932 87 Turi, Z., Lacey, M., Mistrik, M. & Moudry, P. Impaired ribosome biogenesis: mechanisms
933 and relevance to cancer and aging. Aging (Albany NY) **11**, 2512-2540 (2019).
934 https://doi.org:10.18632/aging.101922
935 88 Hay, B. A. Understanding IAP function and regulation: a view from Drosophila. Cell. Death.
936 Differ. **7**, 1045-1056 (2000). https://doi.org:10.1038/sj.cdd.4400765
937 89 Dubrez-Daloz, L., Dupoux, A. & Cartier, J. IAPS : More than just inhibitors of apoptosis
938 proteins. Cell Cycle **7**, 1036-1046 (2008). https://doi.org:10.4161/cc.7.8.5783
939 90 Pflug, K. M. & Sitcheran, R. Targeting NF-κB-inducing kinase (NIK) in immunity,
940 inflammation, and cancer. *Int. J. Mol. Sci.* **21**, 8470 (2020).

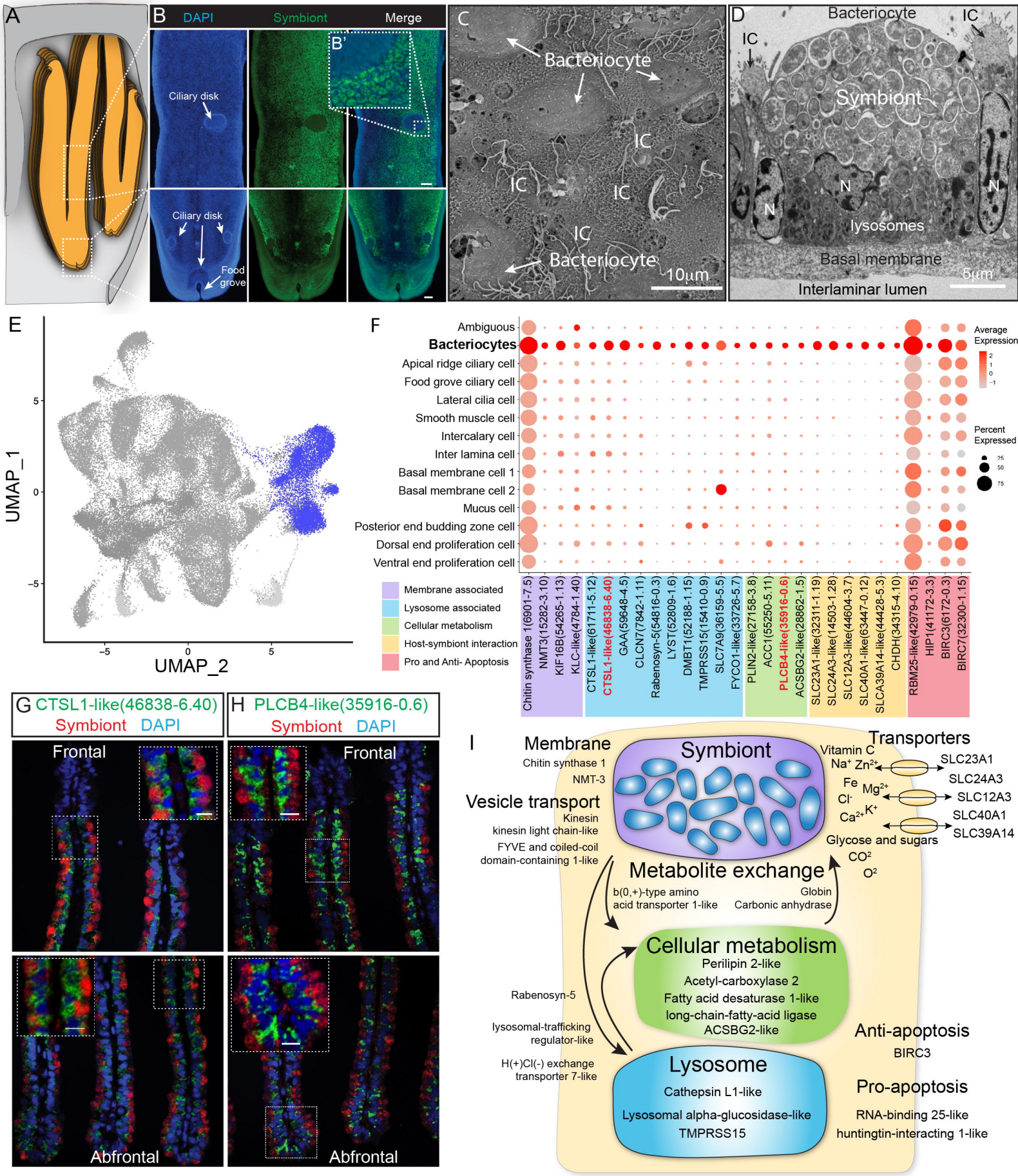
941 91 Otten, E. G. *et al.* Ubiquitylation of lipopolysaccharide by RNF213 during bacterial
942 infection. *Nature* **594**, 111-116 (2021). https://doi.org:10.1038/s41586-021-03566-4
943 92 Tashian, R. E. The carbonic anhydrases: widening perspectives on their evolution,
944 expression and function. *Bioessays* **10**, 186-192 (1989).
945 https://doi.org:10.1002/bies.950100603
946 93 Price, D. R. *et al.* Aphid amino acid transporter regulates glutamine supply to intracellular
947 bacterial symbionts. *Proc. Natl. Acad. Sci. USA.* **111**, 320-325 (2014).
948 https://doi.org:10.1073/pnas.1306068111
949 94 Ponnudurai, R. *et al.* Comparative proteomics of related symbiotic mussel species reveals
950 high variability of host–symbiont interactions. *ISME J.* **14**, 649-656 (2020).
951 https://doi.org:10.1038/s41396-019-0517-6
952 95 Hicks, S. C., Townes, F. W., Teng, M. & Irizarry, R. A. Missing data and technical variability
953 in single-cell RNA-sequencing experiments. *Biostatistics* **19**, 562-578 (2017).
954 https://doi.org:10.1093/biostatistics/kxx053
955 96 Lun, A. T. L. *et al.* EmptyDrops: distinguishing cells from empty droplets in droplet-based
956 single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
957 https://doi.org:10.1186/s13059-019-1662-y
958 97 McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet detection in single-
959 cell RNA sequencing data using artificial nearest neighbors. *Cell Systems* **8**, 329-337
960 (2019). https://doi.org:https://doi.org/10.1016/j.cels.2019.03.003
961 98 Stuart, T. et al. Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821
962 (2019). https://doi.org:https://doi.org/10.1016/j.cell.2019.05.031
963 99 Singh, P. & Zhai, Y. Deciphering Hematopoiesis at single cell level through the lens of
964 reduced dimensions. *bioRxiv*, 2022.2006.2007.495099 (2022).
965 https://doi.org:10.1101/2022.06.07.495099
966 100 Wu, T. et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.
967 *The Innovation* **2**, 100141 (2021).
968 https://doi.org:https://doi.org/10.1016/j.xinn.2021.100141
969 101 Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell
970 transcriptomics. *BMC Genom.* **19**, 477 (2018). https://doi.org:10.1186/s12864-018-4772-0
971 102 Moon, K. R. et al. Visualizing structure and transitions in high-dimensional biological data.
972 *Nat. Biotechnol.* **37**, 1482-1492 (2019). https://doi.org:10.1038/s41587-019-0336-3
973 103 Thisse, C. & Thisse, B. High-resolution *in situ* hybridization to whole-mount zebrafish
974 embryos. *Nat. Protoc.* **3**, 59-69 (2008). https://doi.org:10.1038/nprot.2007.514

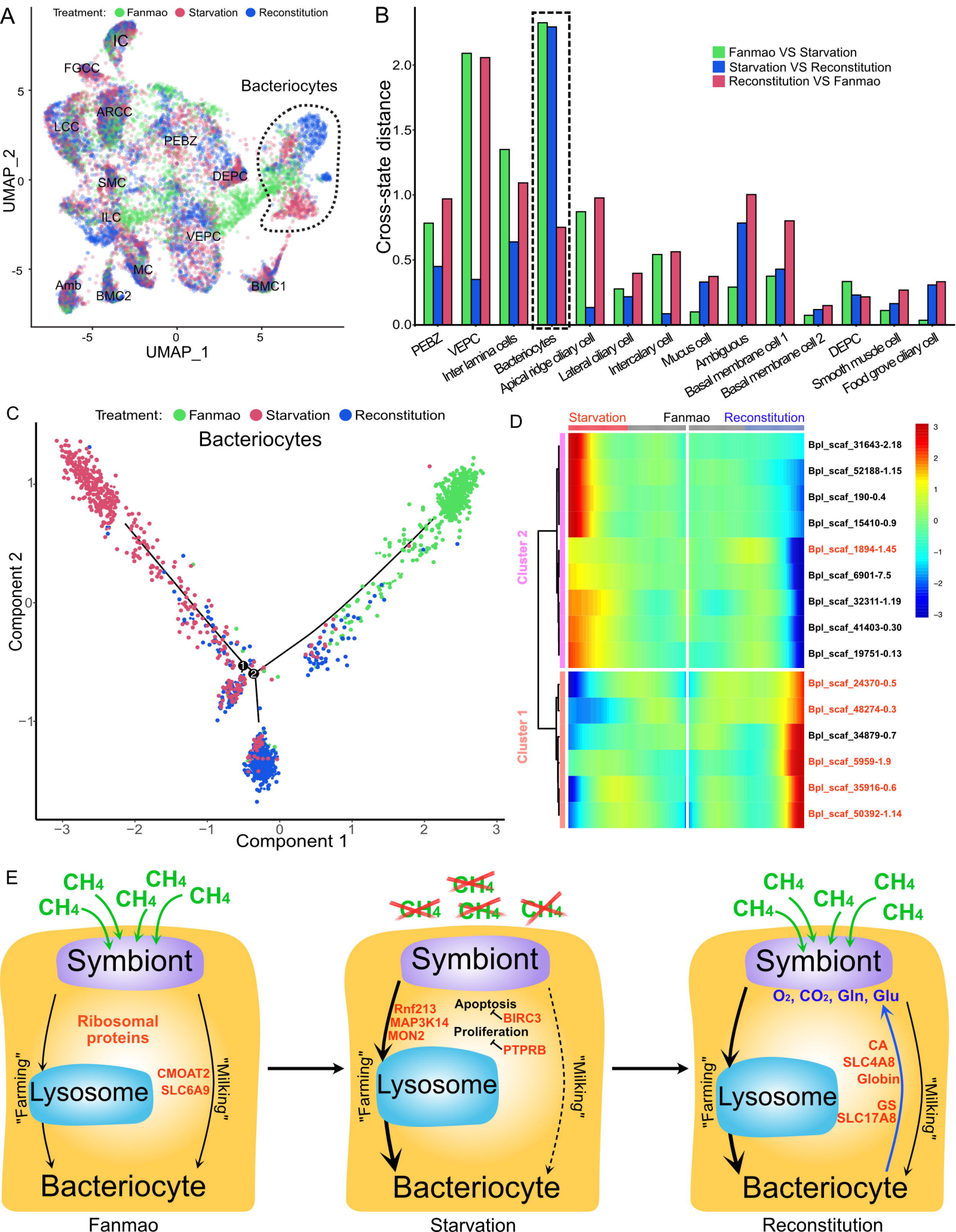

A
In situ experiment
"Fanmao" Site
CH4:~44.23μM
DO :~3.01mg/L
Fanmao
3 days of Recovery
Transplant
Starvation Site
CH4:~0.054μM
DO :~3.20mg/L
11 days of CH4 starvation
14 days of CH4 starvation
Reconstitution
Starvation
Sample dissection, on-board treatment and storage
Liquid nitrogen and store in -80°C
4% PFA fixation and store in MeOH at -20°C
Nuclei extraction and gradient centrifugation
BD Rhapsody snRNA-seq
Illumina sequencing
Bioinformatics pipelines
snRNA-Seq Pipeline
Seurat
Monocle
Integration and cell clustering
Whole mount in situ hybridizaiton (WISH) and ISH and IHC on paraffin slices with Fanmao mussels
Cell clusters validation and spatial mapping using WISH and ISH
Bacteriocytes
Budding zone
Intercalary cells
Food groove
Ciliary cells
.....
Cell cluster specific DEG analysis
Starvation
Reconstitution
Fanmao
Cell cluster specific environmental responds
B
1cm
C
Intercalary cell
Food groove ciliary cells
Apical ridge ciliary cell
Posterior end budding zone cell
Bacteriocyte
Lateral ciliary cell
Smooth musle
Dorsal end proliferation cell
Inter lamina cell
Ventral end proliferation cell
Basal membrane cell 1
Ambiguous
Mucus cell
Basal membrane cell 2
UMAP_2
UMAP_1
D
Posterior end budding zone cell
Ventral end proliferation cell
Inter lamina cell
Bacteriocyte
Lateral ciliary cell
Apical ridge ciliary cell
Intercalary cell
Mucus cell
Ambiguous
Basal membrane cell 1
Basal membrane cell 2
Dorsal end proliferation cell
Smooth musle cell
Food groove ciliary cell


A
UMAP_1
UMAP_2
Inter lamina cell
Basal membrane cell 1
Basal membrane cell 2
Mucus cell
B
Ambiguous
Bacteriocytes
Apical ridge ciliary cell
Food grove ciliary cell
Lateral cilia cell
Smooth muscle cell
Intercalary cell
Inter lamina cell
Basal membrane cell 1
Basal membrane cell 2
Mucus cell
Posterior end budding zone cell
Dorsal end proliferation cell
Ventral end proliferation cell
Average Expression
Percent Expressed
COL2A1 (17465-2.8)
COL11A1 (40907-1.6)
COL6A5 (38266-13.9)
Fibril-forming collagen(11625-1.11)
Filamin-1 (61985-5.10)
SCO-spondin (27638-4.18)
Zonadhesin (1371-2.4)
CACNA1H (33261-0.8)
DSCAMs (43601-0.13)
HSD3B2(56879-1.2)
Cartilage matrix like (3387-5.47)
COL21A1(51350-0.19)
Hemicentin-1(38731-0.50)
Hypothetical (48581-0.7)
PIF-like (1303-2.18)
Mucin 17 (11110-0.2)
ADAMTS gon-1(16935-0.17)
C-type lectin (58694-0.22)
C-type lectin 6 (38133-0.8)
C1qDC (55216-2.11)
FIBCD1 (18519-2.15)
WSCD (23804-0.9)
Hypothetical (20639-0.0)
C
Abfrontal
Frontal
Food grove
D COL2A1 (17465-2.8)
E SCO-spondin (27638-4.18)
F Hypothetical (48581-0.7)
G ADAMTS gon-1 (16935-0.17)
C'
Frontal
Abfrontal
Inter lamina cell
Basal membrane cell 1
Basal membrane cell 2
Mucus cell
H Fibril-forming collagen(11625-1.11)
I SCO-spondin (27638-4.18)
J Hypothetical (48581-0.7)
K ADAMTS gon-1 (16935-0.17)
Symbiont DAPI
Frontal
Abfrontal


A
UMAP 1
UMAP 2
Apical ridge ciliary cell
Food groove ciliary cell
Lateral cilia cell
Smooth muscle cell
Intercalary cell
B
Ambiguous
Bacteriocytes
Apical ridge ciliary cell
Food groove ciliary cell
Lateral ciliary cell
Smooth muscle cell
Intercalary cell
Inter lamina cell
Basal membrane cell 1
Basal membrane cell 2
Mucus cell
Posterior end budding zone cell
Dorsal end proliferation cell
Ventral end proliferation cellv
TUBA1A(3489-0.37)
Dlx6a-like(20631-1.16)
ANXA6(54796-6.4)
caveolin-1-like(63462-0.66)
Hypothetical(5544-0.0)
TPR6(38338-5.8)
Capn2(32409-2.44)
Calpain-5-like(56732-0.16)
CD63 antigen-like(681-5.9)
ATP5G(54862-1.17)
ATPeF1A(34427-7.24)
LDL receptor(2971-0.17)
LDL receptor(33659-0.13)
LRP2(33659-0.9)
ACE-like(29477-5.10)
Krueppel-like factor 5(48777-1.5)
Rootletin-like(22362-6.9)
Trichohyalin-like(55568-3.5)
WSC-like(63348-6.2)
Sox-8(40595-5.30)
Sfrp5-like(57424-6.1)
Average Expression
Percent Expressed
C TUBA1A (3489-0.37)
D Dlx6a-like (20631-1.16)
E Hypothetical (5544-0.0)
F
G CD63 Antigen-like (681-5.9)
H
20µm
100µm
10µm
I LDL receptor (2971-0.17)
J Krueppel-like factor 5 (48777-1.5)
K
25µm


A
UMAP_2
UMAP_1
Posterior end budding zone cell
Dorsal end proliferation cell
Ventral end proliferation cell
B
Ambiguous
Bacteriocytes
Apical ridge ciliary cell
Food grove ciliary cell
Lateral cilia cell
Smooth muscle cell
Intercalary cell
Inter lamina cell
Basal membrane cell 1
Basal membrane cell 2
Mucus cell
Posterior end budding zone cell
Dorsal end proliferation cell
Ventral end proliferation cell
SVEP1(61993-0.4)
RTase(41724-1.2)
DMBT1(19571-0.5)
hmw2(63665-1.12)
hmw2(60067-11.23)
unc-89(1581-3.13)
Titin-like(1581-3.4)
Titin-like(1581-3.10)
Gelsolin(17051-23.15)
SLC6A9(13642-5.8)
ESRP1-like(7779-2.7)
Eif-3-like(61540-1.2)
Gata3-like(36085-1.9)
NOP56(30366-1.6)
NOLC1(13236-7.5)
Histone(17533-5.54)
NASP(26572-2.21)
dom-like(59469-1.36)
roa1-like(1371-5.5)
HnRNP D-like(41044-1.15)
hnRNP Q-like(14738-4.9)
hnRNP U1(64706-0.15)
CHD1-like(58896-4.11)
TOP2A-like(59597-7.24)
LYAR-(37615-0.12)
Nucleolin-like(16646-2.15)
YBX1-like(30175-7.5)
Average Expression
Percent Expressed
C
Growth direction
Pannel E: DEPC
Pannel D: PEBZ
Pannel F: VEPC
D
SVEP1(61993-0.4)
Symbiont
DAPI
D'
D''
E
SLC6A9(13642-5.8)
DAPI
F
Nucleolin-like
(16646-2.15)
F'


A
B
DAPI
Symbiont
Merge
B'
Ciliary disk
Ciliary disk
Food grove
C
Bacteriocyte
IC
IC
IC
IC
Bacteriocyte
10μm
D
Bacteriocyte
IC
IC
Symbiont
N
N
N
lysosomes
Basal membrane
Interlaminar lumen
5μm
E
UMAP_1
UMAP_2
F
Ambiguous
Bacteriocytes
Apical ridge ciliary cell
Food grove ciliary cell
Lateral cilia cell
Smooth muscle cell
Intercalary cell
Inter lamina cell
Basal membrane cell 1
Basal membrane cell 2
Mucus cell
Posterior end budding zone cell
Dorsal end proliferation cell
Ventral end proliferation cell
Average Expression
Percent Expressed
Membrane associated
Lysosome associated
Cellular metabolism
Host-symbiont interaction
Pro and Anti- Apoptosis
Chitin synthase 1(6901-7.5)
NMT3(15282-3.10)
KIF16B(54265-1.13)
KLC-like(4784-1.40)
CTSL1-like(61711-5.12)
CTSL1-like(46838-6.40)
GAA(59648-4.5)
CLCN7(7842-1.11)
Rabenosyn-5(54816-0.3)
LYST(52809-1.6)
DMBT1(52188-1.15)
TMPRSS15(15410-0.9)
SLC7A9(36159-5.5)
FYCO1-like(33726-5.7)
PLIN2-like(27158-3.8)
ACC1(55250-5.11)
PLCB4-like(35916-0.6)
ACSBG2-like(28862-1.5)
SLC23A1-like(32311-1.19)
SLC24A3-like(14503-1.28)
SLC12A3-like(44604-3.7)
SLC40A1-like(63447-0.12)
SLCA39A14-like(44428-5.3)
CHDH(34315-4.10)
RBM25-like(42979-0.15)
HIP1(41172-3.3)
BIRC3(6172-0.3)
BIRC7(32300-1.15)
G CTSL1-like(46838-6.40)
Symbiont DAPI
Frontal
Abfrontal
H PLCB4-like(35916-0.6)
Symbiont DAPI
Frontal
Abfrontal
I
Membrane
Chitin synthase 1
NMT-3
Vesicle transport
Kinesin
kinesin light chain-like
FYVE and coiled-coil domain-containing 1-like
Symbiont
Transporters
Vitamin C
Na+ Zn2+
SLC23A1
SLC24A3
Fe Mg2+
Cl-
SLC12A3
Ca2+ K+
SLC40A1
SLC39A14
Glycose and sugars
CO2
O2
Metabolite exchange
b(0,+)-type amino acid transporter 1-like
Globin
Carbonic anhydrase
Cellular metabolism
Perilipin 2-like
Acetyl-carboxylase 2
Fatty acid desaturase 1-like
long-chain-fatty-acid ligase ACSBG2-like
Rabenosyn-5
lysosomal-trafficking regulator-like
H(+)Cl(-) exchange transporter 7-like
Lysosome
Cathepsin L1-like
Lysosomal alpha-glucosidase-like
TMPRSS15
Anti-apoptosis
BIRC3
Pro-apoptosis
RNA-binding 25-like
huntingtin-interacting 1-like


A
Treatment: Fanmao Starvation Reconstitution
IC
FGCC
Bacteriocytes
ARCC
LCC
PEBZ
SMC
DEPC
ILC
VEPC
MC
Amb
BMC2
BMC1
UMAP_2
UMAP_1
B
Fanmao VS Starvation
Starvation VS Reconstitution
Reconstitution VS Fanmao
Cross-state distance
PEBZ
VEPC
Inter lamina cells
Bacteriocytes
Apical ridge ciliary cell
Lateral ciliary cell
Intercalary cell
Mucus cell
Ambiguous
Basal membrane cell 1
Basal membrane cell 2
DEPC
Smooth muscle cell
Food grove ciliary cell
C
Treatment: Fanmao Starvation Reconstitution
Bacteriocytes
Component 2
Component 1
D
Starvation
Fanmao
Reconstitution
Cluster 2
Cluster 1
Bpl_scaf_31643-2.18
Bpl_scaf_52188-1.15
Bpl_scaf_190-0.4
Bpl_scaf_15410-0.9
Bpl_scaf_1894-1.45
Bpl_scaf_6901-7.5
Bpl_scaf_32311-1.19
Bpl_scaf_41403-0.30
Bpl_scaf_19751-0.13
Bpl_scaf_24370-0.5
Bpl_scaf_48274-0.3
Bpl_scaf_34879-0.7
Bpl_scaf_5959-1.9
Bpl_scaf_35916-0.6
Bpl_scaf_50392-1.14
E
CH4
Symbiont
Ribosomal proteins
Lysosome
CMOAT2
SLC6A9
"Farming"
"Milking"
Bacteriocyte
Fanmao
Rnf213
MAP3K14
MON2
Apoptosis
BIRC3
Proliferation
PTPRB
Starvation
O2, CO2, Gln, Glu
CA
SLC4A8
Globin
GS
SLC17A8
Reconstitution