Title

Diversity and functional specialization of oyster immune cells uncovered by integrative single cell level investigations

Authors

Sébastien de La Forest Divonne[1], Juliette Pouzadoux[1], Océane Romatif[1], Caroline Montagnani[1], Guillaume Mitta[2], Delphine Destoumieux-Garzon[1], Benjamin Gourbal[3], Guillaume M. Charrière[1#], Emmanuel Vignal[1#*]

Affiliations

[1] IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

[2] Ifremer, IRD, Institut Louis-Malardé, Univ Polynésie française, UMR 241 SECOPOL, 98179 Taravao, Tahiti - Polynésie française, France

[3] IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Perpignan, France

EV and GC co-supervised the work.

[*] corresponding author: emmanuel.vignal@umontpellier.fr

Abstract

Mollusks are a major component of animal biodiversity and play a critical role in ecosystems and global food security. The Pacific oyster, *Crassostrea (Magallana) gigas*, is the most farmed bivalve mollusk in the world and is becoming a model species for invertebrate biology. Despite the extensive research on hemocytes, the immune cells of bivalves, their characterization remains elusive. Here we were able to extensively characterize the diverse hemocytes and identified at least seven functionally distinct cell types and three hematopoietic lineages. A combination of single-cell RNA sequencing, quantitative cytology, cell sorting, functional assays and pseudo-time analyses was used to deliver a comprehensive view of the

distinct hemocyte types. This integrative analysis enabled us to reconcile molecular and cellular data and identify distinct cell types performing specialized immune functions, such as phagocytosis, reactive oxygen species production, copper accumulation, and expression of antimicrobial peptides. This study emphasized the need for more in depth studies of cellular immunity in mollusks and non-model invertebrates and set the ground for further comparative immunology studies at the cellular level.

Introduction

Mollusca is the second largest invertebrate phylum, after Arthropoda, and the largest marine phylum, comprising approximately 23 % of all known marine organisms (*1*). Among them, bivalves exhibit a high diversity and a rich evolutionary history (*2*). The Pacific oyster *Crassostrea (Magallana) gigas* (*C. gigas* - Thunberg, 1793) (NCBI:txid29159) is a sessile filter-feeding bivalve that thrives in a variety of stressful environments ranging from intertidal to deep-sea conditions (*3*). It is a key species for the aquaculture industry worldwide (*4*). Several infectious diseases affect *C. gigas* at different life stages, which impacts its production. Given the significant socio-economic value of this species, there has been an increased focus on understanding and mitigating these diseases (*5*). The causes of these mortalities can involve a variety of pathogens, including viruses, bacteria and parasites that can be responsible for the mortality events affecting *C. gigas* (*6*). One of the most extensively researched infectious diseases is POMS (Pacific Oyster Mortality Syndrome), a polymicrobial disease responsible for mass mortalities of juvenile oysters (*7*). The disease is triggered by the OsHV-1 μVar herpesvirus, which alters the immune defenses of oysters, allowing the colonization of opportunistic bacteria, including *Vibrio*, that cause hemocyte lysis and bacteremia, ultimately leading to animal death (*7*). Other bacterial pathogens have also been identified as a contributing factor in mass mortalities of adult Pacific oysters in several countries. The most notable is *Vibrio aestuarianus* which affects adult oysters in Europe (*8*, *9*). To date, the majority of oyster pathogens or opportunistic pathogens that have been

55 characterized in detail have been found to subvert hemocyte defenses for their own benefit.
56 These include the OsHV-1 μVar herpesvirus (*10*) and virulent *Vibrio* strains of the species
57 *Vibrio crassostreae* and *Vibrio tasmaniensis* (*11*), *Vibrio aestuarianus* and *Vibrio harveyi* (*12*),
58 highlighting the critical role of hemocytes in oyster immunity. The development of immune-
59 based prophylactic treatments, such as immune-priming and immune-shaping, represents a
60 promising avenue for enhancing the natural defenses of oysters against pathogens and
61 increasing their survival rate. Nevertheless, the advancement of such therapies is still
62 constrained by a dearth of knowledge regarding the underlying molecular and cellular
63 mechanisms (*13*, *14*).
64 The study of hemocytes has a long history, dating back to the 1970s (for review see (*15*)).
65 Hemocytes are cellular effectors of the immune system. They engage in phagocytosis to
66 engulf and destroy potential pathogens, neutralizing parasites by encapsulation, or preventing
67 pathogen dissemination by cell aggregation and the release of extracellular DNA traps (*16*).
68 Furthermore, they engage in the humoral response by releasing cytokines, antimicrobial
69 peptides, and reactive oxygen species (ROS), which enable them to combat pathogens (*17*). In
70 addition to their role in oyster immunity, hemocytes have been implicated in numerous
71 physiological processes, including shell repair (*18*), wound healing, nutrient transport, and
72 environmental contaminant removal (*19*). Despite this acquired knowledge, hemocytes remain
73 an under-characterized population of circulating immune cells. The lack of a unified
74 classification and of molecular and functional genetic tools hinders our understanding of
75 lineage ontogeny and functional specialization. Several studies have proposed different
76 classifications of hemocytes in the *Ostreidae* family, with 3 to 4 hemocyte types reported (*15*).
77 These classifications are primarily based on either microscopic or flow cytometry analyses. In
78 *C. gigas*, three primary hemocyte cell types have been classically identified : blast,
79 hyalinocyte, and granulocyte cells. Blast-like cells are considered as undifferentiated
80 hemocyte types (*20*), hyalinocytes (*21*) seem to be more involved in wound repair, and
81 granulocytes, more implicated in immune surveillance. The latter are considered as the main

82 immunocompetent hemocyte types (*22*). While the immune response of *C. gigas* has been
83 extensively studied using classical transcriptomics (*23–25*) at the whole animal, tissue, or
84 circulating hemocyte levels, these approaches have failed to consider the diversity of
85 hemocyte cell types and lineages that underpin these responses (*26*). However, it is still
86 imperative to accurately describe the diversity of these cells, understand their ontogeny, and
87 delineate cell lineages to comprehend their specific roles. The advent of single-cell RNA
88 sequencing (scRNA-seq) techniques has enabled the monitoring of global gene expression at
89 the single-cell level with thousands of individual cells in a single experiment. This provides a
90 unique opportunity to overcome these limitations and deepen our understanding of hemocyte
91 diversity and function in bivalves. Recently scRNA-seq was used to provide a first molecular
92 description of a hemocyte population in the oyster *Crassostrea hongkongensis* (*27*). However,
93 in the absence of morphological and/or functional characterization studies, the authors could
94 not deduce the hemocyte cell types to be matched to the transcriptomic profiles generated by
95 scRNA-seq.
96 While numerous transcriptomic analyses have been conducted on *C. gigas* hemocytes, none
97 have adopted a single-cell approach. In this study, we present an integrative analysis of the
98 diversity of *C. gigas* hemocytes at the single-cell level. To this end, we combined scRNA-seq
99 from a pathogen-free adult oyster combined with cytological, cell fractionation and functional
100 assays. This approach allowed us to create a comprehensive transcriptomic, cytological and
101 functional atlas of hemocyte cell types. Our scRNA-seq analyses identified 7 distinct
102 transcriptomic populations and functional annotation revealed distinct populations with
103 specific functions, including phagocytosis, oxidative burst, energetic metabolism, enhanced
104 transcription, translation and cell division. Quantitative cytology enabled the identification of
105 7 morphologically distinct hemocyte cell types, which allowed us to reconcile molecular and
106 cytological data. Density gradients were used to separate hemocyte cell types and qPCR or
107 functional assay analyses were performed to validate cell type-specific markers. By employing
108 this integrated approach, we could identify 1 type of hyalinocyte, 2 types of blasts and 4 types

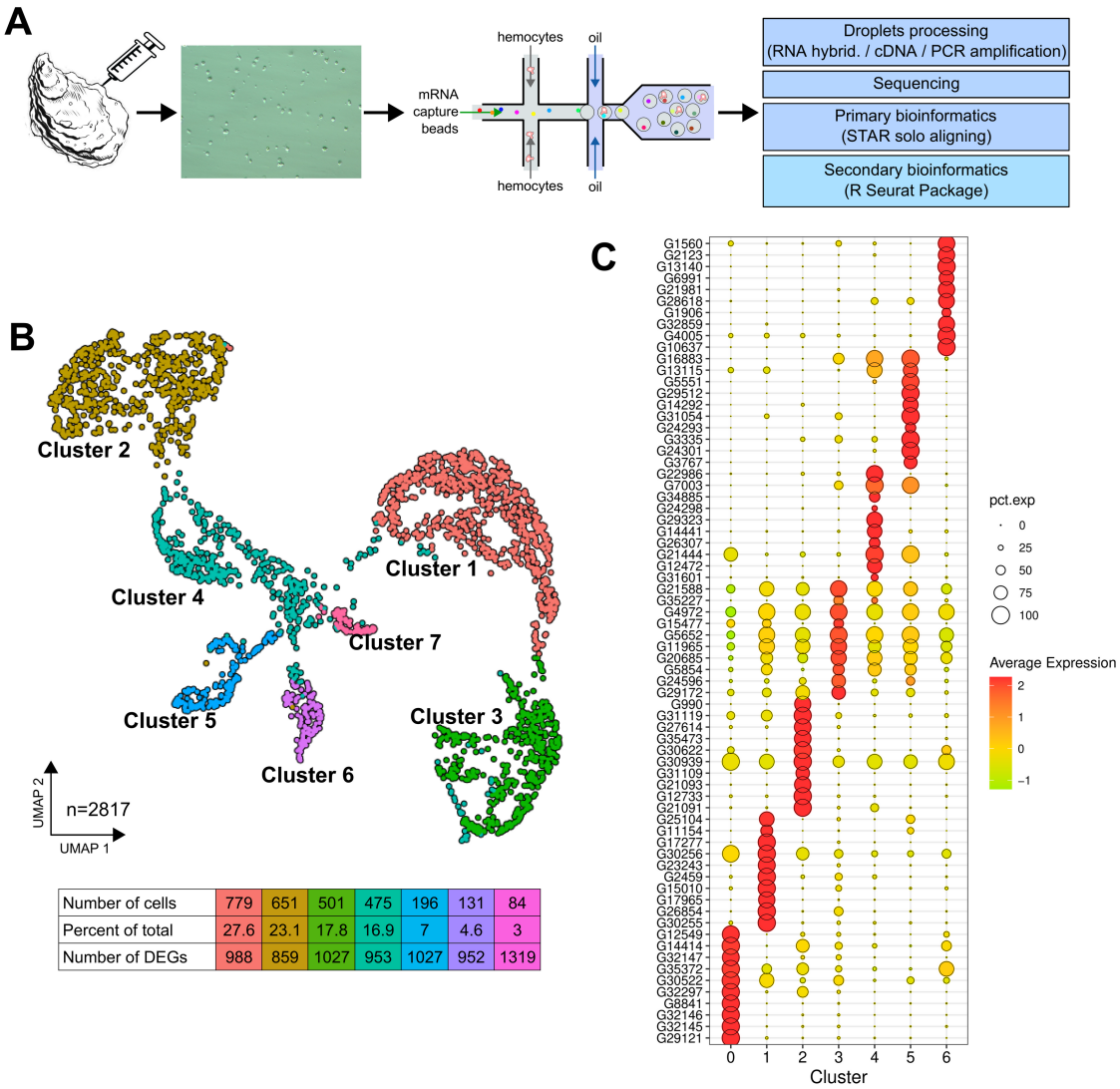
109 of granular cells. Furthermore, we identified cell types that perform antimicrobial functions
110 through phagocytosis, ROS production, copper accumulation, and expression of antimicrobial
111 peptides. Finally, trajectory analysis of scRNA-seq data combined with functional analysis
112 revealed distinct differentiation pathways that may control hemocyte ontology and
113 differentiation processes. Based on these findings, we propose a more comprehensive and up-
114 to-date classification of *C. gigas* hemocytes, with a more accurate description of the different
115 cell types, their potential ontology and a precise description of their sub-functionalization.
116

117 **Results**

118 **Single-cell RNA sequencing reveals 7 distinct transcriptomic clusters of circulating**
119 **immune cells in oysters.**

120 Oysters are known to exhibit a high degree of individual genetic polymorphism, including
121 Copy Number Variation (CNV) and Presence Absence Variation (PAV) (*28*). To prevent
122 misinterpretation of the single-cell transcriptomic data, and to characterize the hemocyte cell
123 types and their heterogeneity, we sampled hemocytes from a unique known pathogen-free
124 animal (Ifremer Standardized Animal, 18-month-old) and applied single-cell drop-seq
125 technology. A total of 4,950 cells were loaded, with a target recovery of 3,000 single
126 hemocytes for analysis. **(Fig. 1A)**. The scRNA-seq library was generated and sequenced,
127 resulting in 127,959,215 high-quality filtered reads available for single-cell analysis (ENA
128 project accession number PRJEB74031). Primary bioinformatics analysis was performed
129 using the STAR solo aligner software (*29*) against the *C. gigas* genome (Genbank reference
130 GCA_902806645.1) from the Roslin Institute (*30*). Of the 127,959,215 reads, 97 % showed a
131 valid barcode and 89.2 % were successfully mapped to the genome with a saturation of 75.6
132 %. A total of 2,937 cells were profiled, yielding a median of 1,578 genes and 4,412 unique
133 molecular identifiers (UMIs) per cell among the 23,841 total genes detected, with a
134 sequencing saturation of 75.6 % **(Supp. Table S1)**. Secondary bioinformatic data processing
135 was conducted using the Seurat R package (version 4.3.0) (*31*). The data set was filtered to

136 remove data corresponding to empty droplets or cell doublets. Cells with a gene number
137 between 750 and 4,000 and less than 5 % mitochondrial genes were retained. After quality
138 control processing, 120 cells corresponding to empty droplets and cell doublets were removed,
139 and 2,817 cells were processed for data normalization. Finally, we performed linear
140 dimensional reduction and clustering on the 3,000 most variable genes from 2,817 cells
141 **(Supp. Fig. S1)**. Dimension reduction and clustering led to identifying 7 transcriptomic
142 clusters, within which hemocytes were distributed. These 7 different clusters represented 27.6,
143 23.1, 17.8, 16.9, 7, 4.6 and 3 % of the total cells **(Fig. 1B)**. For each transcriptomic cluster, the
144 heatmap displays the top 10 positively enriched marker genes identified by Seurat's
145 differential expression analysis, highlighting cluster-specific overexpression **(Fig. 1C)**.
146 Average Log2FC values and percentage of expression in each cluster relative to all other
147 clusters were calculated for the ten most differentially expressed genes **(Fig. 1D and Table 1).**
148 Clear transcriptomic cell clusters were detected as well as specific gene markers for each
149 cluster **(Supp. Data 1)**. Only the transcriptomic signature of cluster 4 was less contrasting
150 than that of the other clusters **(Fig. 1D)**.



Number of cells	779	651	501	475	196	131	84
Percent of total	27.6	23.1	17.8	16.9	7	4.6	3
Number of DEGs	988	859	1027	953	1027	952	1319

Fig. 1. scRNA-seq analysis of *C. gigas* circulating hemocytes reveals 7 transcriptomic cell clusters (A) Schematic of the scRNA-seq 10X Genomics Chromium microfluidic technology and bioinformatics processing workflow used. Dissociated hemocytes were collected from a pathogen-free oyster and encapsulated in droplets for processing. After sequencing, the data were processed bioinformatically. **(B)** Uniform Manifold Approximation and Projection (UMAP) plot for dimensional reduction of the data set and summary of cells and the number of Differentially Expressed Genes (DEGs) in each cluster. The table shows the characteristics (number of cells, percentage of total cells and number of Differentially Expressed Genes in each cluster) of the seven clusters identified. **(C)** Dot plot representing the ten most enriched DEGs per cluster based on average expression (avg_log2FC). The color gradient of the dot

163 represents the expression level, while the size represents the percentage of cells expressing

164 each gene per cluster.

165

Gene	log2FC	Pct 1	Pct 2	Description
				Cluster 1
G29121	4.31	1	0.099	L-galactono-gamma-lactone oxidase
G32145	4.18	0.992	0.079	Angiopoietin-1 receptor
G32146	3.90	0.985	0.063	Angiopoietin-1 receptor
G8841	3.13	0.988	0.049	C-type lectin domain-containing protein
G32297	2.88	0.99	0.232	PNPLA domain-containing protein
G30522	2.76	1	0.515	Cytochrome P450 2D28
G35372	2.69	1	0.483	60S ribosomal protein L26
G32147	2.45	0.969	0.131	Angiopoietin-1 receptor
G14414	2.45	1	0.343	DUF4773 domain-containing protein
G12549	2.38	0.974	0.132	Glutaredoxin domain-containing protein
				Cluster 2
G30255	5.31	1	0.15	Metalloproteinase inhibitor 3
G26854	4.24	1	0.16	Stanniocalcin
G17965	4.21	0.94	0.04	Complement C1q-like protein 2
G15010	3.99	0.99	0.17	Kyphoscoliosis peptidase
G2459	3.82	1	0.21	*Uncharacterized protein*
G23243	3.51	0.98	0.07	Fibronectin type-III domain-containing protein
G30256	3.23	1	0.69	NTR domain-containing protein
G17277	3.21	0.99	0.04	Putative modulator of levamisole receptor-1
G11154	3.16	0.68	0.05	Collagen alpha-1(XII) chain
G25104	3.1	0.84	0.08	C1q domain-containing protein
				Cluster 3
G21091	5.51	0.99	0.13	X-box binding protein-like protein
G12733	5.42	1	0.12	Galectin
G21093	5.17	0.94	0.05	IoIG
G31109	4.36	0.77	0.03	Cystatin domain-containing protein
G30939	4.25	1	0.86	Peptide ABC transporter permease
G30622	3.95	1	0.21	BHLH domain-containing protein
G35473	3.86	0.96	0.04	*Uncharacterized protein*
G27614	3.78	0.96	0.13	G Protein Receptor F1-2 domain-containing protein
G31119	3.77	0.99	0.42	Arrestin C domain-containing protein
G990	3.6	0.93	0.03	LITAF domain-containing protein
				Cluster 4
G29172	3.09	0.78	0.49	G Protein Receptor F1-2 domain-containing protein
G24596	2.37	0.78	0.25	High mobility group protein DSP1
G5854	1.8	0.64	0.38	N-acetyltransferase domain-containing protein
G20685	1.59	0.88	0.53	Ribosomal protein L22
G11965	1.57	0.92	0.7	Ribosome L4 associated C domain-containing protein
G5652	1.54	0.97	0.75	60S ribosomal protein L30
G15477	1.52	0.55	0.3	Metalloendopeptidase
G4972	1.49	0.96	0.8	60S ribosomal protein L12
G35227	1.46	0.55	0.06	ML domain-containing protein
G21588	1.42	0.94	0.69	60S ribosomal protein L13

Gene	log2FC	Pct 1	Pct 2	Description
				Cluster 5
G31601	5.24	0.38	0.06	Polyribonucleotide nucleotidyltransferase
G12472	3.99	0.85	0.03	Galectin
G21444	3.58	1	0.41	*Uncharacterized protein*
G26307	3.42	0.6	0.02	*Uncharacterized protein*
G14441	3.35	0.84	0.08	CUBN
G29323	3.21	0.9	0.01	GTPase IMAP family member 4
G24298	3.18	0.29	0.05	Putative Transmembrane protease serine 9
G34885	3.13	0.59	0.01	Gliding motility-associated C-terminal domain-containing protein
G7003	2.55	0.99	0.16	ETS-related transcription factor Elf-4
G22986	2.37	0.94	0.1	Metalloendopeptidase
				Cluster 6
G3767	6.45	0.74	0.02	Natterin-1
G24301	6.16	0.99	0.02	Aspartate-semialdehyde dehydrogenase
G3335	5.19	1	0.16	*ncRNA*
G24293	4.57	0.59	0.01	*Uncharacterized protein*
G31054	3.56	1	0.13	GATA-binding factor 3
G14292	3.29	0.87	0.03	Membrane-associated guanylate kinase inverted 3
G29512	3.2	0.99	0	Caveolin
G5551	2.93	0.97	0.02	Neuronal acetylcholine receptor subunit alpha-6
G13115	2.73	0.83	0.26	*ncRNA*
G16883	2.62	1	0.2	Allograft inflammatory factor
				Cluster 7
G10637	4.21	0.99	0.03	BZIP domain-containing protein
G4005	4.07	0.99	0.21	Ras-related protein Rab-20
G32859	3.9	0.93	0.04	G Protein Receptor F1-2 domain-containing protein
G1906	3.82	0.5	0.02	Glutathione peroxidase
G28618	3.53	0.91	0.09	Sporozoite and liver stage asparagine-rich protein
G21981	3.5	0.94	0.05	*ncRNA*
G6991	3.4	0.85	0.01	Rho-GAP domain-containing protein
G13140	3.12	1	0.03	DBH-like monooxygenase protein 1
G2123	3.05	0.94	0.02	SPRY domain-containing SOCS box protein 3
G1560	3.02	0.95	0.17	Glutathione peroxidase

167 **Table 1. Top 10 overexpressed genes identified in each transcriptomic cluster.** The first

168 column indicates the gene number according to the annotation. 'log2FC' represents the log2

169 fold change of the gene in the cluster compared to all other cells. 'Pct1' is the percentage of

170 cells expressing the gene in the cluster and 'Pct2' is the fraction of cells expressing the gene in

171 all other clusters. The description is the annotation of the expressed gene (adjusted p-value <

172 0.05).

173 **KEGG pathways and GO-terms analyses reveal functional diversity in *C. gigas***

174 **hemocytes.**

175 The ScRNA-seq data demonstrated that specific functions are carried out by the hemocyte cell

176 types that comprise the seven transcriptomic clusters. A preliminary overview of the functions

177 over-represented in each cluster was obtained through a KEGG pathway analysis on the

178 overrepresented transcripts (Log2FC > 0.25) using the *C. gigas* annotation provided by the

179 DAVID consortium (*32*), thereby identifying pathways specifically enriched in each cluster.
180 Cluster 1 demonstrated enrichment in viral processing and endocytosis **(Figure 2A and**
181 **Supplemental Table S2)**. Additionally, it demonstrated enrichment in pyruvate metabolism,
182 glycolysis/gluconeogenesis and the pentose phosphate pathways. Clusters 1 and 3 exhibited a
183 distinctive enrichment in carbohydrate metabolism and endocytosis. In particular, cluster 3
184 exhibited enriched transcripts associated with glycolysis/gluconeogenesis and TCA cycle
185 activities. Cluster 4 was enriched in protein synthesis, including transcription (spliceosome,
186 ribosome, nucleocytoplasmic transport and mRNA surveillance pathway) folding, sorting and
187 degradation of protein pathways. Clusters 2, 5 and 6 exhibited a shared signature of
188 enrichment in ribosome-related genes. Cluster 2 demonstrated a specific enrichment in motor
189 protein-coding genes responsible for cell motility and xenobiotic metabolism. Remarkably,
190 clusters 1, 3, 4 and 5 exhibited enriched oxidative phosphorylation transcripts, whereas the
191 transcripts of cluster 7 were enriched in vesicular trafficking and endo-lysosomal pathways
192 (endocytosis, endosome, phagosome, lysosome, auto and mitophagy). Cluster 7 displayed also
193 the signature of Wnt pathway components.

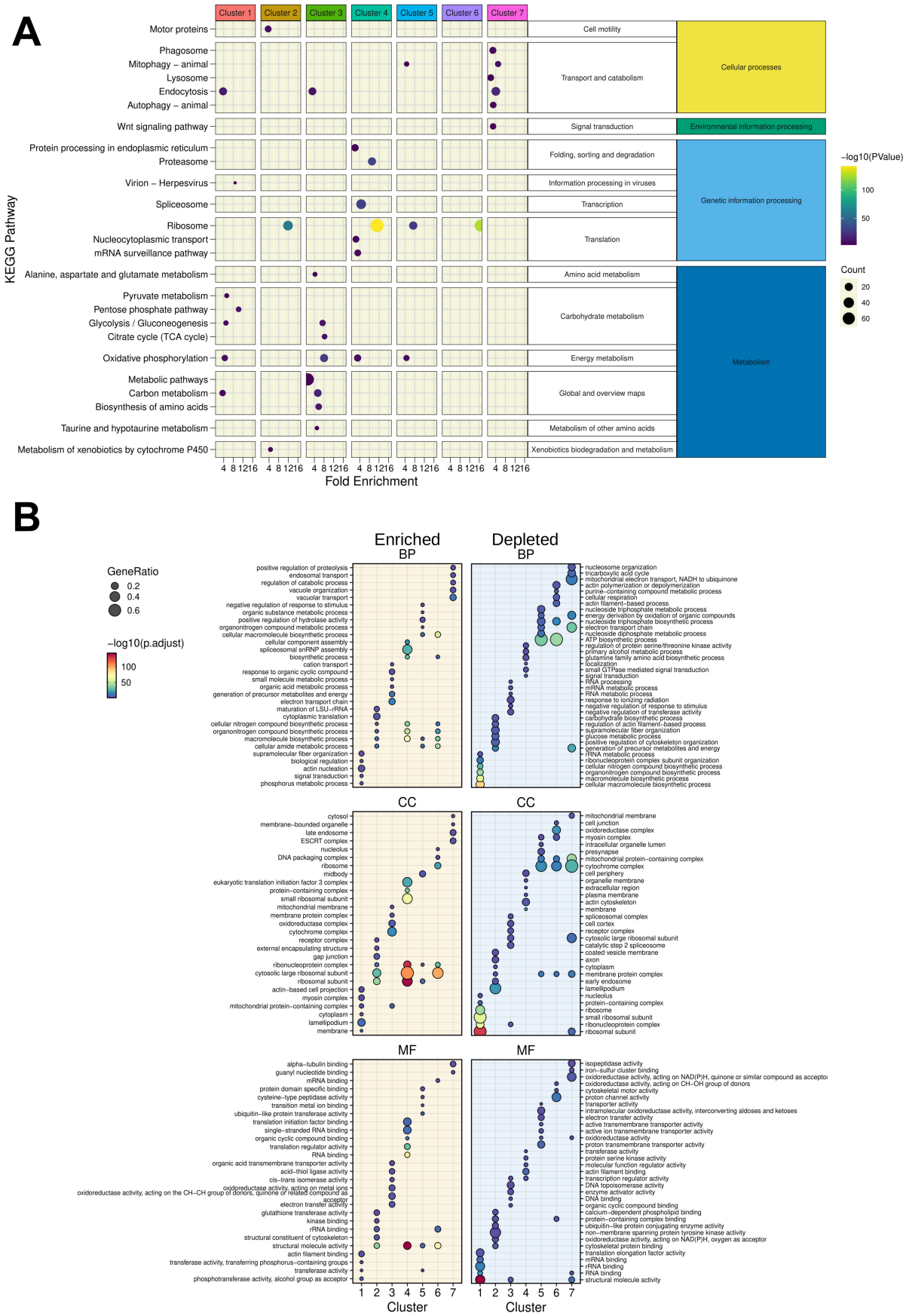


Fig. 2. KEGG and Gene Ontology analysis of the gene signature in each cluster. (A) A synthetic representation of the KEGG pathway analysis is shown. Colored columns represent the 7 transcriptomic clusters. Each row is a KEGG pathway, the colored dot represents the -

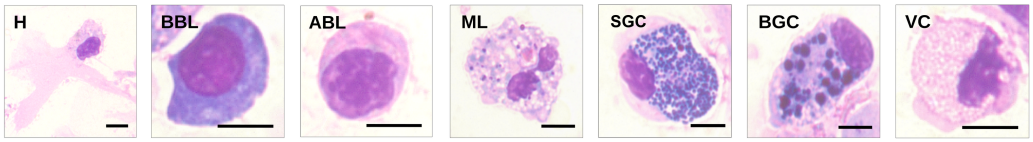
197 log(p-value) and the dot size represents the number (count) of enriched genes in each pathway
198 category. The fold enrichment is shown on the x-axis. Panels **(B)** show the results of Gene
199 Ontology terms (GO-terms) for Biological Processes (BP), Cellular Components (CC) and
200 Molecular Functions (MF) respectively, obtained with the genes of each cluster (Absolute
201 value Log2FC > 0.25 and significant p-value < 0.001) using RBGOA analysis (p-value <=
202 0.001) for three different ontology universes. Each panel corresponds to one ontology
203 universe, and the analysis highlights enriched and depleted terms. The dot size indicate the
204 proportion of significant GO-terms and the gradient scale the p-adjusted value.
205
206 For a more detailed functional characterization of each transcriptomic cluster, we performed a
207 re-annotation of the *C. gigas* genome using a combination of tools to enhance the GO term
208 richness of the existing annotation prior to functional GO term analysis. The original
209 annotation for *C. gigas*, based on the work of Penaloza et al. (*30*), provided GO annotations for
210 18,750 out of 30,724 genes (61%). In order to improve upon these baseline annotations, we
211 used the *C. gigas* genome (Genbank reference GCA_902806645.1) and the associated gff3
212 annotation file from the Roslin Institute (*30*). These files were used to extract and process the
213 longest CDSs for GO-term annotation using the Orson pipeline (see Materials and Methods).
214 Of the 30,724 extracted CDSs, 22,462 were annotated (GO-terms and sequence description),
215 yielding an annotation percentage of 73.1 %. Of the 30,724 CDSs, 22,391 were annotated with
216 Molecular Functions (MF), Biological Processes (BP), and Cellular Components (CC) GO-
217 terms (**Supp. Fig. S2 and Supp. Data 2**). Overall, this improved annotation coverage
218 significantly enhances the resolution of our downstream functional analyses. Using the GO-
219 term annotation and the Log2FC of genes calculated after scRNA-seq processing in each
220 cluster, GO enrichment analysis was performed by rank-based gene ontology analysis
221 (RBGOA) (*33*). RBGOA analysis was performed on the GO-terms identified in each cluster
222 (**Supp. Data 3**). The results are presented in **Figures 2B, C and D**.
223

224 The scRNA-seq-based analysis identified seven distinct transcriptomic profiles for each cell
225 cluster, thereby shedding light on greater heterogeneity and functional diversity of *C. gigas*
226 hemocytes than previously described. Cluster 1, comprising 27.6 % of cells, is characterized
227 by GO-terms related to myosin complex, lamellipodium, membrane and actin cytoskeleton
228 remodelling, as well as phosphotransferase activity. This is evidenced by an enrichment in
229 actin nucleation, phosphorus metabolic process and signal transduction BP, actin filament
230 binding and phosphotransferase activity MF and lamellipodium and actin-based cell projection
231 CC. Cluster 2, comprising 23.1 % of cells, has an increased translation activity, as indicated
232 by enrichment in BP, MF and CC terms related to rRNA binding, cytoplasmic translation,
233 maturation of LSU-rRNA, and ribosomes respectively. Cluster 3, representing 17.8 % of cells,
234 shows enrichment in cellular oxidation and actin nucleation, as evidenced by an enrichment in
235 BP and MF related to oxidoreductase activities acting on metal ions, electron transfer and
236 cellular oxidation, and CC related to cytochromes, oxidoreductase complex, and mitochondrial
237 protein-containing complexes. Cluster 4, representing 16.9 % of cells, has transcriptomic
238 signatures reminiscent of spliceosome assembly, RNA maturation, and macromolecules
239 biosynthesis (BP related to biosynthetic process, macromolecules biosynthesis process, and
240 spliceosome assembly, MF related to RNA and ssRNA binding and translation initiation, and
241 CC related to ribosomes and ribonucleoprotein complexes). Cluster 5, comprising 7 % of
242 cells, demonstrates enrichment in structural molecule activity, hydrolase regulation and
243 ubiquitin-like protein transferase activity, as evidenced by its BP, MF, and CC related to
244 ribosome and midbody localization. Cluster 6, comprising 4.6 % of cells, is characterized by a
245 specific nucleosome organization, translation, and biosynthetic processes, with related BP,
246 MF, and CC tied to nucleolus, ribosomes, ribonucleoprotein complex and rRNA binding.
247 Cluster 7, representing a mere 3 % of cells, is characterized by multivesicular body sorting,
248 vacuolar transport, vacuole organization, and positive regulation of proteolysis, with related
249 BP, MF, and CC of the ESCRT machinery.

250 **Seven morphologically distinct immune cell types are identified by quantitative cytology**
251 **and transcriptomic markers.**

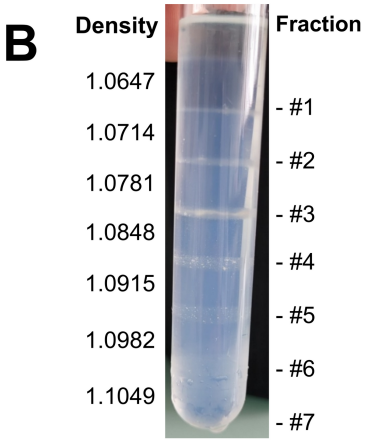
252 Cytological studies were conducted using MCDH staining to better characterize the diversity
253 of circulating hemocytes in *C. gigas*. Seven distinct hemocyte morphotypes were identified:
254 three non-granular (acidophilic blasts, basophilic blasts and hyalinocytes) and four granular
255 (small granule cells, big granule cells, vesicular cells and macrophage-like) **(Fig. 3 and Supp.**
256 **Fig. S3)**. Hyalinocytes (30 % of total hemocytes) are large cells with an irregular spreading
257 membrane. They contain an azurophilic cytoplasm without granulations and their irregular
258 nucleus varies in size (**Fig. 3A panel H**). Basophilic blasts characterized by a basophil
259 cytoplasm (**Fig. 3A panel BBL**) and acidophilic blasts with acidophil cytoplasm (**Fig. 3A**
260 **panel ABL**) accounted for 17 % and 14 %, respectively, of the total hemocytes. They are
261 rounded cells without granulation, with a uniform and regular dense nucleus and a high
262 nucleo-cytoplasmic ratio. Macrophage-like cells (20 % of the hemocytes) present an irregular
263 membrane punctuated by rare pseudopodia with a polylobed nucleus and a basophilic
264 cytoplasm where polychromatic inclusions of various sizes could be observed (**Fig 3A. panel**
265 **ML**). Small granule cells (13 % of total hemocytes) have an irregular membrane punctuated
266 by rare pseudopodia, an acidophilic cytoplasm, and numerous homogeneous purple granules
267 (**Fig. 3A panel SGC**). Big granule cells (4 % of total hemocytes) are rounded cells with a
268 basophilic cytoplasm containing large, slightly dark purple to black vesicles of heterogeneous
269 size (**Fig. 3A panel BGC**). Vesicular cells (2 % of total hemocytes) are rounded cells with
270 acidophilic cytoplasm, rich in homogeneous transparent and fluorescent to UV light vesicles
271 with an irregular nucleus (**Fig. 3A panel VC**). Only blasts (BBL, ABL) have a nucleo-
272 cytoplasmic ratio greater than 0.6, whereas the other hemocytes described have a nucleo-
273 cytoplasmic ratio less than 0.3.
274 To gain further insight into the functional characterization of these diverse hemocyte
275 morphotypes and to associate them with the transcriptomic clusters identified in the scRNA-
276 seq, we enhanced an existing hemocyte fractionation approach using an isopycnic Percoll

density gradient to sort the hemocytes (*34*). Cell sorting was performed using a discontinuous Percoll gradient with densities ranging from 1.0647 to 1.1049. Seven distinct density fractions were established (**Fig. 3B**) and the hemocyte composition was then characterized by cytological analysis of each fraction (**Fig. 3C;** see also **Supp. Fig. S4** for statistical significance). The uneven distribution of hemocyte morphotypes along the density gradient enabled a relative separation of the different cell populations (**Supp. Fig S4**). In summary, hyalinocytes were significantly enriched in the first fraction, while macrophage-like cells were significantly enriched in fraction 3 and fraction 4. Acidophilic blasts were significantly enriched in fraction 2. Basophilic blasts were significantly enriched in fractions 2 and 3. Vesicular cells were enriched in fraction 5, big granule cells were enriched in fraction 4 and small granule cells in fractions 6 and 7 (**Fig. 3C and Supp. Fig. S4**).



Cell types	Hyalinocyte (H)	Basophilic Blast-Like (BBL)	Acidophilic Blast-Like (ABL)	Macrophage Like (ML)	Small Granule Cells (SGC)	Big Granules Cell (BGC)	Vesicular Cell (VC)
Percentage[a]	30.09 ± 6.48%	17.11 ± 5 %	13.79 ± 4.41 %	19.79 ± 5.56 %	12.91 ± 4.55 %	4.11 ± 1.99 %	2.19 ± 1.26 %
Cell Size (µm)	21-31	7-11	6-8	12-17	14-25	10-15	8-13
Nucleus Size (µm)	6-8	4-7	4-5	4-6	5-7	4-6	4-6
Cytoplasm	Acidophilic	Basophilic	Acidophilic	Acidophilic	Acidophilic	Basophilic	Acidophilic
Pseudopods	Long	Unobvious	Unobvious	Rare	Rare	Unobvious	Unobvious
N/C[b] ratio	N/C ≤ 0.1	N/C ≥ 0.6	N/C ≥ 0.7	N/C ≤ 0.2	N/C ≤ 0.2	N/C ≤ 0.2	N/C ≤ 0.3
Granules	-	-	-	Mixed	Basophilic	Basophilic	Transparent
			Size	*0.3-1 µm*	*0.3 µm*	*1 µm*	*0.3 µm*
			Shape	*Round*	*Round*	*Ellipsoid*	*Round*
			Particularities	-	-	*Brown*	*Fluo. at 240nm*

[a] – Percentage of the cell type in total hemolymphe
[b] – Nucleus / Cytoplasm surface ratio



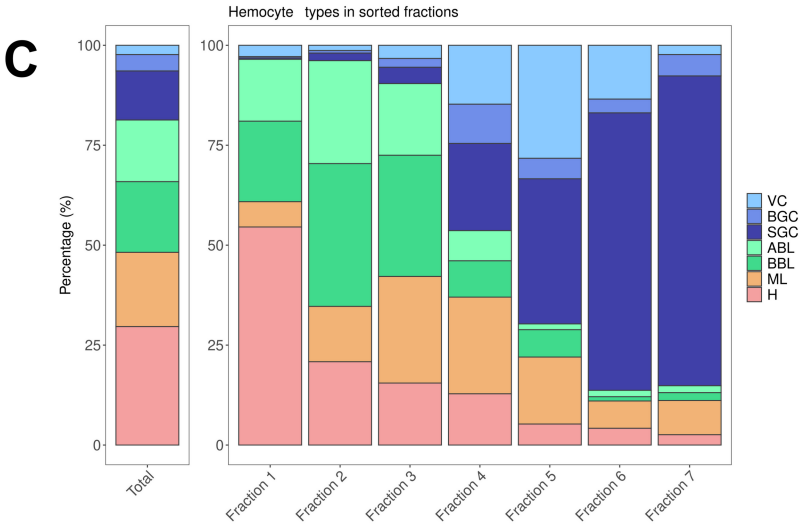


Fig 3. *C. gigas* naive hemocyte formula and Percoll gradient hemolymph fractionation

(A) Morphology, percentages and characteristics of the 7 cell types identified by MCDH staining. **H** : Hyalinocyte, **ML** : Macrophage Like, **BBL** : Basophilic Blast Like cell, **ABL** : Acidophilic Blast Like cell, **SGC** : Small Granule Cell, **BGC** : Big Granule Cell, **VC** : Vesicular Cell. Scale bar : 5 µm. For the ML, SGC, BGC, and VC hemocyte types, size refers to the average granule diameter, shape describes the morphology of the granules (e.g., round, ellipsoid), and particularities highlight distinguishing features such as granule color or fluorescence properties observed under specific staining or imaging conditions. **(B)** Sorting of hemocytes on a discontinuous Percoll gradient. 7 fractions were identified along the gradient at the top of each density cushion (from d=1.0647 at #1 to d= 1.1049 at #7). **(C)** Representation of the average values (from 5 different fractionation experiments) of the different hemocyte types in the seven percoll gradient fractions compared to the average hemolymph composition of a naive oyster (Total). **VC** : Vesicular Cells, **BGC** : Big Granule

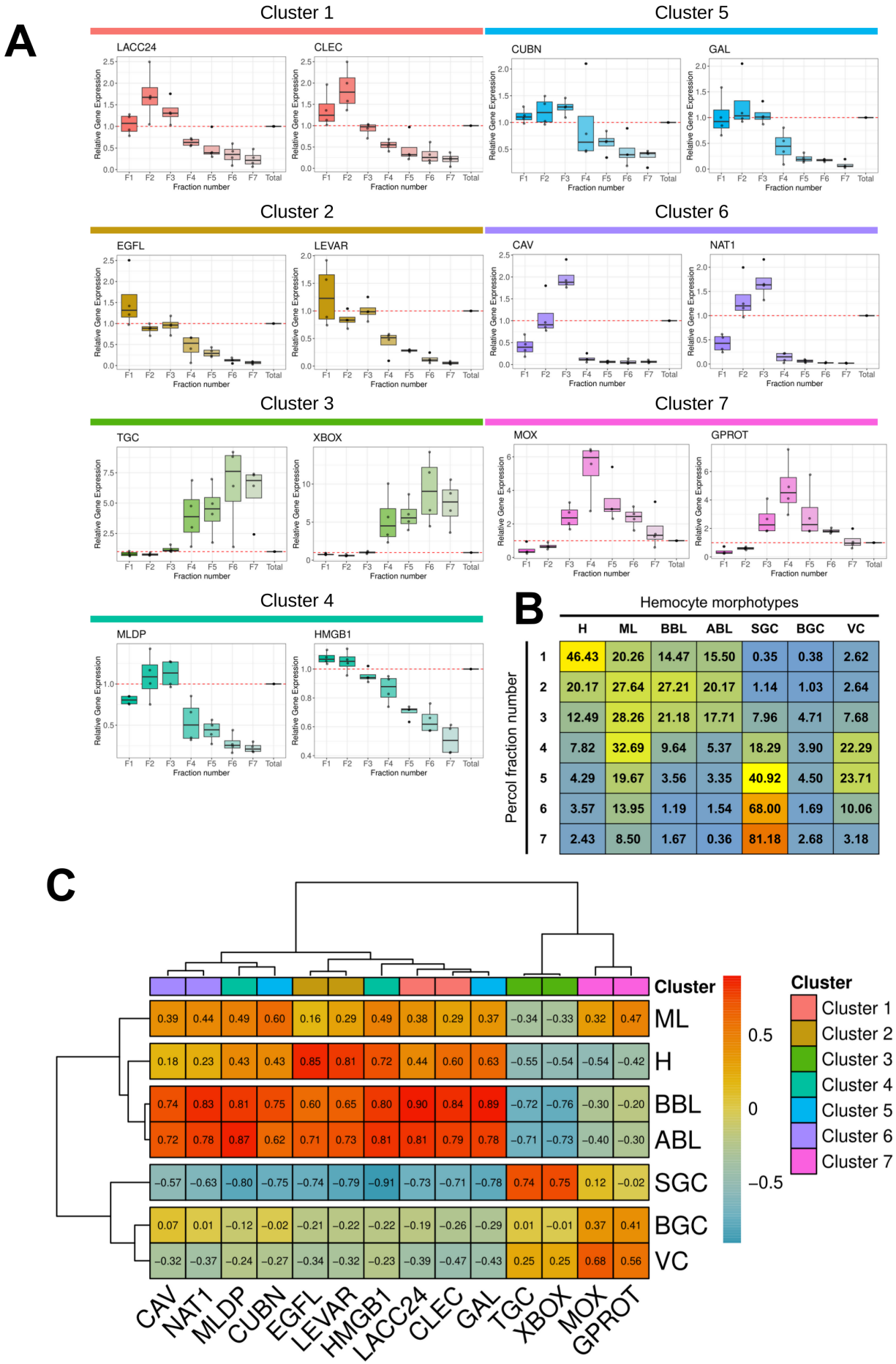
302 Cells, **SGC** : Small Granule Cells, **ABL** : Acidophilic Blast Like cells, **BBL** : Basophilic Blast
303 Like cells, **ML** : Macrophage Like cells and **H** : Hyalinocytes respectively. (**Supp. Fig. S4** for
304 statistics).

305 To identify the transcriptomic cell clusters corresponding to the different hemocyte
306 morphotypes, we used RT-qPCR to detect the expression of different cluster-specific marker
307 genes in the different hemocytes in the Percoll density fractions. The marker genes were
308 selected without any a priori assumptions based solely on their enrichment levels (Log2FC)
309 and their percentage of expression (Pct1 / Pct2 ratio) in the cluster of interest relative to all
310 other cell clusters (**Table 2).** The expression level of each marker in the seven different
311 clusters confirmed that the 14 selected marker genes were differentially expressed in each
312 scRNA-seq transcriptomic cluster (**Supp. Fig. S5**).

Gene Number	Avg. Log2FC	Pct.1/Pct.2 Ratio	Cluster	Description	Marker Name
G8994	2.05	38.20	1	Laccase-24	LACC24
G8841	3.13	20.16		C-type lectin domain-containing protein	CLEC
G24846	2.31	31.52	2	EGF-like domain-containing protein 8	EGFL
G17277	3.21	23.00		Putative modulator of levamisole receptor-1	LEVAR
G22387	3.16	16.76	3	TGc domain-containing protein	TGC
G21091	5.51	7.69		X-box binding protein-like protein	XBOX
G35227	1.46	9.98	4	ML domain-containing protein	MLDP
G31074	0.94	6.51		High mobility group protein B1	HMGB1
G14441	3.35	11.01	5	Cubilin	CUBN
G12472	3.99	24.91		Galectin	GAL
G29512	3.20	246.25	6	Caveolin	CAV
G3767	6.45	41.11		Natterin-1	NAT1
G13140	3.12	40.00	7	DBH-like monooxygenase protein 1	MOX
G32859	3.90	23.82		G protein receptor F1-2 domain-containing protein	GPROT

314 **Table 2. Table of the 14 marker genes specific to the different transcriptomic clusters.**
315 Gene number, average Log2FC, pct1/pct2 ratio (percentage of cells expressing this transcript
316 in the cluster divided by the percentage of all other cells expressing this transcript) and cluster
317 number are reported. The description is taken from our annotation and the marker name is
318 derived from the description.

319 RT-qPCR expression profiles in the hemocyte fractions obtained from the Percoll density
320 gradient revealed distinct patterns according to the different transcriptomic cluster marker
321 genes (**Fig. 4A**). Cluster 1 markers (LACC24 and CLEC) were overexpressed in fractions 1, 2,
322 3 and underexpressed in the remaining fractions relative to total hemocytes. Cluster 2 markers
323 (EGFL and LEVAR) showed a decreasing pattern of expression from fraction 1 to fraction 7.
324 Cluster 3 markers (TGC and XBOX) showed a significant increase in expression in fractions 4
325 to 7. Cluster 4 markers (MLDP and HMGB1) showed a gradually decreasing expression
326 pattern from fraction 2 to fraction 7. Cluster 5 marker genes (GAL and CUBN) were
327 underexpressed in fractions 4 to 7, but not differentially expressed in fractions 1 to 3
328 compared to total hemocytes. Cluster 6 marker genes (CAV and NAT1) were overexpressed
329 in fraction 3, expressed similarly to total hemocytes in fraction 2, and underexpressed in
330 fractions 1, and 4 to 7. Finally, cluster 7 marker genes (MOX and GPROT) were
331 overexpressed only in fractions 3 to 7. Correlation analysis and statistical validation
332 demonstrated a clear association between hemocyte morphotypes and 14 gene markers (**Fig.
333 4B, Fig. 4C and Supp. Table S4**). Cluster 3 marker genes (TGC and XBOX) correlated
334 positively (r = 0.74 & 0.75) and specifically with small granule cells (SGC). Cluster 7 markers
335 (MOX and GPROT) correlated (r = 0.68 and r = 0.56) with vesicular cells (VC). Cluster 2
336 markers (EGFL and LEVAR) correlated positively with hyalinocytes (H) (r=0.85 and r=0.81).
337 No specific markers could be identified for blasts and macrophage-like cells.



Percol fraction number	H	ML	BBL	ABL	SGC	BGC	VC
1	46.43	20.26	14.47	15.50	0.35	0.38	2.62
2	20.17	27.64	27.21	20.17	1.14	1.03	2.64
3	12.49	28.26	21.18	17.71	7.96	4.71	7.68
4	7.82	32.69	9.64	5.37	18.29	3.90	22.29
5	4.29	19.67	3.56	3.35	40.92	4.50	23.71
6	3.57	13.95	1.19	1.54	68.00	1.69	10.06
7	2.43	8.50	1.67	0.36	81.18	2.68	3.18

Cluster	CAV	NAT1	MLDP	CUBN	EGFL	LEVAR	HMGB1	LACC24	CLEC	GAL	TGC	XBOX	MOX	GPROT
ML	0.39	0.44	0.49	0.60	0.16	0.29	0.49	0.38	0.29	0.37	−0.34	−0.33	0.32	0.47
H	0.18	0.23	0.43	0.43	0.85	0.81	0.72	0.44	0.60	0.63	−0.55	−0.54	−0.54	−0.42
BBL	0.74	0.83	0.81	0.75	0.60	0.65	0.80	0.90	0.84	0.89	−0.72	−0.76	−0.30	−0.20
ABL	0.72	0.78	0.87	0.62	0.71	0.73	0.81	0.81	0.79	0.78	−0.71	−0.73	−0.40	−0.30
SGC	−0.57	−0.63	−0.80	−0.75	−0.74	−0.79	−0.91	−0.73	−0.71	−0.78	0.74	0.75	0.12	−0.02
BGC	0.07	0.01	−0.12	−0.02	−0.21	−0.22	−0.22	−0.19	−0.26	−0.29	0.01	−0.01	0.37	0.41
VC	−0.32	−0.37	−0.24	−0.27	−0.34	−0.32	−0.23	−0.39	−0.47	−0.43	0.25	0.25	0.68	0.56

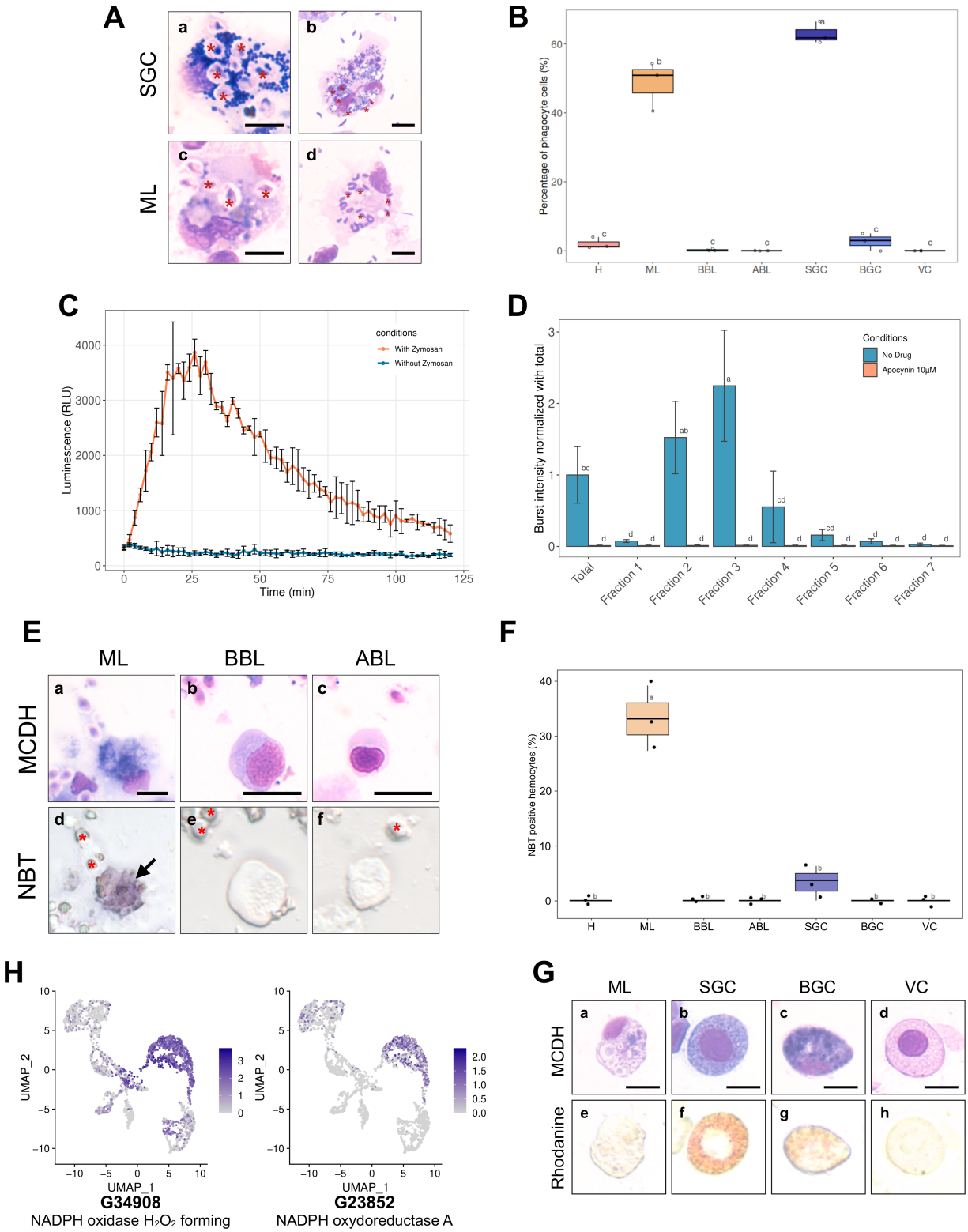
Fig 4. Characterization of molecular markers specific to the different hemocyte morphotypes. (A) Relative gene expression level of the 14 markers in the various fractions

341 after gradient density sorting. The graphs show the relative expression of genes compared to
342 their expression in total hemocytes in the various fractions (red dotted line). Relative gene
343 expression levels were normalized to the reference gene Cg-rps6 and the $2^{-\Delta Ct}$ method was
344 used to calculate relative expression levels, where ΔCt represents the difference between the
345 target gene's Ct value and the reference gene's Ct value. Standard deviations were calculated
346 based on four independent experiments. **(B)** Average percentage of each hemocyte type in the
347 7 Percoll gradient fractions used to quantify marker gene expression by qPCR. **(C)** Correlation
348 matrix between the relative gene expression of each marker gene in each fraction and the
349 percentage of each hemocyte type in the same fractions. Values and color scale represent the
350 Pearson correlation coefficient (r) ranging from -1 (inverse correlation) to +1 (full correlation).
351 **H** : Hyalinocyte, **ML** : Macrophage Like, **BBL** : Basophilic Blast Like cell, **ABL** :
352 Acidophilic Blast Like cell, **SGC** : Small Granule Cell, **BGC** : Big Granule Cell, **VC** :
353 Vesicular Cell. **LACC24** : Laccase 24, **CLEC** : C-type lectin domain-containing protein,
354 **EGFL** : EGF-like domain-containing protein 8, **LEVAR** : Putative regulator of levamisole
355 receptor-1, **TGC** : TGc domain-containing protein, **XBOX** : X-box binding protein-like
356 protein, **MLDP** : ML domain containing protein, **HMGB1** : High mobility group protein B1,
357 **CUBN** : Cubilin, **GAL** : Galectin, **CAV** : Caveolin, **NAT1** : Natterin-1, **MOX** : DBH-like
358 monooxygenase protein 1, **GPROT** : G protein receptor F1-2 domain-containing protein.

359 By employing RT-qPCR, cell sorting on Percoll gradients, and scRNA-seq analysis, we could
360 identify cell types corresponding to three transcriptomic clusters. Cluster 3 corresponds to
361 small granule cells (SGC), cluster 2 to hyalinocytes (H), and cluster 7 to vesicular cells (VC).
362 Cluster 3, cluster 2 and cluster 7 represent 17.8 %, 23 % and 3 % of the total cells analyzed by
363 scRNA-seq, respectively. This is consistent with cytological data, which indicated the
364 presence of 12.5 % +/- 5 % Small Granule Cells, 30 % +/- 9.3 % Hyalinocytes and 2.3 % +/-
365 1.6 % Vesicular Cells (**Fig. 1B and Fig. 3A**). Furthermore, the molecular and cellular
366 functions derived from GO-term and KEGG analyses (**Fig. 2**) were in alignment with the
367 anticipated functions for these three cell types.

368 **Only macrophage-like and small granule cells behave as professional phagocytes.**

369 The hemocytes separated on the Percoll gradient were characterized functionally to gain
370 further insight into their functional specialization. Two hemocyte functions, namely
371 phagocytosis and production of Reactive Oxygen Species (ROS), are known to carry out
372 major cellular antimicrobial activities. Granular cells have been suggested to be the
373 professional phagocytes specialized for these functions (*22*). Phagocytosis and oxidative burst
374 were studied using cell response toward zymosan particles (*35*).



376 **Fig 5. Phagocytosis, Reactive Oxygen Species production capacity and copper storage of**

377 **hemocytes. (A)** Images of small granule cells (**SGC**) and macrophage-like (**ML**) cells with

378 phagocytosed zymosan particles (**panels a & c - red stars**) and *Vibrio tasmaniensis*

379 LMG20012^{T} bacteria (**panels b & d - red stars**) from whole hemolymph sample. Scale bar : 5

380 µm. **(B)** Quantification of the phagocytic activity of zymosan particles by each cell type. The

381 graph shows the result of 3 independent experiments. **(C)** Luminescence recording to detect
382 the production of Reactive Oxygen Species (ROS). In orange, a biphasic curve was obtained
383 on naive oyster hemolymph after zymosan addition at t = 0 min. In blue, the control condition
384 corresponds to hemocytes without zymosan addition. **(D)** Graph showing the intensity of ROS
385 production in each Percoll fraction. Normalized burst intensity was calculated from the
386 luminescence peak obtained from each fraction. In blue, no drug was added to the experiment,
387 in orange, ROS production was impaired by the addition of apocynin. **(E)** NBT
388 (NitroBlueTetrazolium) staining of hemocytes exposed to zymosan particles. Hemocytes
389 morphology after MCDH staining: Macrophage Like **(a**), Basophilic **(b)** and Acidophilic **(c)**
390 Blast cells. NBT staining of the different types of hemocytes **(d-f)**. Red stars show zymosan
391 and bacteria particles. Black arrows indicate Macrophage-Like cells. Scale bar : 10 µm **(F)**
392 Quantification of NBT-positive cells present in the total hemolymph of oysters exposed to
393 zymosan. **(H)** UMAP plots showing cells expressing NADPH oxidase found in the scRNA-
394 seq dataset and their expression level. **(G)** Labeling of intracellular copper stores in *C.gigas*
395 hemocytes. MCDH (upper panels) and rhodanine (lower panels) staining of oyster hemocytes
396 to reveal copper accumulation. Scale bar : 10µm. For panels **(B)**, **(D)** and **(F)** the alphabetic
397 characters displayed above the data points in each plot represent statistically significant
398 differences among the groups, as determined by Tukey's test following ANOVA. Groups
399 denoted by different letters differ significantly at the $p < 0.05$ level of statistical significance.
400 **H** : Hyalinocytes, **ABL** : Acidophilic Blast-Like cells, **BBL** : Basophilic Blast-Like cells,
401 **ML** : Macrophage-Like cells, **SGC** : Small Granule Cells, **VC** : Vesicular Cells and **BGC** :
402 Big Granule Cells.

403 The phagocytic activity of hemocytes was first tested on a sample of total oyster hemolymph.
404 Cells were incubated for 1 hour with either zymosan particles or the bacterial strain *Vibrio*
405 *tasmaniensis* LMG20012^T. Only the small granule cells and the macrophage-like cells
406 exhibited efficient phagocytosis for both zymosan or vibrios, as observed after MCDH
407 staining (**Fig. 5A, Supp. Fig. 6 and Supp. Fig. 7A**). Macrophage-like cells and small granule

408 cells showed a phagocytic activity of 49 % and 55 %, respectively, and a phagocytosis index
409 of 3.5 and 5.2 particles per cell respectively (**Fig. 5B and Supp. Fig. 7B**), as confirmed in 3
410 independent experiments. Very limited phagocytic activity was observed for hyalinocytes (1.7
411 %), basophilic blasts (0.18 %), and big granule cells (2.7 %) with a phagocytosis index of 2, 1,
412 and 2.5 particles per cell, respectively (**Fig. 5B and Supp. Fig. 7B**). These results confirmed
413 that only small granule cells and macrophage-like cells behave as professional phagocytes
414 demonstrating robust phagocytic activity.

415 **Only macrophage-like cells produce Reactive Oxygen Species.**

416 The next step was to assess the capacity of *C. gigas* hemocytes in each Percoll fraction to
417 produce Reactive Oxygen Species (ROS) upon stimulation by zymosan, utilizing the Luminol
418 oxidation assay. Luminol luminescence peaked 25 minutes after exposure of hemocytes
419 (isolated from total hemolymph) to zymosan, indicating a robust oxidative burst (**Fig. 5C**).
420 The production of ROS was dependent on a NADPH oxidase, as it was completely inhibited
421 by the NADPH oxidase-specific inhibitor apocynin (**Fig. 5D**). To ascertain which hemocyte
422 types were involved in ROS production we tested Percoll density-sorted hemocytes for
423 oxidative burst activity. Fractions 2 and 3 displayed higher oxidative burst activity than the
424 other fractions. The burst intensity of fraction 3 was twice that of total hemocytes, and fraction
425 2 also exhibited significantly increased burst activity. In contrast, fraction 4 showed a
426 significant decrease in oxidative burst, and no activity was observed for fractions 1, 6 and 7
427 (**Fig. 5D**). These results indicate that NADPH-dependent oxidative burst activity is carried out
428 by fractions enriched in macrophage-like (ML) cells and blast-like cells (ABL and BBL).

429 Previous studies have shown that granular cells can produce ROS (*22*). In contrast, we found
430 that small granule cells collected in fraction 7 could not produce ROS through an oxidative
431 burst. To identify which type of hemocyte from blast-like cells and macrophage-like cells
432 produces ROS within a few minutes after exposure to zymosan, ROS production was
433 investigated using NitroBlueTetrazolium (NBT) reduction to stain hemocytes directly.

434 Correlative microscopic analysis using MCDH staining can then be conducted after the NBT
435 reduction reaction. Macrophage-like cells were strongly and significantly stained by NBT
436 reduction (33.2 %) (**Fig. 5E, panels a & d**). In contrast, some small granule cells were lightly
437 stained by NBT reduction (**Fig. 5F**). Blast-like cells, big granule cells, vesicular cells, and
438 hyalinocytes were never NBT stained, confirming that these cell types were not involved in
439 ROS production (**Supp. Fig. S7**). Taken together, these observations demonstrate that small
440 granule cells and macrophage-like cells are the two professional phagocytes among
441 hemocytes. However, only macrophage-like cells are capable of oxidative burst upon exposure
442 to zymosan. In light of these new functional data, we further analyzed the expression level of
443 NADPH-oxidase-related enzymes in the scRNA-seq dataset. The cells in cluster 1
444 predominantly expressed two NADPH oxidase isoforms (gene numbers G34908 and G23852)
445 compared to other clusters (**Fig. 5H**). Furthermore, this cluster expresses macrophage-related
446 genes, including the macrophage-expressed gene 1 protein (G30226) (**Supp. Data S1**), along
447 with maturation factors for dual oxidase, an enzyme involved in peroxide formation (**Supp.
448 Fig. S8**), supporting its designation as macrophage-like based on functional characteristics.
449 Collectively, these data indicated that cells of transcriptomic cluster 1 corresponded to
450 macrophage-like cells.

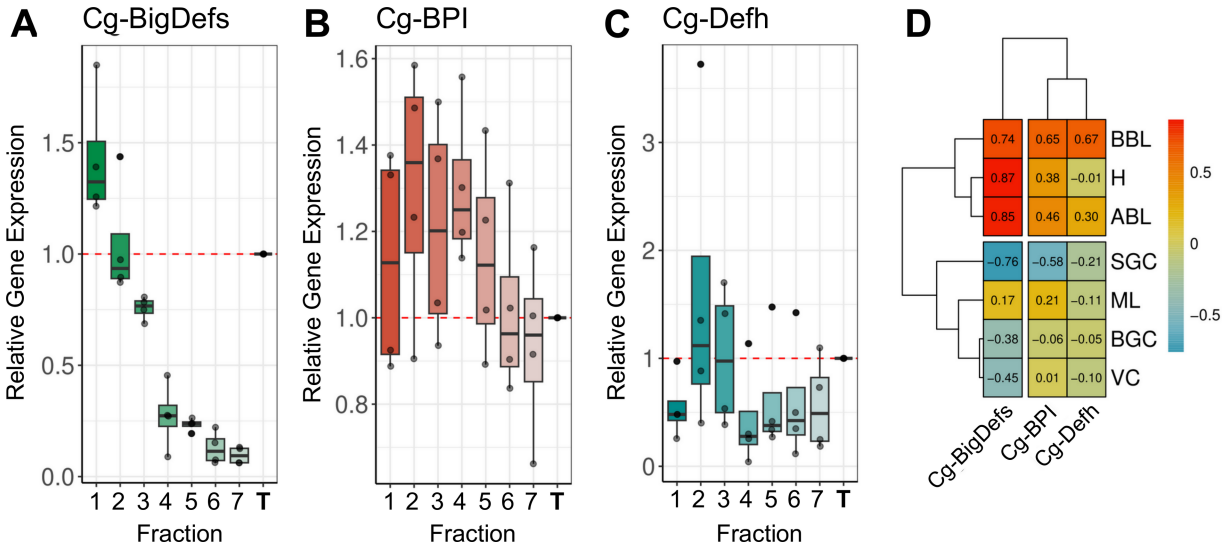
451 **Small granule cells and big granule cells accumulate intracellular copper.**

452 The effects of copper on oyster hemocytes have been studied extensively due to its abundance
453 in polluted marine environments (*19*). In addition, several studies have shown that *Vibrio*
454 species pathogenic for *C. gigas* possess copper resistance genes, which are crucial for their
455 survival within hemocytes (*9*). This prompted us to investigate which hemocyte types are
456 involved in copper metabolism. To this end, total hemocytes were isolated from naive oysters
457 and stained with rhodanine to reveal copper storage in cells. Rhodanine staining revealed that
458 33 % +/- 2 % of small granule cells and 30 % +/- 10 % of big granule cells exhibited a specific
459 reddish/brown staining indicative of a high concentration of copper in their granules (**Fig. 5G**

460 **and Supp. Fig. S9**). These results provide functional proof that SGC are specialized in metal
461 homeostasis in addition to phagocytosis, as suggested by the scRNAseq data (Cluster 3).

462 **Antimicrobial peptides are expressed by agranular cells, blasts and hyalinocytes.**

463 Antimicrobial peptides (AMPs) have long been studied for their role in the invertebrate
464 humoral immune response. They are expressed by hemocytes, including those in *C. gigas (36)*.
465 The aim was to ascertain whether different hemocyte cell types expressed distinct AMPs.
466 However, due to the limited sequencing depth, the scRNA-seq analysis was not sensitive
467 enough to reveal AMP expression. This limitation was addressed by investigating AMP
468 expression through RT-qPCR on Percoll density-fractionated hemocytes. The results indicated
469 that Cg-Bigdefs 1-2, Cg-BPI and hemocyte defensin were predominantly expressed by
470 agranular cells, blasts and hyalinocytes (ABL, BBL and H) **(Fig. 6A, B and C)**. The
471 expression of these AMPs was associated with Blasts abundance, while the expression of Cg-
472 BigDefs 1-2 was only associated with hyalinocytes **(Fig. 6D)**. Granular cells (VC, ML, BGC
473 and SGC) did not seem to express any of the analyzed AMPs. These data suggest that some of
474 the agranular cells are specialized in the production of humoral effectors.

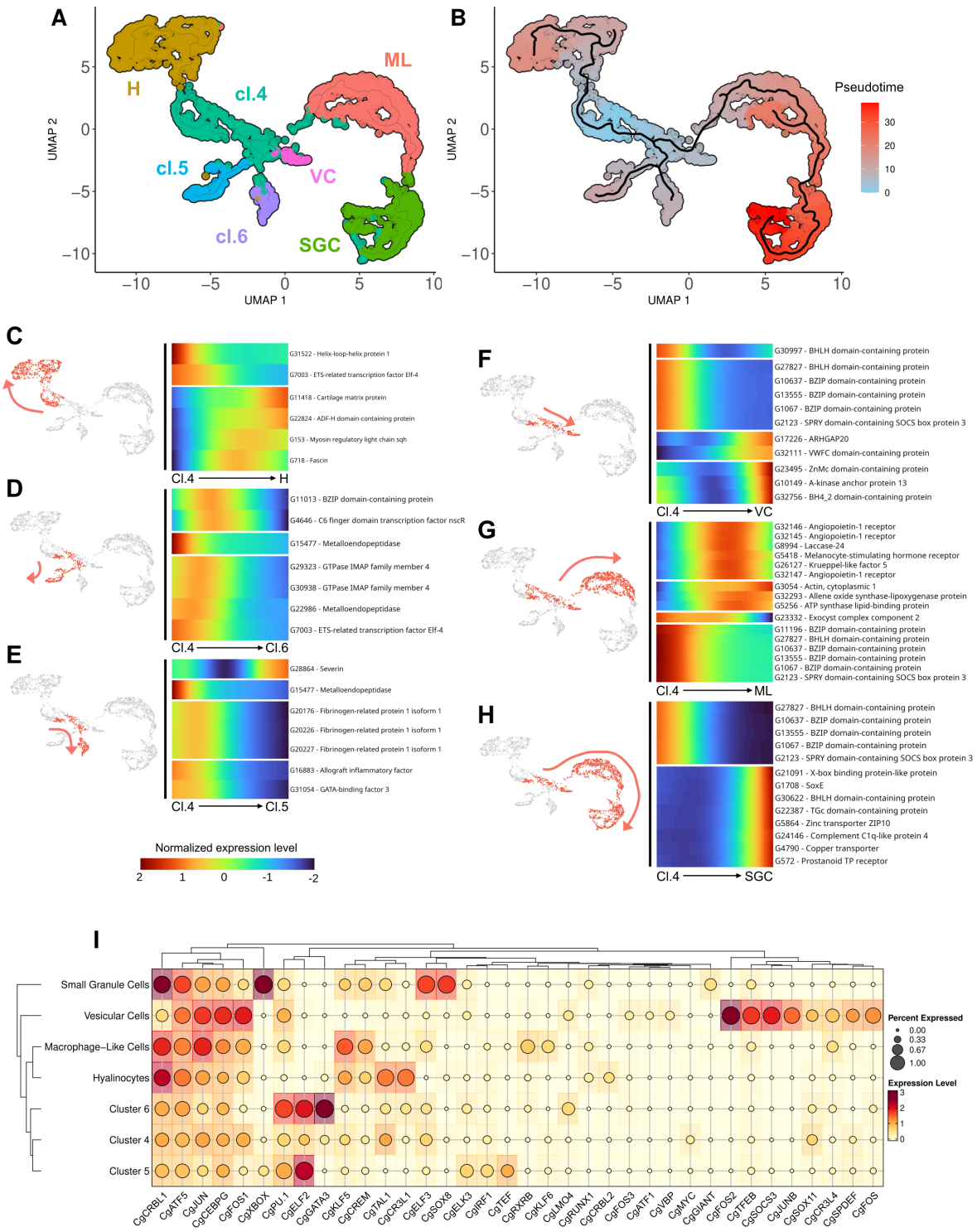


476 **Fig 6. Hemocyte expression profiles of some antimicrobial peptides. (A) (B)** and **(C)**
477 Relative gene expression in the 7 Percoll hemocyte fractions of Big-Defensin1 & 2 (Cg-
478 BigDefs), BPI (Cg-BPI) and hemocyte defensin (Cg-Defh), respectively, in comparison to the

479 gene expression level in unfractionated hemolymph. Relative gene expression levels were
480 normalized to the reference gene Cg-rps6. The $2^{-\Delta Ct}$ method was used to calculate relative
481 expression levels, where ΔCt represents the difference between the target gene's Ct value and
482 the reference gene's Ct value. **(D)** Correlation matrix between the relative gene expression of
483 BigDefensin1 & 2, BPI and hemocyte defensin gene in each fraction and the percentage of
484 each hemocyte type in each fraction (**H** : Hyalinocytes, **ABL** : Acidophilic Blast Like, **BBL** :
485 Basophilic Blast Like, **SGC** : Small Granule Cell, **ML** : Macrophage Like, **BGC** : Big
486 Granule Cell, **VC** : Vesicular Cell. Values and color scale represent the Pearson correlation
487 coefficient (r) ranging from -1 (inverse correlation) to +1 (full correlation).

488 **Tentative model of hemocyte lineages and differentiation pathways in *C. gigas.***

489 The ontogeny, lineage and differentiation pathways of bivalves remain largely unknown (*37*).
490 However, there are some indications of circulating and proliferating hemocyte progenitors in
491 the hemolymph of *C. gigas* (*38*). GO-terms analysis of the 7 transcriptomic clusters revealed
492 different functional signatures, including the transcriptomic signature of cluster 4, which
493 showed a high expression level of ribosomal proteins (**Supp. Fig. S1J**) . This particularity has
494 been observed in hematopoietic stem cells in vertebrates (*39–41*). Furthermore, scRNA-seq
495 approaches using bioinformatic tools like Monocle3 (*42*) can now be used to deduce
496 differentiation pathways from trajectory inference using gene expression profiles analysis.
497 This revealed an overrepresentation of functions involved in the splicing, transcription and
498 translation continuum in the same fourth cluster (see cluster 4 in **Fig. 2A** and **Fig. 2B**), thereby
499 supporting the hypothesis of a pool of quiescent or immature cells that can differentiate upon
500 stimulation.
501



502 **Figure 7. Pseudotime ordering of cells revealed 6 potential differentiation pathways of**
503 **hemocytes. (A)** UMAP plot of scRNA-seq analysis showing the 7 transcriptomic clusters used
504 for pseudotime analysis. 4 clusters were identified cytologically (SGC for small granule cells -
505 cluster 3, H for hyalinocytes - cluster 2, ML for Macrophage Like - cluster 1 and VC for
506 vesicular cells - cluster 7), cl.4, cl.5, and cl.6 represent clusters 4, 5, and 6, respectively. **(B)**
507 Graphical representation (UMAP projection) of the Monocle 3 pseudo-time order of the
508 clustered cells. Cluster 4 (cl.4) was used as the origin for the pseudotime analysis. **(C) (D) (E)**

509 **(F) (G) and (H)** show the gene expression level of selected marker genes obtained from the
510 monocle3 trajectory analysis at the beginning and end of the modelized differentiation
511 pathways (in red on the UMAP plot) from cluster 4 to hyalinocytes, to cluster 5 cells, to
512 cluster 6 cells, to Vesicular Cells (VC), to Macrophage-Like cells (ML) and to Small Granule
513 Cells (SGC) respectively. The color scale represents the normalized expression level of each
514 gene. **(I)** Dot plot showing the average expression and the percentage of cells expressing
515 identified transcripts encoding for transcription factors in the scRNA-seq dataset. Expression
516 level expression is expressed in Log2FC and dot diameter represents the percent of cells
517 expressing the gene.

518 Cluster 4 was chosen to enroot the pseudotime analysis to deduce differentiation pathways and
519 cell lineages using Monocle3. (**Fig. 7A**). By temporally ordering the 2817 cells analyzed by
520 scRNA-seq, 6 cell lineages could be defined (**Fig. 7B**). Differentiation pathway 1 leads to
521 hyalinocytes (H) (**Fig. 7C**). This transition is characterized by the downregulation of 8 genes
522 (**Supp. Fig. S10A**), two of which are transcription factors (G31522 and G7003). The
523 hyalinocyte cluster was also characterized by the overexpression of genes involved in cell
524 contractility (G11418 , G22824, G153 and G718). Differentiation pathway 2 leads to cells of
525 cluster 5 and is characterized by the downregulation of about 23 genes (**Supp. Fig. S10B**),
526 including two genes related to metalloendopeptidase activity (G15477 and G22986), two
527 GTPase IMAP genes (G29323 and G30938), and one transcription factor that is specifically
528 downregulated (G7003). Additionally, two other transcription factors are transiently
529 overexpressed (G11013 and G4646) (**Fig. 7D**). Differentiation pathway 3 leads to cells of
530 cluster 6 and is characterized by the down-expression of 21 genes (**Supp. Fig. S10C**). Three
531 fibrinogen-related protein genes (G20176, G20226 and G20227), a metalloendopeptidase gene
532 (G15477) and a GATA-binding factor (G31054) were downregulated. The severin gene was
533 upregulated (G28864) (**Fig. 7E**). The pathways between clusters 4, 5 and 6 found by
534 Monocle3 analysis were pseudo temporally short and few specific markers were identified,
535 suggesting that they were transcriptionally close. Differentiation pathway 4 leads to vesicular

536 cells (VC) (**Fig. 7F**) where a large number of genes were silenced to give rise to these cells
537 (**Supp. Fig. S10D**). Five genes were specifically overexpressed in VC (G32111, G17226,
538 G32756, G10149 and G23495). Interestingly, 6 genes encoding potential transcription factors
539 were downregulated in this lineage (G30997, G10637, G1067, G13555, G2123 and G27827).
540 Differentiation pathway 5 leads to macrophage-like cells (ML), and is characterized by the
541 underexpression of 35 genes and the overexpression of 13 genes (**Supp. Fig. S10E**). Among
542 the 35 genes, 6 are putative transcription factors (G2123, G10637, G13555, G27827, G1067
543 and G11196) (**Fig. 7G**). Finally, we can outline a lineage ending in small granule cells (SGC)
544 (**Fig. 7H**). This pathway involved cluster 4 (immature cells), VC, ML and SGC and was
545 characterized by the downregulation of 27 genes (**Supp. Fig. S10F**), including 5 potential
546 transcription factors (G10637, G13555, G27827, G1067 and G2123). SGC showed a distinct
547 transcriptomic signature with 61 overexpressed genes, including 3 transcription factors
548 (G1708, G21091 and G30622). Based on these findings, we postulate that immature cluster 4
549 cells possess pluripotent potential and can give rise to four terminally differentiated cell
550 types : cluster 5 and 6 cells, hyalinocytes and small granule cells, and two other transient
551 hemocyte cell types : vesicular cells and macrophage-like cells. All data from the Monocle3
552 analyses are available as Supplementary Data (**Supp. Fig. S10**).

553 Differentiation pathway analysis thus revealed the over- or under-expression of various
554 transcription factors in the identified pathways. Given the established role of transcription
555 factors as master regulators of cell differentiation and their utility in delineating cell lineages,
556 we investigated the combinatorial expression patterns of transcription factors among the
557 different transcriptomic cell clusters. Based on GO-terms annotation, 28 different sequences
558 corresponding to transcription factors were isolated from the scRNA-seq dataset. The
559 transcription factor function was confirmed by manual annotation (**Supp. Table S5**) and
560 **Figure 7I** shows the average expression profiles of these factors in the different clusters.
561 **Supplementary Figure S11** illustrates their expression levels in single cells. Two
562 transcription factors, CgATF5 and CgCRBL1, exhibited a contrasting expression profile with

563 an increased average expression in macrophage-like cells, hyalinocytes and small granule cells
564 versus a low expression profile in cells in clusters 4, 5, 6 and vesicular cells. We also
565 identified transcription factors that were specifically expressed in the different transcriptomic
566 clusters : CgSPDEF, CgSOCS3, CgFOS and CgTFEB were specific for vesicular cells,
567 CgSOX8, CgXBOX and CgELF3 were specific for small granule cells, CgCR3L1 and
568 CgTAL1 for hyalinocytes, and CgJUN, CgKLF5, CgKLF6 and CgCREM for macrophage-like
569 cells. Eight additional transcription factors were specifically identified in cluster 6
570 (CgGATA3, CgPU.1 and CgELF2), cluster 5 (CgELF2, CgELK3 and CgIRF1) and cluster 4
571 (CgTAL1 and CgFOS1). These data potentially define four distinct hematopoietic lineages
572 originating from one type of immature blast cells and give rise to hyalinocytes, SGC (via VC
573 and ML), or two distinct differentiated blast-like cells. We also identified a combination of
574 transcription factors specific to lineages and cell types that are potential master regulators of
575 cell fate during hematopoiesis.

576 **Discussion**
577
578 The findings of our study represent a significant advancement in our understanding of the
579 functional diversity and lineages of *C. gigas* hemocytes. Single-cell RNA-seq and cytology
580 were combined to identify seven distinct hemocyte transcriptomic cell clusters and an
581 equivalent number of morphotypes. These include four granular cell types (big granule cells,
582 macrophage-like cells, small granule cells and vesicular cells), two distinct blast-like cells
583 (basophilic and acidophilic blast-like cells), and one agranular epithelial-like cell type
584 (hyalinocytes). A significant challenge was to identify correlations between transcriptomic and
585 cytological data to fully define each cytological cell type. This challenge was overcome by
586 combining multiple approaches, including isopycnic Percoll density gradient fractionation
587 combined with the analysis of transcriptomic markers expression, and functional assays
588 including phagocytosis, oxidative burst, copper accumulation, AMP expression and finally
589 pseudotime analysis of gene expression. These results confirmed the historical classification
590 of the 3 main cell groups : blasts, hyalinocytes, and granular cells (*15*) and deepened our

591 understanding of the functional specificities of poorly characterized cell types. In particular,
592 we identified distinct transcriptional and functional subtypes among blasts and granular cells
593 with complementary immune specialization and lineage relationships between cells.
594
595 One significant outcome of the present study is the identification of cell types involved in
596 antimicrobial activities, including phagocytosis, intracellular copper accumulation, an
597 oxidative burst and antimicrobial peptide production. These cell types have been extensively
598 studied for their role in antibacterial and antiparasitic defenses, as they are found in the large
599 majority of invertebrates (*17*). Nevertheless, there has been considerable debate surrounding
600 the cell types specialized for these critical immune functions in bivalves, particularly oysters.
601 Moreover, the involvement of the different hemocyte subpopulations in immune functions is
602 not yet fully understood.
603
604 Our findings reveal that the macrophage-like (ML) cells and small granule cells (SGC) are the
605 sole hemocyte cell types that function as professional phagocytes, as demonstrated against
606 Zymosan or *Vibrio*. These two distinct cell types could be distinguished functionally. First,
607 two distinct transcriptomic clusters were identified for each cell type (Cluster 1 for ML / BGC
608 and Cluster 3 for SGC). Secondly, only ML induced a measurable oxidative burst. Thirdly,
609 only SGC accumulated intracellular copper in specific granules. Interestingly, our scRNA-seq
610 analysis indicates that SGC (cluster 3) expresses the scavenger receptor cysteine-rich (SRCR)
611 gene G3876, annotated as an Low-density lipoprotein receptor-related protein with a Log2
612 fold change (Log2FC) of 0.77 linking them to scavenger receptor-mediated pathogen
613 recognition and clearance. This aligns with findings by Wang et al. (*43*), who demonstrated
614 significant expansion and dynamic regulation of SRCR genes in response to pathogen-
615 associated molecular patterns. The pseudotime trajectory linking macrophage-like (ML) and
616 small granule-like (SGC) cells suggests a potential developmental relationship within the
617 granular cell lineage; however, this hypothesis requires further validation. The notion that ML

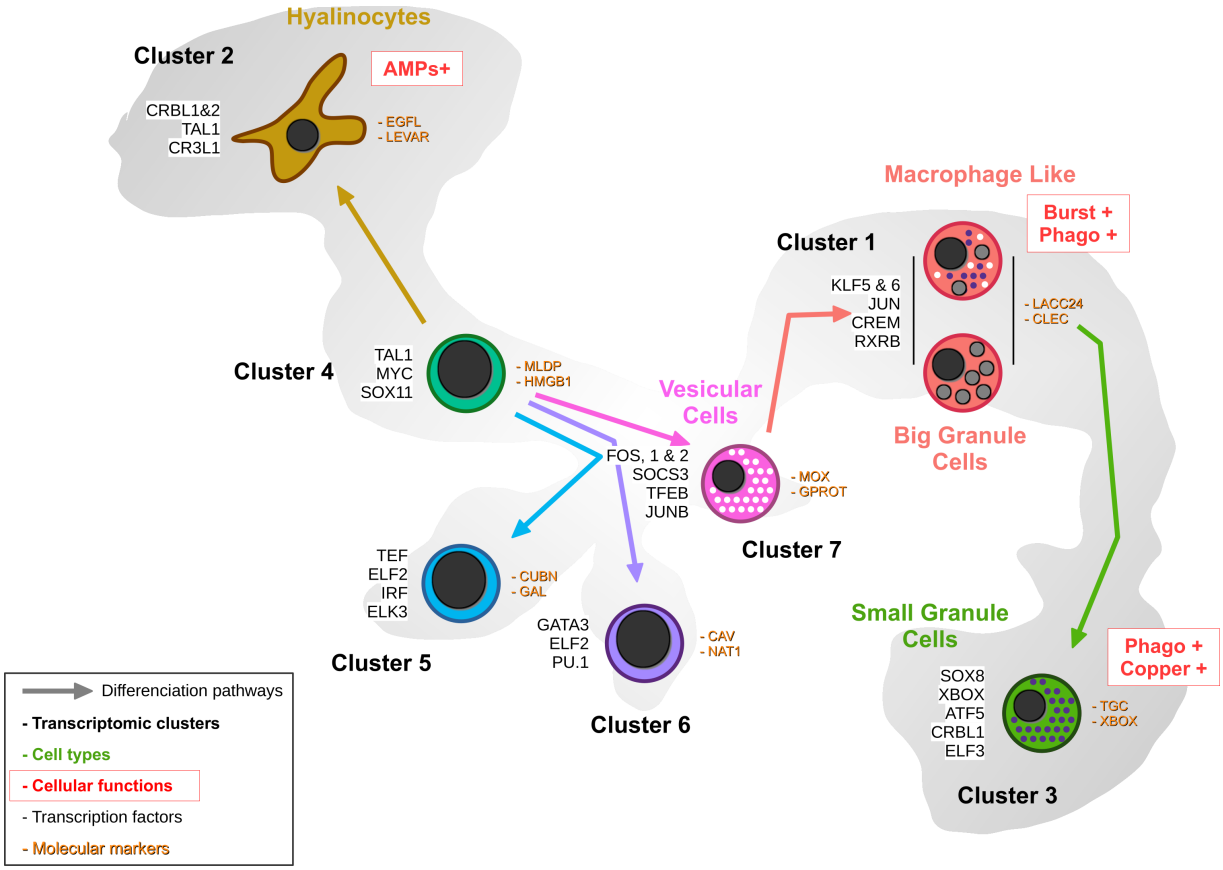
618 could serve as a precursor for SGC may seem counterintuitive. However, this is not an isolated
619 phenomenon among invertebrate hemocytes. For instance, in *Drosophila* larvae (*44*), some
620 populations of professional phagocytes, the sessile plasmatocytes, give rise to crystal cells or
621 lamellocytes that are morphologically and functionally distinct from plasmatocytes (*45*).
622 Similarly, the existence of multiple professional
623 phagocytes in the oyster is reminiscent of vertebrate macrophages, polynuclear neutrophils
624 and dendritic cells, which possess distinct functional specializations, including efferocytosis
625 (*46*), oxidative burst, NETosis, or antigen presentation (*47*).
626
627 The characterization of professional phagocytes in oysters is of particular importance for a
628 deeper understanding of oyster-*Vibrio* interactions during pathogenesis. Some of the most
629 extensively studied oyster pathogens, including strains of *V. tasmaniensis* and *V. aestuarianus*
630 *francensis,* harbor virulence traits that enable them to disrupt the phagocytic activity of
631 hemocytes during pathogenesis. For instance, *V. tasmaniensis* behaves as a facultative
632 intracellular pathogen with phagocytosis-dependent cytotoxicity (*11*). Additionally, both *V.*
633 *tasmaniensis* and *V. aestuarianus* display resistance to copper toxicity through CopA and
634 CusA/B transporters. This trait is essential for the survival and virulence of these pathogens in
635 oysters (*9*, *48*). Since SGCs have been demonstrated to be professional phagocytes with copper-
636 rich granules, the cellular interactions between SGCs and these vibrios are likely to be critical
637 during the antibacterial host response and pathogenesis. ML are phagocytes that possess a
638 very potent NADPH-oxidase-dependent oxidative burst. The oxidative burst is a rapid and
639 potent antimicrobial response observed in professional phagocytes, such as polynuclear
640 neutrophils in mammals. It is worth noting that NADPH-dependent ETosis has been observed
641 in *C. gigas* in a manner analogous to that observed in human neutrophils (*49*). This cell death is
642 characterized by the projection of DNA extracellular traps that capture and kill some
643 pathogens like *Vibrio* (*49*). Therefore, it is reasonable to hypothesize that ML may also be
644 involved in ETosis.

645

646 The specialized functions of the two other types of granular cells, the BGCs and the VCs,
647 remain unclear. Despite the difficulty in identifying a specific scRNAseq transcriptomic
648 cluster for BGCs, the level of expression of laccase 24 was found to be higher in a particular
649 subcluster among ML (**Supp. Fig. S8**) and pseudo-time analysis highlighted the same
650 subcluster as an alternative differentiation state among ML (**Fig. 7B**). The enrichment of
651 transcripts involved in oxido-reduction pathways, particularly laccase 24, aligns with their
652 potential role in melanization and response to oxidative stress (*50*). While melanin-like
653 deposits have been observed in some cases of infestation by the parasites *Martelia* or B*onamia*
654 in the Sydney rock oyster (*51*), this mechanism is not as robust as that described in arthropods,
655 which perform melanization through a prophenoloxidase activation cascade (*52*, *53*). In many
656 marine invertebrates (*50*), a type of hemocyte known as Brown Cells could be related to the
657 BGCs described here. When observed without any staining (as in **Fig. 5F**), their big granules
658 with a yellow to dark brown content appear to align with the historical description of brown
659 cells that often infiltrate tissues (like gills) in animals exposed to polluted waters (*50*). It has
660 therefore been theorized that these cells are involved in detoxification processes. Our
661 pseudotime analysis indicates that they likely originate from ML, with limited phagocytosis
662 activity and a specialized role in melanization and potentially heavy metal detoxification, as
663 evidenced by rhodanine staining showing copper accumulation in some of their granules (**Fig.**
664 **5G**). Further studies are recommended to clarify the role of BGCs, particularly in the context
665 of parasitic infestation or exposure to toxic stresses. Lastly, VC are granular cells that remain
666 to be functionally characterized. Their transcriptomic profile suggests strong intracellular
667 vesicular trafficking and autophagy activity. However, our functional assays did not reveal
668 any particular immune-related function. Their clear granules appear auto-fluorescent when
669 illuminated with UV light (**Supp. Fig. S12**) but the biochemical nature of the content of these
670 granules remains to be characterized. As they also appear as cell intermediates along the
671 granular cell differentiation pathway in pseudotime analysis, they could represent functionally

672 immature precursors of the other three granular cell types, much like promyelocytes which
673 possess specific azurophilic granules but are functionally immature precursors during
674 granulocyte differentiation in humans. However, the enrichment of autophagy-related
675 transcripts in VC calls for further investigation into a potential role in antiviral immunity, as
676 autophagy has been suggested to play a role in the response to OsHV-1 virus (*54*).
677
678 Hyalinocytes are a homogenous cell type, with only one morphotype matching one
679 transcriptomic cluster. It can be deduced from pseudotime analysis that they originate from a
680 specific and very different differentiation pathway than the granular cells. According to the
681 literature, hyalinocytes are involved in the early stages of inflammation and can infiltrate
682 wounds and interact with foreign particles. In the flat oyster *Ostrea edulis*, they contribute to
683 shell production and wound healing (*55*). In the Sydney rock oyster they play a role in cell
684 aggregation (*56*), while in *C. virginica* (*57*) they contribute to encapsulation, reminiscent of
685 lamellocytes in *Drosophila*. Our results suggest an important role of AMP expression in the
686 immune response. Indeed, *Cg*-BigDefs, which participate in the control of oyster microbiota
687 (*58*), were found to be expressed in both hyalinocytes and blast-like cells. Moreover,
688 hyalinocytes from the oyster *O. edulis* have been shown to express the AMP Myticin C (*59*),
689 which lends further support to this immune function. Among the AMPs, we also found that
690 Cg-BPI and *Cg*-Defh, are more expressed in BBL, ABL than in other cell types. These results
691 are somewhat unexpected, given the prevailing assumption that AMPs are stored in granules
692 of granular cells, rather than agranular cells (*16*). These results highlight the necessity to
693 reassess the role of specific agranular cell types in the active production of humoral immune
694 effectors. Nevertheless, further staining approaches would be necessary to confirm these
695 transcriptomic results. Our findings suggest that hyalinocytes and/or the blast-like cells may
696 be a cellular target of the OsHV-1 virus, the causal agent of POMS, which dampens the
697 expression of certain AMPs (*7*), thereby inducing bacterial dysbiosis. It is still unclear whether

698 this is due to a decreased expression of AMPs and/or inhibition of immature blast cell
699 differentiation involved in the renewal of agranular cell types.
700
701 It should be noted that the complexity of blast-like cells could not be fully elucidated in this
702 study, as 3 clusters and only 2 morphotypes were identified. Cell fractionation using Percoll
703 gradient failed to yield pure blast-enriched fractions (ABL and BBL), preventing precise
704 functional characterization. The enrichment in transcripts of the transcription synthesis
705 degradation continuum aligns with the definition of undifferentiated blast-type cells and with
706 the basophilic staining obtained in MCDH for BBL cells, as immature blast cells are
707 characterized by a basophilic cytoplasmic staining in humans. However, our results show that
708 certain blast populations can produce AMPs, suggesting that these cells may also play a role in
709 the production of humoral effectors. Ultimately, it remains unclear whether these circulating
710 immature cells are indeed the stem cells from which all hemocytes originate.



712 **Figure 8. Proposed hemocyte ontology in *Crassostrea gigas* based on the transcriptomic,**
713 **cytological and functional results obtained.** Cells are colored according to the same color

714 code as the transcriptomic clusters. Cluster numbers and cell types are indicated. To the left of
715 the cells are the overexpressed transcription factors and to the right are the identified marker
716 genes in each cluster. Functional characteristics of hyalinocytes, macrophage-like cells and
717 small granule cells are marked in red. (**AMP** : AntiMicrobial Peptide, **Burst** : ROS
718 production, **Phago** : phagocytosis)

719 In the animal kingdom, the innate immune system relies on specialized cells derived from
720 pluripotent precursors through hematopoiesis. Transcription factors in particular have been
721 found to exhibit a high degree of conservation throughout the animal kingdom, from
722 invertebrates to vertebrates. However, the mechanisms underlying the functional
723 differentiation of bivalves are only partially available or understood (*60*). Recent research has
724 indicated the potential existence of hemocyte progenitors, also known as blast-like cells in
725 several bivalve species. These include clams, mussels, scallops, marine mussels, freshwater
726 mussels, oysters, pearl oysters, and wing shells (*60*). However, various models for
727 hematopoiesis in bivalves have been proposed and extensively debated without a clear
728 consensus or definitive proof. Single-cell RNA sequencing and pseudotime analysis have
729 enabled us to propose a refined model of hematopoiesis at an unprecedented level of detail. In
730 this model, the different types of hemocytes are likely produced through four differentiation
731 pathways that originate from a common progenitor. One pathway results in the formation of
732 agranular hyalinocytes, while another independent pathway gives rise to the granular cells,
733 including VC, ML, BGC, and SGC. Furthermore, two additional differentiation pathways
734 have been identified that may lead to terminally differentiated blast-like cells. The
735 differentiation of mature hemocytes involves the establishment of lineage-specific gene
736 expression profiles, which rely on transcription factors to modulate the expression of their
737 target genes. Our study identified a combination of transcription factors that are differentially
738 expressed during *C. gigas* hematopoiesis and are specific to differentiation stages and cell
739 lineages, including many well-known hematopoietic transcription factors such as GATA, PU-
740 1, TAL1 and SOX factors, which are positioned in this model of *C. gigas* hematopoiesis (*61*–

741 *63*) (**Fig. 8**). Nevertheless, we did not detect any canonical stem or progenitor cell populations
742 in our dataset, underscoring the need for future investigations - potentially involving
743 immunological challenges and lineage-tracing assays - to clarify whether proliferative cells
744 circulate in the hemolymph or instead reside primarily in tissue compartments.
745 The comparison of hemocyte populations in *C. hongkongensis* and *C. gigas* highlights a
746 shared reliance on granulocytes as primary immune effectors, with *C. hongkongensis* (*27*). Our
747 study emphasizes functional specialization in *C. gigas*, focusing on phagocytosis, oxidative
748 stress responses, and vesicular trafficking, likely reflecting differences in transcriptomic and
749 functional characterization methodologies. Moreover, the identification of conserved
750 transcription factors such as Tal1, Sox, Runx and GATA in both datasets underscores the
751 evolutionary conservation of hematopoiesis pathways across metazoans. While direct
752 comparisons between scRNA-seq datasets are hindered by differing annotation frameworks,
753 parallels can still be drawn with studies in other invertebrates. For example, shrimp and
754 crayfish scRNAseq datasets (*64*) describe hemocytes with immune-activated states and
755 integrin-related markers, respectively, while Anopheles gambiae (*65*)exhibits clusters with
756 high ribosomal activity and transcription factors such as bHLH and myc. These findings
757 highlight both the shared molecular foundations and species-specific differences that define
758 hemocyte diversity and specialization.
759 This study significantly advances our understanding of bivalve immunity, especially in
760 comparison to arthropods. By introducing a standardized reference hemocytogram for oysters
761 using MCDH staining, defining cell type-specific markers and key transcription factors likely
762 involved in cell fate determination, as well as clarifying the functions of different hemocytes,
763 we have paved the way for future in-depth studies. Moreover, the identified clusters represent
764 functionally and transcriptionally distinct cell populations under the conditions studied, yet
765 they may reflect dynamic states rather than fixed subtypes, thereby acknowledging the
766 potential for hemocyte populations to transition between states in response to environmental
767 or physiological changes—even under homeostatic conditions. This will facilitate further

768 studies of the oyster's immune response to various biotic and abiotic stress at the cellular
769 level. Improved comprehension of antiviral and antibacterial responses in bivalves, along with
770 an enhanced understanding of immune priming and immune memory at the cellular level in
771 bivalves, will benefit health and population management practices for sustainable aquaculture
772 production. These findings will also contribute to the broader field of evolutionary
773 immunology by enabling comparative studies and elucidating the diversification of immune
774 cells and immunity-related genes in a protostome.
775

776 **Materials and Methods**
777

778 **Conservation of oysters**

779 The work described here was performed using two different sources of oysters of the same
780 species *Crassostrea (Magallana) gigas*. ISA (Ifremer Standardized Oysters - La tremblade -
781 France) oysters for the scRNA-seq experiment and oysters provided by a local supplier
782 (https://www.huitres-bouzigues.com). Animals were washed and kept in 10 L tanks containing
783 seawater and a bubbler to oxygenate and homogenize the water. Water was changed daily and
784 oyster health was monitored. All animals used in this study were 18 months of age.

785 **Hemocyte collection and processing**

786 *Crassostrea gigas* hemocytes were collected from live animals by puncture of the adductor
787 muscle. The oyster shell was incised on the posterior side with forceps. Hemolymph was
788 collected using a 23Gx1" needle mounted on a 5 mL syringe prefilled with 2 mL of ice-cold
789 Alsever modified medium (20.8 g glucose – 8 g trisodium citrate - 22.5 g sodium chloride -
790 0.4 g BSA - pH=7.5). Samples were centrifuged for 4 minutes at 200 g – 4 °C and the
791 supernatant was removed and replaced with 1 mL of fresh Alsever modified medium. Each
792 hemocyte sample was thoroughly checked for quality, counted under a microscope using
793 KOVA (Kova International, USA) slides, and stored on ice prior to processing. For scRNA-
794 seq analysis, resuspended hemocytes were filtered on 30 µm filters, counted, the solution was
795 adjusted to 1.10^6 cells per mL and stored on ice prior to 10X genomic library preparation.

796 ***C. gigas* genome annotation**

797 The *C. gigas* genome (Genbank GCA_902806645.1) (*30*) was used as a reference. Prior to
798 annotation, the longest CDS sequences were extracted from the gff3 file. Annotation was
799 realized using the ORSON script (https://gitlab.ifremer.fr/bioinfo/workflows/orson). ORSON
800 combines cutting-edge tools for annotation processes within a Nextflow pipeline. The ORSON
801 script performs a sequence similarity search with BLAST (*66*) against the Uniprot-Swissprot
802 and Uniprot-trEMBL databases, and functional prediction with InterProScan (*67*) and eggNOG
803 (*68*) orthogroup annotation. Interproscan analysis was performed against Pfam, Prosite, CDD,
804 TIGR, SMART, SuperFamily, PRINTS and Hamap databases. Results were collected and
805 processed using Blast2GO (*69*) for annotation mapping and validation.

806 **Drop Seq-based scRNA-seq library generation**

807 The 10X Genomics protocol, Single Cell 3' Reagent Kits v2 User Guide from the
808 manufacturer (10X Genomics, USA) was followed to prepare gel in emulsion beads (GEM)
809 containing single cells, hydrogel beads, and reverse transcription reagents, perform barcoded
810 cDNA synthesis, and generate sequencing libraries from pooled cDNAs. The concentration of
811 single-cell suspensions was approximately 1000 cells / µL, as estimated by manual counting,
812 and cells were loaded according to the 10X protocol to capture approximately 3000 cells per
813 reaction. Library construction (after GEM digestion) was performed using 10X reagents
814 according to the manufacturer's instructions. Libraries (paired-end reads 75 bp) were
815 sequenced on an Illumina NovaSeq (Illumina, USA) using two sequencing lanes per sample.

816 **scRNA-seq analysis**

817 Reads were aligned to the *C. gigas* reference genome (Genbank GCA_902806645.1) (*27*)
818 using STAR solo software (v 2.7.10) (*29*). Unique molecular identifiers (UMIs) were extracted
819 and counted for each cell, and an expression matrix was generated for further analysis. Single-
820 cell RNA sequencing (scRNA-seq) data analysis was performed using the R programming
821 language (version 4.2.1) (R Core Team, 2018) and the Seurat package (version 4.3.0) (*70*). The
822 data were then pre-processed to remove unwanted sources of variation and to normalize gene

823 expression. Cells with small library sizes and a high proportion of mitochondrial genes were
824 excluded. Data normalization was performed using the SCTransform method. After
825 normalization, highly variable genes were identified using the FindVariableFeatures function
826 and the top variable genes were selected for downstream analyses. Dimensionality reduction
827 was performed using Principal Component Analysis (PCA), followed by Uniform Manifold
828 Approximation and Projection (UMAP) to visualize the data. Cell clustering was performed
829 using the 'FindClusters' function, using the previously identified significant principal
830 components (dims = 6) and a resolution parameter (r = 0.1) to define cluster granularity.
831 Differential expression analysis was performed using the 'FindAllMarkers' function (pct.min
832 = 0.25) to identify genes differentially expressed between clusters, with statistical significance
833 determined using the Wilcoxon rank sum test. Functional enrichment analysis of differentially
834 expressed genes was performed using gene set enrichment analysis (RBGOA) (*33*).

835 **KEGG pathway analysis**

836 KEGG analysis was performed using DAVID Bioinformatics Resources (NIAID/NIH) (*32*).
837 Gene lists of specifically overexpressed genes in each cluster were obtained after scRNA-seq
838 processing (genes with Log2FC > 0.25 and significant p-value < 0.001) and used for KEGG
839 annotation. The *C. gigas* reference genome from the DAVID bioinformatics resource was
840 used for this analysis, with thresholds of 2 for counts and 0.05 for EASE value (p-value).
841 KEGG annotation results were post-processed and presented as a heatmap showing the KEGG
842 pathway, fold enrichment, p-value significance and number of positive terms.

843 **Rank-Based Gene Ontology Analysis**

844 RBGOA first clusters GOs according to their representative genes to identify the most
845 significant GOs, and then ranks the identified biological processes according to their average
846 expression levels (over all representative genes). Finally, biological processes, molecular
847 functions and cellular components significantly enriched in DEGs are identified by a Mann-
848 Whitney rank test with a strict FDR correction. The files generated by the GO-MWU scripts
849 (https://github.com/z0on/GO_MWU) were post-processed to extract the category names and

850 the fraction indicating the number of "good candidates" relative to the total number of genes
851 belonging to that category. The "good candidates" were the genes that exceeded an arbitrary
852 'absValue' cutoff (defined as 0.001) in their significance measure. The results were presented
853 as a heatmap.

854 **Percoll Density Gradient Separation of Hemocytes10.3390/cells13242106**

855 A concentration series of Percoll (Cytiva, Sweden) diluted in Alsever modified medium was
856 prepared as follows: 10, 20, 30, 40, 50, 60 and 70 % (vol/vol). Discontinuous density gradients
857 (from 10 % Percoll with a density of 1.0647 to 70 % Percoll with d=1.1049) were made using
858 1.5 mL of each concentration and loaded with 1 mL of the hemocyte suspension
859 corresponding to approximately 2.10^7 cells. Centrifugation was performed (30 min, 800 g, 4
860 °C) in a swinging bucket on a Beckman Coulter JS-13.1 rotor (Beckman Coulter, USA).
861 Hemocytes concentrated at each density interface were collected separately with a 70 mm long
862 20Gx2.75" needle mounted on a 1 mL syringe. The hemocytes were then washed from the
863 Percoll by adding 10 mL of ice-cold Alsever modified medium, pelleted by centrifugation (10
864 min, 200 g, 4°C) and resuspended in Alsever modified medium or filtered seawater.

865 **Cytological description of the hemocyte populations**

866 200,000 fresh hemocytes were seeded onto a slide using a Cytospin 4 centrifuge (Thermo
867 Scientific, USA). The samples were then stained using the panoptic MCDH (Micro Chromatic
868 Detection for Hematology) (Cellavision, Sweden) staining protocol. This protocol produces
869 purple hues typical of Romanowsky-Giemsa staining results. After staining, the samples were
870 observed using a LEICA DMR (Leica AG, Germany) transmitted light microscope with a 40x
871 magnification objective. Each slide was imaged and the hemocytes were counted and
872 characterized based on their morphology.

873 **Real Time-quantitative Polymerase Chain Reaction (RT-qPCR)**

874 Total RNA was extracted using the RNeasy kit (Qiagen, the Netherlands) and cDNA was
875 synthesized from 1 µg total RNA using the Superscript IV kit (ThermoFisher Scientific, USA)
876 with oligo(dT) primers. RT-PCR was performed on LightCycler© 480 thermocycler using the

877 SYBR Green 1 Master kit (Roche, Switzerland). Primers were used at 200 nM. Primer
878 sequences are listed in **Supplementary Table S3**. Expression was normalized to *Cg-rps6*
879 reference gene. The standard cycling conditions were 95°C pre-incubation for 3 minutes
880 followed by 40 cycles of 95°C for 10 seconds, 60°C for 20 seconds, and 72°C for 25 seconds.
881 Each pair of primers was first tested and validated on a total hemocyte RNA preparation to
882 control melting curves and establish calibration lines.

883 **Oxidative Burst Assay**

884 The production of reactive oxygen species was quantified by luminescence assay. Briefly,
885 hemocytes from hemolymph puncture or Percoll density gradient were washed once with
886 filtered sterile water. 50 µL of hemocytes were plated in triplicate on a 96-well plate (3.10^5
887 cells/cm²). 50 µL of 40 mM luminol (Merck and Co, USA) was added to each well. After 45
888 minutes of incubation at room temperature, the oxidative burst was induced by adding 100 µL
889 of zymosan (Merck and Co, USA) at a multiplicity of infection (MOI) of 50:1. The plate was
890 immediately placed in a Berthold Centro XS3 LB 960 luminescence microplate reader
891 (Berthold GmbH, Germany) to measure luminescence emission every 2 minutes for 2 hours.

892 **Phagocytosis assay**

893 400 µL of filtered, sterile water-washed hemocytes were seeded in a 24-well plate at a
894 concentration of 3.10^5 cells/cm². After 15 minutes, 50 µL of zymosan was added to the
895 fractions at an MOI of 20:1. For LMG20012^T Vibrio, 50 µL of bacteria was added to a total
896 hemolymph at an MOI of 5:1. After 1 hour of contact at room temperature, the cells were
897 resuspended and 200 µL of the suspension was applied to a microscope slide using a Cytospin
898 centrifuge. The slides were then stained with MCDH and observed under a LEICA DMR
899 transmitted light microscope with a 40x magnification objective (Leica AG, Germany) to
900 count phagocytic cells.

901 **NitroBlueTetrazolium (NBT) staining**

902 ROS production was measured using NitroBlue Tetrazolium reduction after zymosan
903 stimulation. Briefly, 1 mL (1.10^6 cells) of hemocyte solution was mixed with 50 µL of filtered

904 NBT solution (15 mg/mL in water). Zymosan was added at a 4:1 MOI and the mixture was
905 incubated at room temperature for 10 minutes on a rocking shaker. Then, 50 µL of each
906 sample was plated onto glass coverslips and observed under a transmitted light microscope
907 LEICA DMR with a 40x magnification objective (Leica AG, Germany) to count NBT-positive
908 cells. The positions of positive NBT cells were recorded prior to MCDH staining to identify
909 hemocyte types.

910 **Rhodanine copper staining of hemocytes**

911 The storage of copper by hemocytes was examined by Rhodanine staining. Briefly, 1.10^5
912 hemocytes were plated on Superfrost slides using a cytospin and circled with a hydrophobic
913 pen to retain the staining solution. The Copper Stain Kit (SkyTek, USA) was used to stain the
914 hemocytes. As described by the kit manufacturer, one drop of Rhodanine solution was added
915 to 9 drops of acetate to form the working solution. Five drops were placed on the cytospin
916 cells and a 5 mL Eppendorf tube with the cap removed was placed over the cells to prevent
917 evaporation of the working solution. The slide and the balanced Eppendorf tube were placed
918 in a beaker of boiling distilled water for 20 minutes. The slide was then washed with 5 drops
919 of acetate and 3 drops of hematoxylin were placed on the slide for 1 minute at room
920 temperature. The slides were then washed a final time with acetate and observed under a
921 LEICA DMR transmitted light microscope with a 40x magnification objective (Leica AG,
922 Germany) to count rhodanine-positive cells. The positions of rhodanine-positive cells were
923 recorded prior to MCDH staining to identify hemocyte types.

924 **Pseudotemporal ordering of cells with Monocle3**

925 Cells from the *C. gigas* dataset were analyzed using Monocle3 (https://github.com/cole-
926 trapnell-lab/monocle3) (*42*). The Monocle3 analysis was performed on the Seurat object
927 following the aforementioned processing steps. Clustering information (features, genes,
928 partitions, clusters and UMAP coordinates) was transferred to a CDS object. The cell
929 trajectory was calculated using the *learn_graph* function. The *choose_graph_segments*
930 function was used to select three lineages. The gene expression along pseudotime data was

931 extracted from the result. Then, the data were used to plot genes along pseudotime in three
932 lineages using ggplot2 v3.4.4 R package and the heatmap was generated using the pheatmap
933 v1.0.12 R package.

934 **Statistical Analysis**

935 To evaluate differences between samples, a statistical analysis was performed using (version
936 4.2.1) (R Core Team, 2018) and appropriate packages. All data were examined for normality,
937 and statistical tests were selected accordingly. One-way analysis of variance (ANOVA) was
938 used for normally distributed data. Seven different hemolymph samples were used for
939 cytological analysis. Oysters were provided by our local supplier and approximately 300
940 hemocytes were counted per sample. Seven independent experiments were performed for
941 Percoll density gradient separation. A Tukey test was used to evaluate the statistical difference
942 between the proportions of hemocyte types. The phagocytic capacity of hemocytes was tested
943 in three independent experiments and statistical differences were evaluated using the Tukey
944 test for both the phagocytic capacity between hemocyte types and the number of particles per
945 phagocyte. Finally, oxidative burst capacity was tested 3 times on Percoll-separated
946 hemocytes. The Tukey test was also used to assess statistical differences between conditions.
947 The null hypothesis was rejected at a significance level of $p = 0.05$.

948
949

950 **References**

951

1. Y. M. Bar-On, R. Phillips, R. Milo, The biomass distribution on Earth. *Proceedings of the National Academy of Sciences* **115**, 6506–6511 (2018).

2. B. Leicester Bayne, *Biology of Oysters* (Academic Press, 1st Edition., 2017)vol. 41.

3. G. Zhang, L. Li, J. Meng, H. Qi, T. Qu, F. Xu, L. Zhang, Molecular Basis for Adaptation of Oysters to Stressful Marine Intertidal Environments. *Annu. Rev. Anim. Biosci.* **4**, 357–381 (2016).

4. FAO, Global aquaculture production. Fisheries and aquaculture division (2023).

5. F. Pernet, C. Lupo, C. Bacher, R. J. Whittington, Infectious diseases in oyster aquaculture require a new integrated approach. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150213 (2016).

6. A. Alfjorden, M. Areskog, D. Bruno, R. Carnegie, D. Cheslett, S. Feist, S. Ford, S. Jones, A. Lillehaug, L. Madsen, T. Renault, N. Ruane, P. Vennerstrom, New Trends in Important Diseases Affecting the Culture of Fish and Molluscs in the ICES Area 2002– 2015. *ICES cooperative research report*, doi: 10.17895/ices.pub.2800 (2017).

7. J. de Lorgeril, A. Lucasson, B. Petton, E. Toulza, C. Montagnani, C. Clerissi, J. Vidal-Dupiol, C. Chaparro, R. Galinier, J.-M. Escoubas, P. Haffner, L. Dégremont, G. M. Charrière, M. Lafont, A. Delort, A. Vergnes, M. Chiarello, N. Faury, T. Rubio, M. A. Leroy, A. Pérignon, D. Régler, B. Morga, M. Alunno-Bruscia, P. Boudry, F. Le Roux, D. Destoumieux-Garzón, Y. Gueguen, G. Mitta, Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. *Nat Commun* **9**, 4215 (2018).

8. N. M. Coyle, C. O'Toole, J. C. L. Thomas, D. Ryder, E. J. Feil, M. Geary, T. P. Bean, A. W. Joseph, A. Waine, D. Cheslett, D. W. Verner-Jeffreys, Vibrio aestuarianus clade A and clade B isolates are associated with Pacific oyster (Magallana gigas) disease outbreaks across Ireland. *Microbial Genomics* **9**, 001078 (2023).

9. A. Mesnil, M. Jacquot, C. Garcia, D. Tourbiez, L. Canier, A. Bidois, L. Dégremont, D. Cheslett, M. Geary, A. Vetri, A. Roque, D. Furones, A. Garden, P. Orozova, I. Arzul, M. Sicard, G. M. Charrière, D. Destoumieux-Garzón, M.-A. Travers, Emergence and clonal expansion of Vibrio aestuarianus lineages pathogenic for oysters in Europe. *Molecular Ecology* **32**, 2869–2883 (2023).

10. D. Oyanedel, A. Lagorce, M. Bruto, P. Haffner, A. Morot, Y. Labreuche, Y. Dorant, S. de La Forest Divonne, F. Delavat, N. Inguimbert, C. Montagnani, B. Morga, E. Toulza, C. Chaparro, J.-M. Escoubas, Y. Gueguen, J. Vidal-Dupiol, J. de Lorgeril, B. Petton, L. Degremont, D. Tourbiez, L.-L. Pimparé, M. Leroy, O. Romatif, J. Pouzadoux, G. Mitta, F. Le Roux, G. M. Charrière, M.-A. Travers, D. Destoumieux-Garzón, Cooperation and cheating orchestrate Vibrio assemblages and polymicrobial synergy in oysters infected with OsHV-1 virus. *Proceedings of the National Academy of Sciences* **120**, e2305195120 (2023).

11. T. Rubio, D. Oyanedel, Y. Labreuche, E. Toulza, X. Luo, M. Bruto, C. Chaparro, M. Torres, J. de Lorgeril, P. Haffner, J. Vidal-Dupiol, A. Lagorce, B. Petton, G. Mitta, A. Jacq, F. Le Roux, G. M. Charrière, D. Destoumieux-Garzón, Species-specific mechanisms of cytotoxicity toward immune cells determine the successful outcome of Vibrio infections. *Proceedings of the National Academy of Sciences* **116**, 14238–14247 (2019).

12. Y. Labreuche, F. Le Roux, J. Henry, C. Zatylny, A. Huvet, C. Lambert, P. Soudant, D. Mazel, J.-L. Nicolas, Vibrio aestuarianus zinc metalloprotease causes lethality in the Pacific oyster Crassostrea gigas and impairs the host cellular immune defenses. *Fish Shellfish Immunol* **29**, 753–758 (2010).

13. M. Lafont, A. Vergnes, J. Vidal-Dupiol, J. de Lorgeril, Y. Gueguen, P. Haffner, B. Petton, C. Chaparro, C. Barrachina, D. Destoumieux-Garzon, G. Mitta, B. Gourbal, C. Montagnani, A Sustained Immune Response Supports Long-Term Antiviral Immune Priming in the Pacific Oyster, Crassostrea gigas. *mBio* **11**, 10.1128/mbio.02777-19 (2020).

14. M. Fallet, C. Montagnani, B. Petton, L. Dantan, J. de Lorgeril, S. Comarmond, C. Chaparro, E. Toulza, S. Boitard, J.-M. Escoubas, A. Vergnes, J. Le Grand, I. Bulla, Y. Gueguen, J. Vidal-Dupiol, C. Grunau, G. Mitta, C. Cosseau, Early life microbial exposures shape the Crassostrea gigas immune system for lifelong and intergenerational disease protection. *Microbiome* **10**, 85 (2022).

15. N. R. de la Ballina, F. Maresca, A. Cao, A. Villalba, Bivalve Haemocyte Subpopulations: A Review. *Front. Immunol.* **13**, 826255 (2022).

16. E. Bachère, R. D. Rosa, P. Schmitt, A. C. Poirier, N. Merou, G. M. Charrière, D. Destoumieux-Garzón, The new insights into the oyster antimicrobial defense: Cellular, molecular and genetic view. *Fish & Shellfish Immunology* **46**, 50–64 (2015).

17. J.-M. Escoubas, B. Gourbal, D. Duval, T. J. Green, G. M. Charrière, D. Destoumieux-Garzón, C. Montagnani, "Immunity in Molluscs" in *Encyclopedia of Immunobiology*, M. J. H. Ratcliffe, Ed. (Academic Press, Oxford, 2016; https://www.sciencedirect.com/science/article/pii/B9780123742797120041), pp. 417–436.

18. A. S. Mount, A. P. Wheeler, R. P. Paradkar, D. Snider, Hemocyte-Mediated Shell Mineralization in the Eastern Oyster. *Science* **304**, 297–300 (2004).

19. W. S. S. Fischer, "Environmental influence on bivalve hemocyte function" in *Am Fish Soc Symp* (1988)vol. 18, pp. 225–237.

20. L. Donaghy, H.-K. Hong, H.-J. Lee, J.-C. Jun, Y.-J. Park, K.-S. Choi, Hemocyte parameters of the Pacific oyster Crassostrea gigas a year after the Hebei Spirit oil spill off the west coast of Korea. *Helgol Mar Res* **64**, 349–355 (2010).

21. N. R. de la Ballina, A. Villalba, A. Cao, Differences in proteomic profile between two haemocyte types, granulocytes and hyalinocytes, of the flat oyster *Ostrea edulis*. *Fish & Shellfish Immunology* **100**, 456–466 (2020).

22. W. Wang, M. Li, L. Wang, H. Chen, Z. Liu, Z. Jia, L. Qiu, L. Song, The granulocytes are the main immunocompetent hemocytes in Crassostrea gigas. *Developmental & Comparative Immunology* **67**, 221–228 (2017).

23. F. Mao, N.-K. Wong, Y. Lin, X. Zhang, K. Liu, M. Huang, D. Xu, Z. Xiang, J. Li, Y. Zhang, Z. Yu, Transcriptomic Evidence Reveals the Molecular Basis for Functional Differentiation of Hemocytes in a Marine Invertebrate, Crassostrea gigas. *Front. Immunol.* **11** (2020).

24. H. Zhang, H. Wang, H. Chen, M. Wang, Z. Zhou, L. Qiu, L. Wang, L. Song, The transcriptional response of the Pacific oyster *Crassostrea gigas* under simultaneous bacterial and heat stresses. *Developmental & Comparative Immunology* **94**, 1–10 (2019).

25. E. Zhang, Z. Li, L. Dong, Y. Feng, G. Sun, X. Xu, Z. Wang, C. Cui, W. Wang, J. Yang, Exploration of Molecular Mechanisms of Immunity in the Pacific Oyster (Crassostrea gigas) in Response to Vibrio alginolyticus Invasion. *Animals* **14**, 1707 (2024).

26. T. V. Nguyen, A. C. Alfaro, F. Merien, Omics approaches to investigate host–pathogen interactions in mass mortality outbreaks of Crassostrea gigas. *Reviews in Aquaculture* **11**, 1308–1324 (2019).

27. J. Meng, G. Zhang, W.-X. Wang, Functional heterogeneity of immune defenses in molluscan oysters Crassostrea hongkongensis revealed by high-throughput single-cell transcriptome. *Fish & Shellfish Immunology* **120**, 202–213 (2022).

28. R. D. Rosa, P. Alonso, A. Santini, A. Vergnes, E. Bachère, High polymorphism in big defensin gene expression reveals presence–absence gene variability (PAV) in the oyster Crassostrea gigas. *Developmental & Comparative Immunology* **49**, 231–238 (2015).

29. B. Kaminow, D. Yunusov, A. Dobin, STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. bioRxiv [Preprint] (2021). https://doi.org/10.1101/2021.05.05.442755.

30. C. Peñaloza, A. P. Gutierrez, L. Eöry, S. Wang, X. Guo, A. L. Archibald, T. P. Bean, R. D. Houston, A chromosome-level genome assembly for the Pacific oyster Crassostrea gigas. *GigaScience* **10**, giab020 (2021).

31. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* **33**, 495–502 (2015).

32. B. T. Sherman, M. Hao, J. Qiu, X. Jiao, M. W. Baseler, H. C. Lane, T. Imamichi, W. Chang, DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research* **50**, W216–W221 (2022).

33. R. M. Wright, G. V. Aglyamova, E. Meyer, M. V. Matz, Gene expression associated with white syndromes in a reef building coral, Acropora hyacinthus. *BMC Genomics* **16**, 371 (2015).

34. E. Bachère, D. Chagot, H. Grizel, Separation of Crassostrea gigas hemocytes by density gradient centrifugation and counterflow centrifugal elutriation. *Developmental & Comparative Immunology* **12**, 549–559 (1988).

35. C. Lambert, P. Soudant, G. Choquet, C. Paillard, Measurement of Crassostrea gigas hemocyte oxidative metabolism by flow cytometry and the inhibiting capacity of pathogenic vibrios. *Fish & Shellfish Immunology* **15**, 225–240 (2003).

36. P. Schmitt, R. Rosa, M. Duperthuy, J. de Lorgeril, E. Bachère, D. Destoumieux-Garzon, The Antimicrobial Defense of the Pacific Oyster, Crassostrea gigas. How Diversity may Compensate for Scarcity in the Regulation of Resident/Pathogenic Microflora. *Frontiers in Microbiology* **3** (2012).

37. G. Vogt, Hidden Treasures in Stem Cells of Indeterminately Growing Bilaterian Invertebrates. *Stem Cell Rev and Rep* **8**, 305–317 (2012).

38. M. Jemaà, N. Morin, P. Cavelier, J. Cau, J. M. Strub, C. Delsert, Adult somatic progenitor cells and hematopoiesis in oysters. *Journal of Experimental Biology* **217**, 3067–3077 (2014).

39. S. P. Fahl, M. Wang, Y. Zhang, A.-C. E. Duc, D. L. Wiest, Regulatory Roles of Rpl22 in Hematopoiesis: An Old Dog with New Tricks. *Crit Rev Immunol* **35**, 379–400 (2015).

40. M. Rehn, A. Wenzel, A.-K. Frank, M. B. Schuster, S. Pundhir, N. Jørgensen, K. Vitting-Seerup, Y. Ge, J. Jendholm, M. Michaut, E. M. Schoof, T. L. Jensen, N. Rapin, R. T. Sapio, K. L. Andersen, A. H. Lund, M. Solimena, M. Holzenberger, D. G. Pestov, B. T. Porse, PTBP1 promotes hematopoietic stem cell maintenance and red blood cell development by ensuring sufficient availability of ribosomal constituents. *Cell Reports* **39**, 110793 (2022).

41. E. I. Athanasiadis, J. G. Botthof, H. Andres, L. Ferreira, P. Lio, A. Cvejic, Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat Commun* **8**, 2045 (2017).

42. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).

43. L. Wang, H. Zhang, M. Wang, Z. Zhou, W. Wang, R. Liu, M. Huang, C. Yang, L. Qiu, L. Song, The transcriptomic expression of pattern recognition receptors: Insight into molecular recognition of various invading pathogens in Oyster *Crassostrea gigas*. *Developmental & Comparative Immunology* **91**, 1–7 (2019).

44. D. Hultmark, I. Andó, Hematopoietic plasticity mapped in Drosophila and other insects. *eLife* **11**, e78906 (2022).

45. U. Banerjee, J. R. Girard, L. M. Goins, C. M. Spratford, Drosophila as a Genetic Model for Hematopoiesis. *Genetics* **211**, 367–417 (2019).

46. E. Boada-Romero, J. Martinez, B. L. Heckmann, D. R. Green, The clearance of dead cells by efferocytosis. *Nat Rev Mol Cell Biol* **21**, 398–414 (2020).

47. D. C. Dale, L. Boxer, W. C. Liles, The phagocytes: neutrophils and monocytes. *Blood* **112**, 935–945 (2008).

48. A. S. Vanhove, T. P. Rubio, A. N. Nguyen, A. Lemire, D. Roche, J. Nicod, A. Vergnes, A. C. Poirier, E. Disconzi, E. Bachère, F. Le Roux, A. Jacq, G. M. Charrière, D. Destoumieux-Garzón, Copper homeostasis at the host vibrio interface: lessons from intracellular vibrio transcriptomics. *Environmental Microbiology* **18**, 875–888 (2016).

49. A. C. Poirier, P. Schmitt, R. D. Rosa, A. S. Vanhove, S. Kieffer-Jaquinod, T. P. Rubio, G. M. Charrière, D. Destoumieux-Garzón, Antimicrobial Histones and DNA Traps in Invertebrate Immunity: Evidences in Crassostrea Gigas *. *Journal of Biological Chemistry* **289**, 24821–24831 (2014).

50. G. Zaroogian, P. Yevich, Cytology and biochemistry of brown cells in *Crassostrea virginica* collected at clean and contaminated stations. *Environmental Pollution* **79**, 191–197 (1993).

51. B. Allam, D. Raftos, Immune responses to infectious diseases in bivalves. *Journal of Invertebrate Pathology* **131**, 121–136 (2015).

52. J. Nakhleh, L. El Moussawi, M. A. Osta, "Chapter Three - The Melanization Response in Insect Immunity" in *Advances in Insect Physiology*, P. Ligoxygakis, Ed. (Academic Press, 2017)vol. 52 of *Insect Immunity*, pp. 83–109.

53. P. Amparyup, W. Charoensapsri, A. Tassanakajon, Prophenoloxidase system and its role in shrimp immune responses against major pathogens. *Fish & Shellfish Immunology* **34**, 990–1001 (2013).

54. S. Picot, N. Faury, C. Pelletier, I. Arzul, B. Chollet, L. Dégremont, T. Renault, B. Morga, Monitoring Autophagy at Cellular and Molecular Level in Crassostrea gigas During an Experimental Ostreid Herpesvirus 1 (OsHV-1) Infection. *Front. Cell. Infect. Microbiol.* **12** (2022).

55. N. R. de la Ballina, A. Villalba, A. Cao, Differences in proteomic profile between two haemocyte types, granulocytes and hyalinocytes, of the flat oyster *Ostrea edulis*. *Fish & Shellfish Immunology* **100**, 456–466 (2020).

56. S. Aladaileh, S. V. Nair, D. Birch, D. A. Raftos, Sydney rock oyster (*Saccostrea glomerata*) hemocytes: Morphology and function. *Journal of Invertebrate Pathology* **96**, 48–63 (2007).

57. E. Rifkin, T. C. Cheng, H. R. Hohl, An electron-microscope study of the constituents of encapsulating cysts in the American oyster, *Crassostrea virginica*, formed in response to *Tylocephalum* metacestodes. *Journal of Invertebrate Pathology* **14**, 211–226 (1969).

58. N. De San Nicolas, A. Asokan, R. D. Rosa, S. N. Voisin, M.-A. Travers, G. Rocha, L. Dantan, Y. Dorant, G. Mitta, B. Petton, G. M. Charrière, J.-M. Escoubas, V. Boulo, J. Pouzadoux, H. Meudal, K. Loth, V. Aucagne, A. F. Delmas, P. Bulet, C. Montagnani, D. Destoumieux-Garzón, Functional Diversification of Oyster Big Defensins Generates Antimicrobial Specificity and Synergy against Members of the Microbiota. *Marine Drugs* **20**, 745 (2022).

59. N. R. de la Ballina, A. Villalba, A. Cao, Shotgun analysis to identify differences in protein expression between granulocytes and hyalinocytes of the European flat oyster Ostrea edulis. *Fish Shellfish Immunol* **119**, 678–691 (2021).

60. E. A. Pila, J. T. Sullivan, X. Z. Wu, J. Fang, S. P. Rudko, M. A. Gordy, P. C. Hanington, Haematopoiesis in molluscs: A review of haemocyte development and function in gastropods, cephalopods and bivalves. *Developmental & Comparative Immunology* **58**, 119–128 (2016).

61. M. de Bruijn, E. Dzierzak, Runx transcription factors in the development and function of the definitive hematopoietic system. *Blood* **129**, 2061–2069 (2017).

62. E. W. Scott, M. C. Simon, J. Anastasi, H. Singh, Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573–1577 (1994).

63. I. Bergiers, T. Andrews, Ö. Vargel Bölükbaşı, A. Buness, E. Janosz, N. Lopez-Anguita, K. Ganter, K. Kosim, C. Celen, G. Itır Perçin, P. Collier, B. Baying, V. Benes, M. Hemberg, C. Lancrin, Single-cell transcriptomics reveals a new dynamical function of transcription factors during embryonic hematopoiesis. *eLife* **7**, e29312 (2018).

64. F. Xin, X. Zhang, Hallmarks of crustacean immune hemocytes at single-cell resolution. *Front. Immunol.* **14** (2023).

65. H. Kwon, M. Mohammed, O. Franzén, J. Ankarklev, R. C. Smith, Single-cell analysis of mosquito hemocytes identifies signatures of immune cell subtypes and cell differentiation. *eLife* **10**, e66192 (2021).

66. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410 (1990).

67. P. Jones, D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.-Y. Yong, R. Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).

68. C. P. Cantalapiedra, A. Hernández-Plaza, I. Letunic, P. Bork, J. Huerta-Cepas, eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution* **38**, 5825–5829 (2021).

69. S. Götz, J. M. García-Gómez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talón, J. Dopazo, A. Conesa, High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research* **36**, 3420–3435 (2008).

70. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).

Acknowledgments

We are grateful to the staff of the Ifremer platform of "La Tremblade" for technical support in animal housing. scRNA-seq data generated and used in this work were produced through the MGX platform (University of Montpellier, CNRS ,INSERM). We thank the bioinformatic service of Ifremer (SEBIMER) for their help in bioinformatics and the qPCR platform GenSeq (University of Montpellier). We would like to thank Viviane Boulo and Danielle Mello for their enriching discussions, as well as all the members of the 2MAP laboratory team for the fruitful discussions throughout this project.

Ethics

This work did not require ethical approval from a human subject or animal welfare committee.

Funding

This work was funded by the Agence Nationale de la Recherche (MOSAR-DEF, ANR-19-CE18-0025; DECICOMP, ANR-19-CE20-0004 and TRANSCAN ANR-18-CE35-0009), the University of Montpellier, iSite MUSE and the "GT-Huître" initiative from Ifremer. This study falls within the framework of the "Laboratoires d'Excellence (LABEX)" Tulip (ANR-10-LABX-41). Sébastien De La Forest Divonne was awarded a PhD grant from the Region Occitanie (HemoFight project) and the University of Perpignan Via Domitia graduate school ED305.

Author contributions

Conceptualization: G.M., D.D.-G., B.G., G.M.C. and E.V.

Methodology: S.d.L.F.D., J.P., G.M., D.D.-G., B.G., G.M.C. and E.V.

Investigation: S.d.L.F.D., J.P., G.M.C. and E.V.

Supervision: G.M.C. and E.V.

Writing (original draft): S.d.L.F.D. and E.V.

Writing (review and editing): all authors.

Competing interests

The authors declare that they have no competing interests.

Declaration of AI use

We have not used AI-assisted technologies in creating this article.

Data and materials availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Raw reads are available at ENA under the project accession number PRJEB74031.