



HAL
open science

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

Sébastien de la Forest Divonne, Juliette Pouzadoux, Océane Romatif, Caroline Montagnani, Guillaume Mitta, Delphine Destoumieux-Garzon, Benjamin Gourbal, Guillaume M Charrière, Emmanuel Vignal

► To cite this version:

Sébastien de la Forest Divonne, Juliette Pouzadoux, Océane Romatif, Caroline Montagnani, Guillaume Mitta, et al.. Diversity and functional specialization of oyster immune cells uncovered by integrative single cell level investigations. *eLife*, 2024, 10.1101/2024.07.19.604245 . hal-04869462

HAL Id: hal-04869462

<https://hal.science/hal-04869462v1>

Submitted on 7 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.




Distributed under a Creative Commons Attribution 4.0 International License

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

Reviewed Preprint

v1 • November 13, 2024

Not revised

Sébastien de La Forest Divonne, Juliette Pouzadoux, Océane Romatif, Caroline Montagnani, Guillaume Mitta, Delphine Destoumieux-Garzon, Benjamin Gourbal, Guillaume M Charrière, Emmanuel Vignal 

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France • Ifremer, IRD, Institut Louis-Malardé, Univ Polynésie française, UMR 241 SECOPOL, Taravao, Tahiti - Polynésie française, France • IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Perpignan, France

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

eLife Assessment

This manuscript offers an exploration of the immune cells in the oyster *Crassostrea gigas*, by correlating distinct hemocyte morphotypes with specific single-cell transcriptional profiles. The evidence supporting the conclusion is **convincing**, deriving from the comprehensive dataset that not only captures unicellular diversity but also associates these cells with distinct immune roles, making it an **important** resource for the broader research community. There are some concerns on the data presentation that leave some questions.

<https://doi.org/10.7554/eLife.102622.1.sa3>

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 μ Var 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 characterized in detail have been found to subvert hemocyte defenses for their own benefit. These include the OsHV-1 μ Var herpesvirus (10) and virulent *Vibrio* strains of the species *Vibrio crassostreae* and *Vibrio tasmaniensis* (11), *Vibrio aestuarianus* and *Vibrio harveyi* (12), highlighting the critical role of hemocytes in oyster immunity. The development of immune-based prophylactic treatments, such as immune-priming and immune-shaping, represents a promising avenue for enhancing the natural defenses of oysters against pathogens and increasing their survival rate. Nevertheless, the advancement of such therapies is still constrained by a dearth of knowledge regarding the underlying molecular and cellular mechanisms (13, 14).

The study of hemocytes has a long history, dating back to the 1970s (for review see (15)). Hemocytes are cellular effectors of the immune system. They engage in phagocytosis to engulf and destroy potential pathogens, neutralizing parasites by encapsulation, or preventing pathogen dissemination by cell aggregation and the release of extracellular DNA traps (16). Furthermore, they engage in the humoral response by releasing cytokines, antimicrobial peptides, and reactive oxygen species (ROS), which enable them to combat pathogens (17). In addition to their role in oyster immunity, hemocytes have been implicated in numerous physiological processes, including shell repair (18), wound healing, nutrient transport, and environmental contaminant removal (19). Despite this acquired knowledge, hemocytes remain an under-characterized population of circulating immune cells. The lack of a unified classification and of molecular and functional genetic tools hinders our understanding of lineage ontogeny and functional specialization. Several studies have proposed different classifications of hemocytes in the *Ostreidae* family, with 3 to 4 hemocyte types reported (15). These classifications are primarily based on either microscopic or flow cytometry analyses. In *C. gigas*, three primary hemocyte cell types have been classically identified : blast, hyalinocyte, and granulocyte cells. While the immune response of *C. gigas* has been extensively studied using classical transcriptomics at the whole animal, tissue, or circulating hemocyte levels, these approaches have failed to consider the diversity of hemocyte cell types and lineages that underpin these responses (20). However, it is still imperative to accurately describe the diversity of these cells, understand their ontogeny, and delineate cell lineages to comprehend their specific roles. The advent of single-cell RNA sequencing (scRNA-seq) techniques has enabled the monitoring of global gene expression at the single-cell level with thousands of individual cells in a single experiment. This provides a unique opportunity to overcome these

limitations and deepen our understanding of hemocyte diversity and function in bivalves. Recently scRNA-seq was used to provide a first molecular description of a hemocyte population in the oyster *Crassostrea hongkongensis* (21 [↗](#)). However, in the absence of morphological and/or functional characterization studies, the authors could not deduce the hemocyte cell types to be matched to the transcriptomic profiles generated by scRNA-seq. While numerous transcriptomic analyses have been conducted on *C. gigas* hemocytes, none have adopted a single-cell approach. In this study, we present an integrative analysis of the diversity of *C. gigas* hemocytes at the single-cell level. To this end, we combined scRNA-seq from a pathogen-free adult oyster combined with cytological, cell fractionation and functional assays. This approach allowed us to create a comprehensive transcriptomic, cytological and functional atlas of hemocyte cell types. Our scRNA-seq analyses identified 7 distinct transcriptomic populations and functional annotation revealed distinct populations with specific functions, including phagocytosis, oxidative burst, energetic metabolism, enhanced transcription, translation and cell division. Quantitative cytology enabled the identification of 7 morphologically distinct hemocyte cell types, which allowed us to reconcile molecular and cytological data. Density gradients were used to separate hemocyte cell types and qPCR or functional assay analyses were performed to validate cell type-specific markers. By employing this integrated approach, we could identify 1 type of hyalinocyte, 2 types of blasts and 4 types of granular cells. Furthermore, we identified cell types that perform antimicrobial functions through phagocytosis, ROS production, copper accumulation, and expression of antimicrobial peptides. Finally, trajectory analysis of scRNA-seq data combined with functional analysis revealed distinct differentiation pathways that may control hemocyte ontology and differentiation processes. Based on these findings, we propose a more comprehensive and up-to-date classification of *C. gigas* hemocytes, with a more accurate description of the different cell types, their potential ontology and a precise description of their sub-functionalization.

Results

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

Oysters are known to exhibit a high degree of individual genetic polymorphism, including Copy Number Variation (CNV) and Presence Absence Variation (PAV) (22 [↗](#)). To prevent misinterpretation of the single-cell transcriptomic data, and to characterize the hemocyte cell types and their heterogeneity, we sampled hemocytes from a unique pathogen-free animal (Ifremer Standardized Animal, 18-month-old) and applied single-cell drop-seq technology to 3,000 single hemocytes (Fig. 1A [↗](#)). The scRNA-seq library was generated and sequenced, resulting in 127,959,215 high-quality filtered reads available for single-cell analysis (ENA project accession number PRJEB74031). Primary bioinformatics analysis was performed using the STAR solo aligner software (23 [↗](#)) against the *C. gigas* genome (Genbank reference GCA_902806645.1) from the Roslin Institute (24 [↗](#)). Of the 127,959,215 reads, 97 % showed a valid barcode and 89.2 % were successfully mapped to the genome with a saturation of 75.6 %. A total of 2,937 cells were profiled, yielding a median of 1,578 genes and 4,412 unique molecular identifiers (UMIs) per cell among the 23,841 total genes detected, with a sequencing saturation of 75.6 % (Supp. Table 1 [↗](#)). Secondary bioinformatic data processing was conducted using the Seurat R package (version 4.3.0) (25 [↗](#)). The data set was filtered to remove data corresponding to empty droplets or cell doublets. Cells with a gene number between 750 and 4,000 and less than 5 % mitochondrial genes were retained. After quality control processing, 120 cells corresponding to empty droplets and cell doublets were removed, and 2,817 cells were processed for data normalization. Finally, we performed linear dimensional reduction and clustering on the 3,000 most variable genes from 2,817 cells (Supp. Fig. S1 [↗](#)). Dimension reduction and clustering led to identifying 7 transcriptomic clusters, within which hemocytes were distributed. These 7 different clusters represented 27.6, 23.1, 17.8, 16.9, 7, 4.6 and 3 % of the total cells (Fig. 1B [↗](#)). For each transcriptomic cluster, a pattern of over- and under-represented transcripts in each cell was identified (Fig. 1C [↗](#)). Average Log2FC values and

percentage of expression in each cluster relative to all other clusters were calculated for the ten most differentially expressed genes (**Fig. 1D** [↗](#) and **Table 1** [↗](#)). Clear transcriptomic cell clusters were detected as well as specific gene markers for each cluster (**Supp. Data 1**). Only the transcriptomic signature of cluster 4 was less contrasting than that of the other clusters (**Fig. 1D** [↗](#)).

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

The scRNA-seq data demonstrated that specific functions are carried out by the hemocyte cell types that comprise the seven transcriptomic clusters. A preliminary overview of the functions over-represented in each cluster was obtained through a KEGG pathway analysis on the overrepresented transcripts ($\text{Log}_2\text{FC} > 0.25$) using the *C. gigas* annotation provided by the DAVID consortium ([26](#) [↗](#)), thereby identifying pathways specifically enriched in each cluster. Cluster 1 demonstrated enrichment in viral processing and endocytosis (**Figure 2A** [↗](#) and **Supplemental Table S2** [↗](#)). Additionally, it demonstrated enrichment in pyruvate metabolism, glycolysis/gluconeogenesis and the pentose phosphate pathways. Clusters 1 and 3 exhibited a distinctive enrichment in carbohydrate metabolism and endocytosis. In particular, cluster 3 exhibited enriched transcripts associated with glycolysis/gluconeogenesis and TCA cycle activities. Cluster 4 was enriched in protein synthesis, including transcription (spliceosome, ribosome, nucleocytoplasmic transport and mRNA surveillance pathway) folding, sorting and degradation of protein pathways. Clusters 2, 5 and 6 exhibited a shared signature of enrichment in ribosome-related genes. Cluster 2 demonstrated a specific enrichment in motor protein-coding genes responsible for cell motility and xenobiotic metabolism. Remarkably, clusters 1, 3, 4 and 5 exhibited enriched oxidative phosphorylation transcripts, whereas the transcripts of cluster 7 were enriched in vesicular trafficking and endo-lysosomal pathways (endocytosis, endosome, phagosome, lysosome, auto and mitophagy).

For a more detailed functional characterization of each transcriptomic cluster, we performed a re-annotation of the *C. gigas* genome using a combination of tools to enhance the GO term richness of the existing annotation prior to functional GO term analysis. To this end, we used the *C. gigas* genome (Genbank reference GCA_902806645.1) and the associated gff3 annotation file from the Roslin Institute ([24](#) [↗](#)). These files were used to extract and process the longest CDSs for GO-term annotation using the Orson pipeline (see Materials and Methods). Of the 30,724 extracted CDSs, 22,462 were annotated (GO-terms and sequence description), yielding an annotation percentage of 73.1 %. Of the 30,724 CDSs, 22,391 were annotated with Molecular Functions (MF), Biological Processes (BP), and Cellular Components (CC) GO-terms (**Supp. Fig. S2** [↗](#) and **Supp. Data 2**). Using the GO-term annotation and the Log_2FC of genes calculated after scRNA-seq processing in each cluster, GO enrichment analysis was performed by rank-based gene ontology analysis (RBGOA) ([27](#) [↗](#)). RBGOA analysis was performed on the GO-terms identified in each cluster (**Supp. Data 3**). The results are presented in **Figures 2B, C** [↗](#) and **D** [↗](#).

The scRNA-seq-based analysis identified seven distinct transcriptomic profiles for each cell cluster, thereby shedding light on greater heterogeneity and functional diversity of *C. gigas* hemocytes than previously described. Cluster 1, comprising 27.6 % of cells, is characterized by its morphology and capacity to remodel actin cytoskeleton, as well as oxidoreductase activity. This is evidenced by an enrichment in oxidoreductase activity acting on NAD(P)H and kinase activities BP, actin nucleation MF and lamellipodium and early endosomes CC. Cluster 2, comprising 23.1 % of cells, has an increased translation activity, as indicated by enrichment in BP, MF and CC terms related to rRNA binding, cytoplasmic translation, maturation of LSU-rRNA, and ribosomes respectively. Cluster 3, representing 17.8 % of cells, shows enrichment in cellular oxidation and actin nucleation, as evidenced by an enrichment in BP and MF related to oxidoreductase activities acting on metal ions, electron transfer, cellular oxidation and actin nucleation, and CC related to

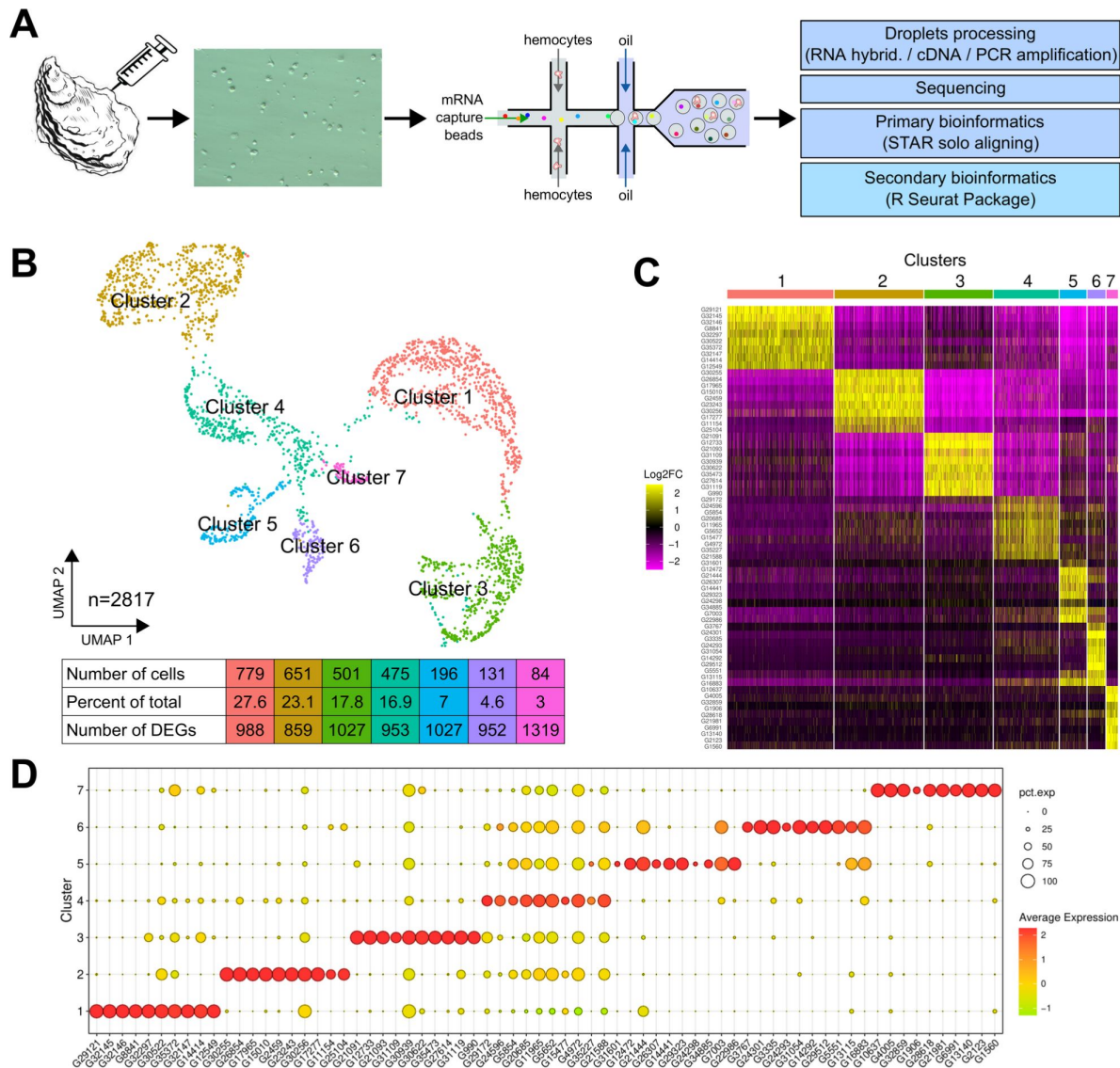


Fig. 1.

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)** Heatmap showing the top 10 overexpressed genes in each cell per cluster as determined by FindAllMarkers() function in Seurat, corresponding to clusters in UMAP plots from **Fig. 1B**. **(D)** Dot plot representing the ten most enriched DEGs per cluster based on average expression (avg_log2FC). The color gradient of the dot represents the expression level, while the size represents the percentage of cells expressing each gene per cluster.

Gene	log2FC	Pct 1	Pct 2	Description	Cluster
G29121	4.31	1	0.099	L-galactono-gamma-lactone oxidase	Cluster 1
G32145	4.18	0.992	0.079	Angiotensin-1 receptor	
G32146	3.90	0.985	0.063	Angiotensin-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	Angiotensin-1 receptor	
G14414	2.45	1	0.343	DUF4773 domain-containing protein	
G12549	2.38	0.974	0.132	Glutaredoxin domain-containing protein	
G30255	5.31	1	0.15	Metalloproteinase inhibitor 3	Cluster 2
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	
G21091	5.51	0.99	0.13	X-box binding protein-like protein	Cluster 3
G12733	5.42	1	0.12	Galectin	
G21093	5.17	0.94	0.05	loIG	
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	
G29172	3.09	0.78	0.49	G Protein Receptor F1-2 domain-containing protein	Cluster 4
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	
G6652	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
G31601	5.24	0.38	0.06	Polyribonucleotide nucleotidyltransferase	Cluster 5
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	
G3767	6.45	0.74	0.02	Natterin-1	Cluster 6
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	
G10637	4.21	0.99	0.03	BZIP domain-containing protein	Cluster 7
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	Sporozooite 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	

Table 1.

Top 10 overexpressed genes identified in each transcriptomic cluster.

The first column indicates the gene number according to the annotation. 'log2FC' represents the log2 fold change of the gene in the cluster compared to all other cells. 'Pct1' is the percentage of cells expressing the gene in the cluster and 'Pct2' is the fraction of cells expressing the gene in all other clusters. The description is the annotation of the expressed gen.

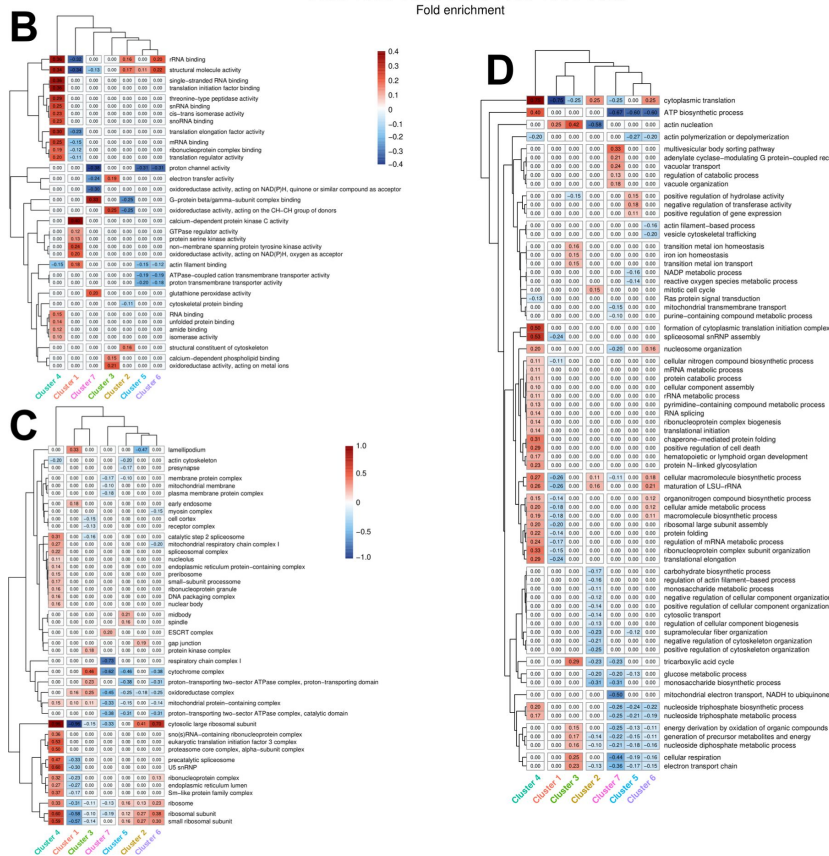
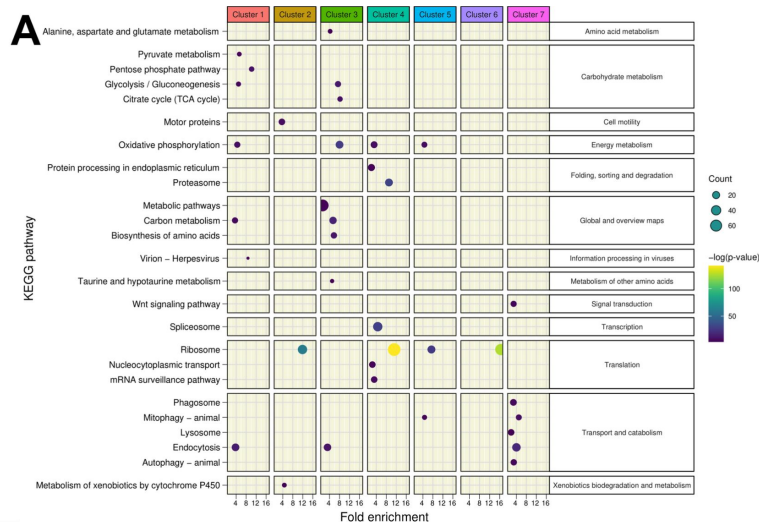


Fig. 2.

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 $-\log(p\text{-value})$ and the dot size represents the number (count) of enriched genes in each pathway category. The fold enrichment is shown on the x-axis. Panels (B), (C) and (D) show the results of Gene Ontology terms (GO-terms) for Biological Processes (BP), Cellular Components (CC) and Molecular Functions (MF) respectively, obtained with the overexpressed genes of each cluster (Absolute value $\text{Log}_2\text{FC} > 0.25$ and significant $p\text{-value} < 0.001$) using RBGOA analysis ($p\text{-value} \leq 0.001$) for three different ontology universes. Each panel corresponds to one ontology universe, and the analysis highlights over- and under-enriched terms. The number in the heatmap and the scale indicate the proportion of significant positive GO-terms. Positive values indicate over-enrichment and negative values indicate under-enrichment of the respective BP, CC and MF ontologies.

cytochromes, oxidoreductase complex, and mitochondrial protein-containing complexes. Cluster 4, representing 16.9 % of cells, has transcriptomic signatures reminiscent of spliceosome assembly, rRNA maturation, and ATP biosynthesis (BP related to RNA binding, translation, and proteasome, MF related to spliceosome assembly, rRNA maturation, and ATP biosynthesis, and CC related to nuclear bodies and ribosomes). Cluster 5, comprising 7 % of cells, demonstrates enrichment in structural molecule activity, hydrolase regulation, and gene expression, as evidenced by its BP, MF, and CC related to ribosome, midbody, and spindle localization. Cluster 6, comprising 4.6 % of cells, is characterized by a specific nucleosome organization, translation, and biosynthetic processes, with related BP, MF, and CC tied to ribosomes, ribonucleoprotein complex and rRNA binding. Cluster 7, representing a mere 3 % of cells, is characterized by multivesicular body sorting, vacuolar transport, vacuole organization, and G protein-coupled receptor signaling, with related BP, MF, and CC of the ESCRT machinery.

Seven morphologically distinct immune cell types are identified by quantitative cytology and transcriptomic markers

Cytological studies were conducted using MCDH staining to better characterize the diversity of circulating hemocytes in *C. gigas*. Seven distinct hemocyte morphotypes were identified: 3 non-granular (acidophilic blasts, basophilic blasts and hyalinocytes) and 4 granular (small granule cells, big granule cells, vesicular cells and macrophage-like) (Fig. 3A and Supp. Fig. S3). Hyalinocytes (30 % of total hemocytes) are large cells with an irregular spreading membrane. They contain an azurophilic cytoplasm without granulations and their irregular nucleus varies in size (Fig. 3A panel H). Macrophage-like cells (19 % of the hemocytes) present an irregular membrane punctuated by rare pseudopodia with a polylobed nucleus and a basophilic cytoplasm where polychromatic inclusions of various sizes could be observed (Fig. 3A panel ML). Basophilic blasts characterized by a basophil cytoplasm (Fig. 3A panel BBL) and acidophilic blasts with acidophil cytoplasm (Fig. 3A panel ABL) accounted for 18 % and 15 %, respectively, of the total hemocytes. They are rounded cells without granulation, with a uniform and regular dense nucleus and a high nucleo-cytoplasmic ratio. Small granule cells (12 % of total hemocytes) have an irregular membrane punctuated by rare pseudopodia, an acidophilic cytoplasm, and numerous homogeneous purple granules (Fig. 3A panel SGC). Big granule cells (4 % of total hemocytes) are rounded cells with a basophilic cytoplasm containing large, slightly dark purple to black vesicles of heterogeneous size (Fig. 3A panel BGC). Vesicular cells (2 % of total hemocytes) are rounded cells with acidophilic cytoplasm, rich in homogeneous transparent and fluorescent to UV light vesicles with an irregular nucleus (Fig. 3A panel VC). Only blasts (BBL, ABL) have a nucleo-cytoplasmic ratio less than 1, whereas the other hemocytes described have a nucleo-cytoplasmic ratio greater than 1. To gain further insight into the functional characterization of these diverse hemocyte morphotypes and to associate them with the transcriptomic clusters identified in the scRNA-seq, we enhanced an existing hemocyte fractionation approach using an isopycnic Percoll density gradient to sort the hemocytes (28). 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. 3A; 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).

To identify the transcriptomic cell clusters corresponding to the different hemocyte morphotypes, we used RT-qPCR to detect the expression of different cluster-specific marker genes in the different hemocytes in the Percoll density fractions. The marker genes were selected without any a

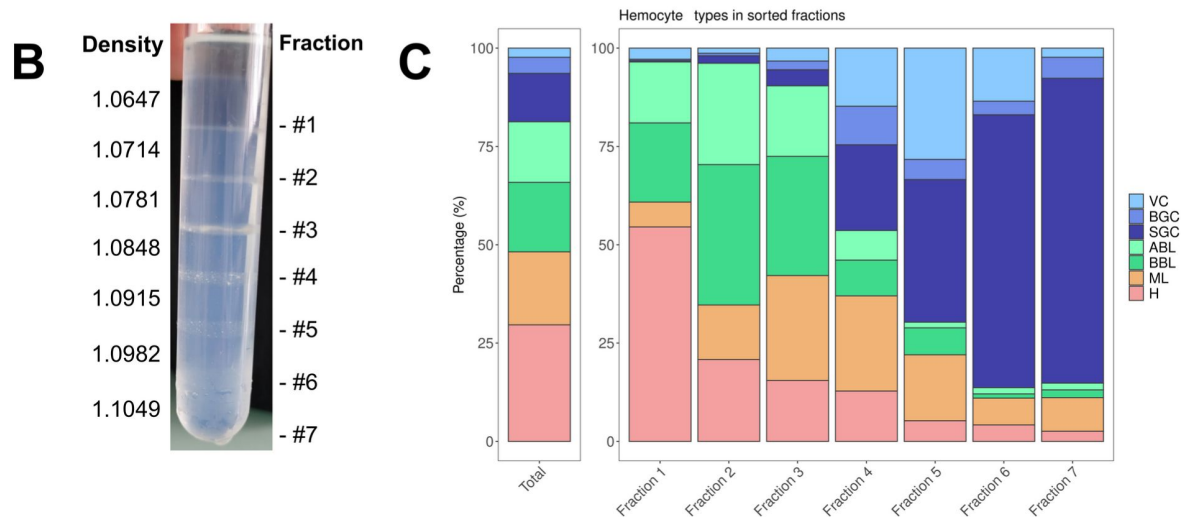
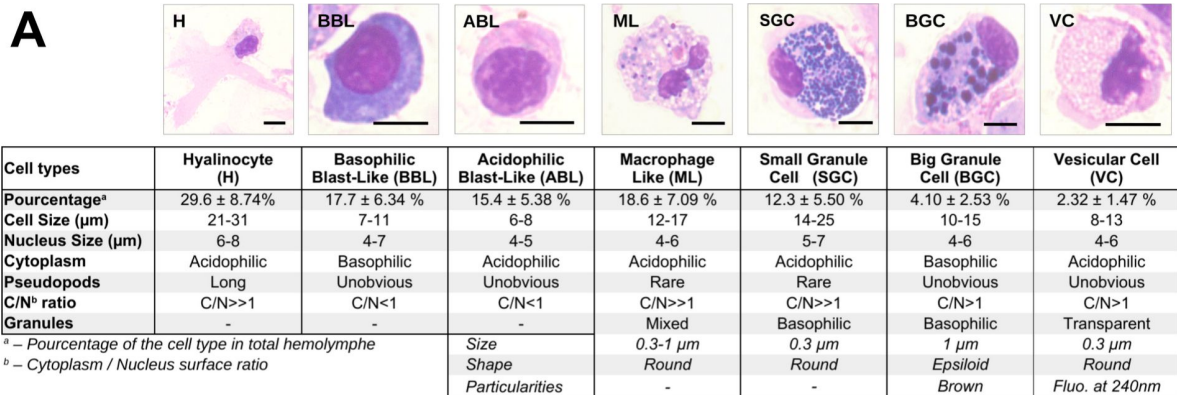


Fig 3.

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. Hyalinocytes (54 %) and blast cells (ABL & BBL) (35 %) were predominant in the first fraction. The second fraction showed an increase in blasts (ABL 26 %, BBL 36 %) and a decrease in hyalinocytes (21 %). Fraction 3 had a mixed content with a decrease in hyalinocytes and blasts (15 %), and a majority of blasts and macrophage-like cells. Fraction 4 had an increase in granular cells (SGC 22 %, BGC 10 %, VC 15 %) and a decrease in blasts (ABL 8 %, BBL 9 %). Fraction 5 showed an increase in small granule cells (36 %) and vesicular cells (28 %), and a decrease in big granule cells (5%). Fraction 6 had fewer hyalinocytes, macrophage-like cells, and blast cells compared to small granule cells and vesicular cells. The last fraction consisted mainly of small granule cells (81 %). **(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 Cells, **SGC** : Small Granule Cells, **ABL** : Acidophilic Blast Like cells, **BBL** : Basophilic Blast Like cells, **ML** : Macrophage Like cells and **H** : Hyalinocytes respectively. (**Supp. Fig. S4** [for statistics](#)).

priori assumptions based solely on their expression levels (Log2FC) and their percentage of expression (Pct1 / Pct2 ratio) in the cluster of interest relative to all other cell clusters (**Table 2**). The expression level of each marker in the seven different clusters confirmed that the 14 selected marker genes were differentially expressed in each scRNA-seq transcriptomic cluster (**Supp. Fig. S5**).

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

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

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

The hemocytes separated on the Percoll gradient were characterized functionally to gain further insight into their functional specialization. Two hemocyte functions, namely phagocytosis and production of Reactive Oxygen Species (ROS), are known to carry out major cellular antimicrobial activities. Granular cells have been suggested to be the professional phagocytes specialized for these functions (29). Phagocytosis and oxidative burst were studied using cell response toward zymosan particles (30).

The phagocytic activity of hemocytes was first tested on a sample of total oyster hemolymph. Cells were incubated for 1 hour with either zymosan particles or the bacterial strain *Vibrio tasmaniensis* LMG20012^T. Only the small granule cells and the macrophage-like cells exhibited efficient phagocytosis for both zymosan or vibrios, as observed after MCDH staining (**Fig. 5A**, **Supp. Fig. 6** and **Supp. Fig. 7A**). Macrophage-like cells and small granule cells showed a phagocytic activity of 49 % and 55 %, respectively, and a phagocytosis index of 3.5 and 5.2 particles per cell respectively (**Fig. 5B** and **Supp. Fig. 7B**), as confirmed in 3 independent experiments. Very limited phagocytic activity was observed for hyalinocytes (1.7 %), basophilic blasts (0.18 %), and big granule cells (2.7 %) with a phagocytosis index of 2, 1, and 2.5 particles per cell,

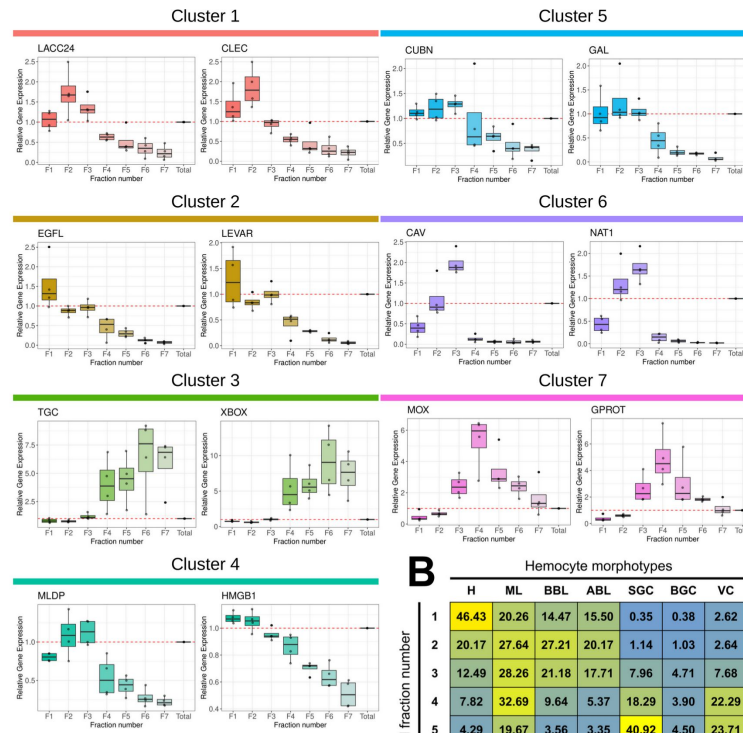
Gene Number	Avg. Log2FC	Pct.1/Pct.2 Ratio	Cluster	Description	Marker Name
G8994	2.05	38.20	1	Laccase-24	LACC24
G8841	3.13	20.16		C-type lectin domain-containing protein	CLEC
G24846	2.31	31.52	2	EGF-like domain-containing protein 8	EGFL
G17277	3.21	23.00		Putative modulator of levamisole receptor-1	LEVAR
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

Table 2.

Table of the 14 marker genes specific to the different transcriptomic clusters.

Gene number, average Log2FC, pct1/pct2 ratio (percentage of cells expressing this transcript in the cluster divided by the percentage of all other cells expressing this transcript) and cluster number are reported. The description is taken from our annotation and the marker name is derived from the description.

A



B

		Hemocyte morphotypes						
		H	ML	BBL	ABL	SGC	BGC	VC
Percol fraction number	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

C

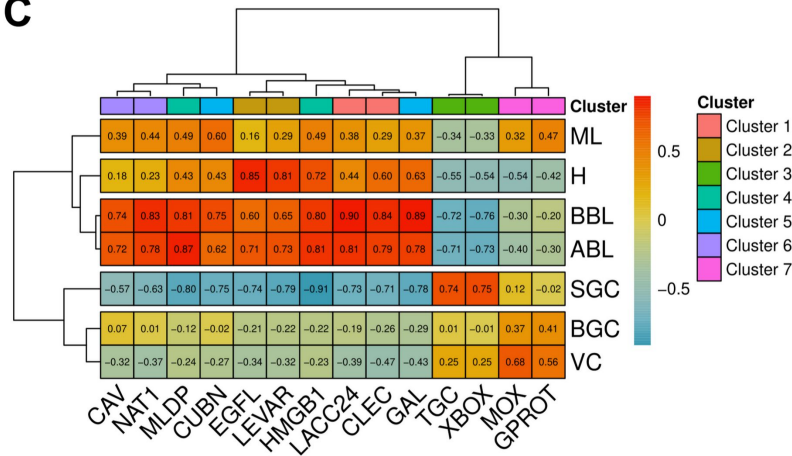


Fig 4.

Characterization of molecular markers specific to the different hemocyte morphotypes.

(A) Relative expression level of the 14 markers in the various fractions after gradient density sorting. The graphs show the relative expression of genes compared to their expression in total hemocytes in the various fractions (red dotted line). Standard deviations were calculated based on four independent experiments. **(B)** Average percentage of each hemocyte type in the 7 Percoll gradient fractions used to quantify marker gene expression by qPCR. **(C)** Correlation matrix between the relative gene expression of each marker gene in each fraction and the percentage of each hemocyte type in the same fractions. Values and color scale represent the Pearson correlation coefficient (r) ranging from -1 (inverse correlation) to $+1$ (full correlation). **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. **LACC24** : Laccase 24, **CLEC** : C-type lectin domain-containing protein, **EGFL** : EGF-like domain-containing protein 8, **LEVAR** : Putative regulator of levamisole receptor-1, **TGC** : TgC domain-containing protein, **XBOX** : X-box binding protein-like protein, **MLDP** : ML domain containing protein, **HMGB1** : High mobility group protein B1, **CUBN** : Cubilin, **GAL** : Galectin, **CAV** : Caveolin, **NAT1** : Natterin-1, **MOX** : DBH-like monooxygenase protein 1, **GPROT** : G protein receptor F1-2 domain-containing protein.

respectively (**Fig. 5B** [↗](#) and **Supp. Fig. 7B** [↗](#)). These results confirmed that only small granule cells and macrophage-like cells behave as professional phagocytes demonstrating robust phagocytic activity.

Only macrophage-like cells produce Reactive Oxygen Species

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

Previous studies have shown that granular cells can produce ROS ([29](#) [↗](#)). In contrast, we found that small granule cells collected in fraction 7 could not produce ROS through an oxidative burst. To identify which type of hemocyte from blast-like cells and macrophage-like cells produces ROS within a few minutes after exposure to zymosan, ROS production was investigated using NitroBlueTetrazolium (NBT) reduction to stain hemocytes directly. Correlative microscopic analysis using MCDH staining can then be conducted after the NBT reduction reaction. Macrophage-like cells were strongly and significantly stained by NBT reduction (33.2 %) (**Fig. 5E** [↗](#), **panels a & d**). In contrast, some small granule cells were lightly stained by NBT reduction (**Fig. 5F** [↗](#)). Blast-like cells, big granule cells, vesicular cells, and hyalinocytes were never NBT stained, confirming that these cell types were not involved in ROS production (**Supp. Fig. S7** [↗](#)). Taken together, these observations demonstrate that small granule cells and macrophage-like cells are the two professional phagocytes among hemocytes. However, only macrophage-like cells are capable of oxidative burst upon exposure to zymosan. In light of these new functional data, we further analyzed the expression level of NADPH-oxidase-related enzymes in the scRNA-seq dataset. The cells in cluster 1 predominantly expressed two NADPH oxidase isoforms (gene numbers G34908 and G23852) compared to other clusters (**Fig. 5H** [↗](#)). Furthermore, this cluster expresses macrophage-specific genes such as the angiotensin receptor (**Table 1** [↗](#)), as well as maturation factors for dual oxidase, an enzyme involved in peroxide formation (**Supp. Fig. S8** [↗](#)). Collectively, these data indicated that cells of transcriptomic cluster 1 corresponded to macrophage-like cells.

Small granule cells and big granule cells accumulate intracellular copper

The effects of copper on oyster hemocytes have been studied extensively due to its abundance in polluted marine environments ([19](#) [↗](#)). In addition, several studies have shown that *Vibrio* species pathogenic for *C. gigas* possess copper resistance genes, which are crucial for their survival within hemocytes ([9](#) [↗](#)). This prompted us to investigate which hemocyte types are involved in copper metabolism. To this end, total hemocytes were isolated from naive oysters and stained with rhodanine to reveal copper storage in cells. Rhodanine staining revealed that 33 % +/- 2 % of small granule cells and 30 % +/- 10 % of big granule cells exhibited a specific reddish/brown staining indicative of a high concentration of copper in their granules (**Fig. 5G** [↗](#) and **Supp. Fig. S9** [↗](#)). These results provide functional proof that SGC are specialized in metal homeostasis in addition to phagocytosis, as suggested by the scRNAseq data (Cluster 3).

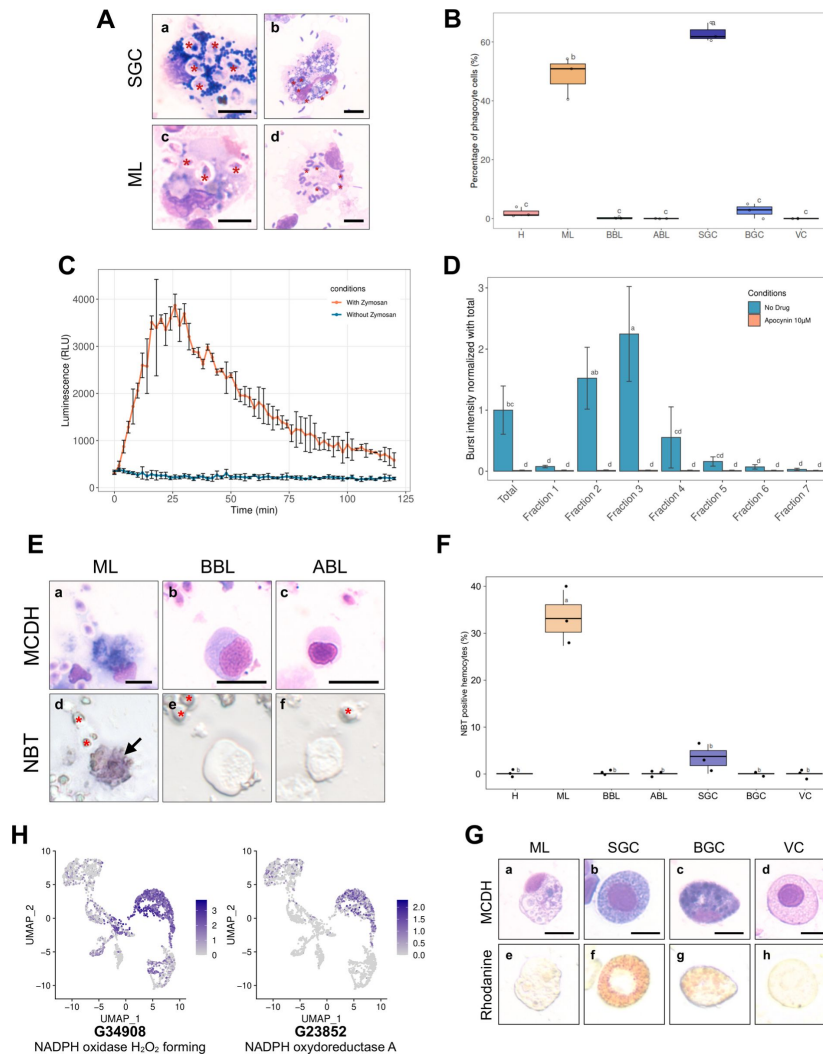


Fig 5.

Phagocytosis, Reactive Oxygen Species production capacity and copper storage of hemocytes.

(A) Images of small granule cells (SGC) and macrophage-like (ML) cells with phagocytosed zymosan particles (panels a & c - red stars) and *Vibrio tasmaniensis* LMG20012^T bacteria (panels b & d - red stars) from whole hemolymph sample. Scale bar: 5 μ m. (B) Quantification of the phagocytic activity of zymosan particles by each cell type. The graph shows the result of 3 independent experiments. (C) Luminescence recording to detect the production of Reactive Oxygen Species (ROS). In orange, a biphasic curve was obtained on naive oyster hemolymph after zymosan addition at t = 0 min. In blue, the control condition corresponds to hemocytes without zymosan addition. (D) Graph showing the intensity of ROS production in each Percoll fraction. Normalized burst intensity was calculated from the luminescence peak obtained from each fraction. In blue, no drug was added to the experiment, in orange, ROS production was impaired by the addition of apocynin. (E) NBT (NitroBlueTetrazolium) staining of hemocytes exposed to zymosan particles. Hemocytes morphology after MCDH staining: Macrophage Like (a), Basophilic (b) and Acidophilic (c) Blast cells. NBT staining of the different types of hemocytes (d-f). Red stars show zymosan and bacteria particles. Black arrows indicate Macrophage-Like cells. Scale bar: 10 μ m (F) Quantification of NBT-positive cells in the total hemolymph of oysters exposed to zymosan. (H) UMAP plots showing cells expressing NADPH oxidase found in the scRNA-seq dataset and their expression level. (G) Labeling of intracellular copper stores in *C.gigas* hemocytes. MCDH (upper panels) and rhodanine (lower panels) staining of oyster hemocytes to reveal copper accumulation. Scale bar: 10 μ m. For panels (B), (D) and (F) the alphabetic characters displayed above the data points in each plot represent statistically significant differences among the groups, as determined by Tukey's test following ANOVA. Groups denoted by different letters differ significantly at the p < 0.05 level of statistical significance. H : Hyalinocytes, ABL : Acidophilic Blast-Like cells, BBL : Basophilic Blast-Like cells, ML : Macrophage-Like cells, SGC : Small Granule Cells, VC : Vesicular Cells and BGC : Big Granule Cells.

Antimicrobial peptides are expressed by agranular cells, blasts and hyalinocytes

Antimicrobial peptides (AMPs) have long been studied for their role in the invertebrate humoral immune response. They are expressed by hemocytes, including those in *C. gigas* (31). The aim was to ascertain whether different hemocyte cell types expressed distinct AMPs. However, due to the limited sequencing depth, the scRNA-seq analysis was not sensitive enough to reveal AMP expression. This limitation was addressed by investigating AMP expression through RT-qPCR on Percoll density-fractionated hemocytes. The results indicated that Cg-Bigdefs 1-2, Cg-BPI and hemocyte defensin were predominantly expressed by agranular cells, blasts and hyalinocytes (ABL, BBL and H) (Fig. 6A, B and C). The expression of these AMPs was associated with Blasts abundance, while the expression of Cg-BigDefs 1-2 was only associated with hyalinocytes (Fig. 6D). Granular cells (VC, ML, BGC and SGC) did not seem to express any of the analyzed AMPs. These data suggest that some of the agranular cells are specialized in the production of humoral effectors.

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

The ontogeny, lineage and differentiation pathways of bivalves remain largely unknown (32). However, there are some indications of circulating and proliferating hemocyte progenitors in the hemolymph of *C. gigas* (33). GO-terms analysis of the 7 transcriptomic clusters revealed different functional signatures, including the transcriptomic signature of cluster 4, which showed a high expression level of ribosomal proteins (Supp. Fig. S1J). This particularity has been observed in hematopoietic stem cells in vertebrates (34–36). Furthermore, scRNA-seq approaches can now be used to deduce differentiation pathways from mRNA splicing variant analysis using bioinformatic tools like Monocle3 (37). This revealed an overexpression of genes involved in the splicing, transcription and translation continuum in the same fourth cluster (Fig. 2), thereby reinforcing the hypothesis of a pool of quiescent or immature cells that can differentiate upon stimulation.

Cluster 4 was chosen to enroot the pseudotime analysis to deduce differentiation pathways and cell lineages using Monocle3. (Fig. 7A). By temporally ordering the 2817 cells analyzed by scRNA-seq, 5 cell lineages could be defined (Fig. 7B). Differentiation pathway 1 leads to hyalinocytes (H) (Fig. 7C). This transition is characterized by the downregulation of 8 genes, two of which are transcription factors (G7003 and G31522). The hyalinocyte cluster was also characterized by the overexpression of genes involved in cell contractility (G22824, G153 and G11418). Differentiation pathway 2 leads to cells of cluster 5 and is characterized by the downregulation of about 30 genes, and 3 genes related to Zn finger protein, actin cleavage and Ig-like domain proteins were specifically overexpressed (G30887, G28864 and G32340) (Fig. 7D). Differentiation pathway 3 leads to cells of cluster 6 and is characterized by the down-expression of 12 genes, and the upregulation of a GATA family transcription factor (G31054) (Fig. 7E). The pathways between clusters 4, 5 and 6 found by Monocle3 analysis were pseudo temporally short and few specific markers were identified, suggesting that they were transcriptionally close. Differentiation pathway 4 leads to vesicular cells (VC) (Fig. 7F) where a large number of genes were silenced to give rise to these cells. Seven genes were specifically overexpressed in VC (G32111, G17226, G687, G16200, G32756, G10149 and G23495). Interestingly, 6 genes encoding potential transcription factors were downregulated in this lineage (G30997, G10637, G1067, G13555, G2123 and G27827). Differentiation pathway 5 leads to macrophage-like cells (ML), and is characterized by the underexpression of 35 genes and the overexpression of 13 genes. Among the 35 genes, 6 are putative transcription factors (G2123, G10637, G13555, G27827, G1067 and G11196) (Fig. 7G). Finally, we can outline a lineage ending in small granule cells (SGC) (Fig. 7H). This

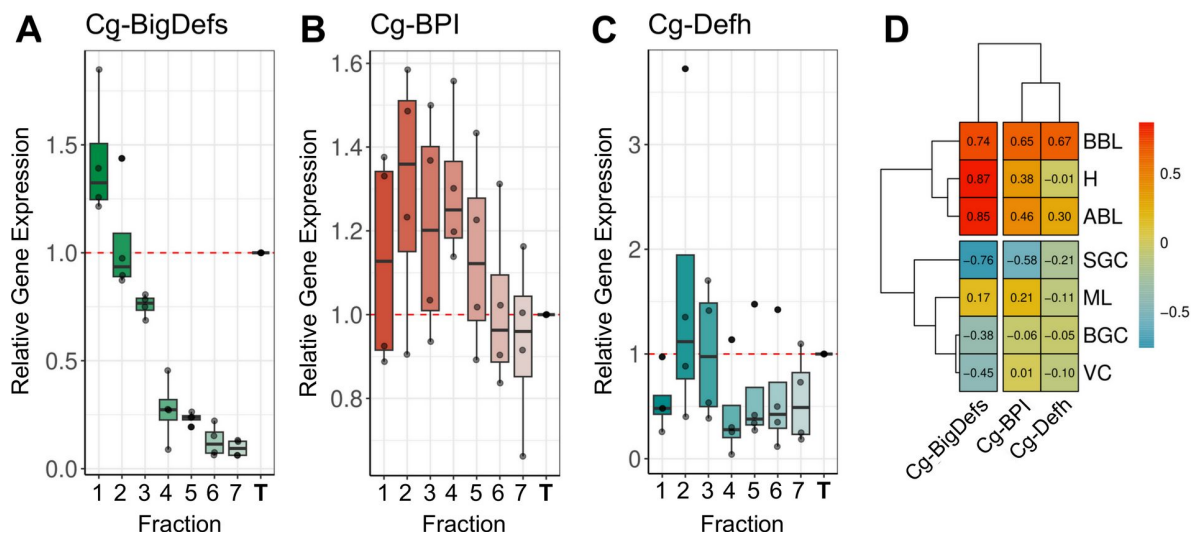


Fig 6.

Hemocyte expression profiles of some antimicrobial peptides.

(A) (B) and (C) Relative Gene Expression in the 7 Percoll hemocyte fractions of Big-Defensin1 & 2 (Cg-BigDefs), BPI (Cg-BPI) and hemocyte defensin (Cg-Defh), respectively, in comparison to the gene expression level in unfractionated hemolymph. (D) Correlation matrix between the relative gene expression of BigDefensin1 & 2, BPI and hemocyte defensin gene in each fraction and the percentage of each hemocyte type in each fraction (**H** : Hyalinocytes, **ABL** : Acidophilic Blast Like, **BBL** : Basophilic Blast Like, **SGC** : Small Granule Cell, **ML** : Macrophage Like, **BGC** : Big Granule Cell, **VC** : Vesicular Cell). Values and color scale represent the Pearson correlation coefficient (r) ranging from -1 (inverse correlation) to $+1$ (full correlation).

pathway involved cluster 4 (immature cells), VC, ML and SGC and was characterized by the downregulation of 27 genes, including 5 potential transcription factors (G10637, G13555, G27827, G1067 and G2123). SGC showed a distinct transcriptomic signature with 61 overexpressed genes, including 3 transcription factors (G1708, G21091 and G30622). Based on these findings, we postulate that immature cluster 4 cells possess pluripotent potential and can give rise to four terminally differentiated cell types : cluster 5 and 6 cells, hyalinocytes and small granule cells, and two other transient hemocyte cell types : vesicular cells and macrophage-like cells. All data from the Monocle3 analyses are available as Supplementary Data (**Supp. Fig. S10** [↗](#)).

Differentiation pathway analysis thus revealed the over- or under-expression of various transcription factors in the identified pathways. Given the established role of transcription factors as master regulators of cell differentiation and their utility in delineating cell lineages, we investigated the combinatorial expression patterns of transcription factors among the different transcriptomic cell clusters. Based on GO-terms annotation, 28 different sequences corresponding to transcription factors were isolated from the scRNA-seq dataset. The transcription factor function was confirmed by manual annotation (**Supp. Table S5** [↗](#)) and **Figure 71** [↗](#) shows the average expression profiles of these factors in the different clusters. **Supplementary Figure S11** [↗](#) illustrates their expression levels in single cells. Two transcription factors, CgATF5 and CgCRBL1, exhibited a contrasting expression profile with an increased average expression in macrophage-like cells, hyalinocytes and small granule cells versus a low expression profile in cells in clusters 4, 5, 6 and vesicular cells. We also identified transcription factors that were specifically expressed in the different transcriptomic clusters : CgSPDEF, CgSOCS3, CgFOS and CgTFEB were specific for vesicular cells, CgSOX8, CgXBOX and CgELF3 were specific for small granule cells, CgCR3L1 and CgTAL1 for hyalinocytes, and CgJUN, CgKLF5, CgKLF6 and CgCREM for macrophage-like cells. Eight additional transcription factors were specifically identified in cluster 6 (CgGATA3, CgPU.1 and CgELF2), cluster 5 (CgELF2, CgELK3 and CgIRF1) and cluster 4 (CgTAL1 and CgFOS1). These data potentially define four distinct hematopoietic lineages originating from one type of immature blast cells and give rise to hyalinocytes, SGC (via VC and ML), or two distinct differentiated blast-like cells. We also identified a combination of transcription factors specific to lineages and cell types that are potential master regulators of cell fate during hematopoiesis.

Discussion

The findings of our study represent a significant advancement in our understanding of the functional diversity and lineages of *C. gigas* hemocytes. Single-cell RNA-seq and cytology were combined to identify seven distinct hemocyte transcriptomic cell clusters and an equivalent number of morphotypes. These include four granular cell types (big granule cells, macrophage-like cells, small granule cells and vesicular cells), two distinct blast-like cells (basophilic and acidophilic blast-like cells), and one agranular epithelial-like cell type (hyalinocytes). A significant challenge was to identify correlations between transcriptomic and cytological data to fully define each cytological cell type. This challenge was overcome by combining multiple approaches, including isopycnic Percoll density gradient fractionation combined with the analysis of transcriptomic markers expression, and functional assays including phagocytosis, oxidative burst, copper accumulation, AMP expression and finally pseudotime analysis of gene expression. These results confirmed the historical classification of the 3 main cell groups : blasts, hyalinocytes, and granular cells ([15](#) [↗](#)) and deepened our understanding of the functional specificities of poorly characterized cell types. In particular, we identified distinct transcriptional and functional subtypes among blasts and granular cells with complementary immune specialization and lineage relationships between cells.

One significant outcome of the present study is the identification of cell types involved in antimicrobial activities, including phagocytosis, intracellular copper accumulation, an oxidative burst and antimicrobial peptide production. These cell types have been extensively studied for

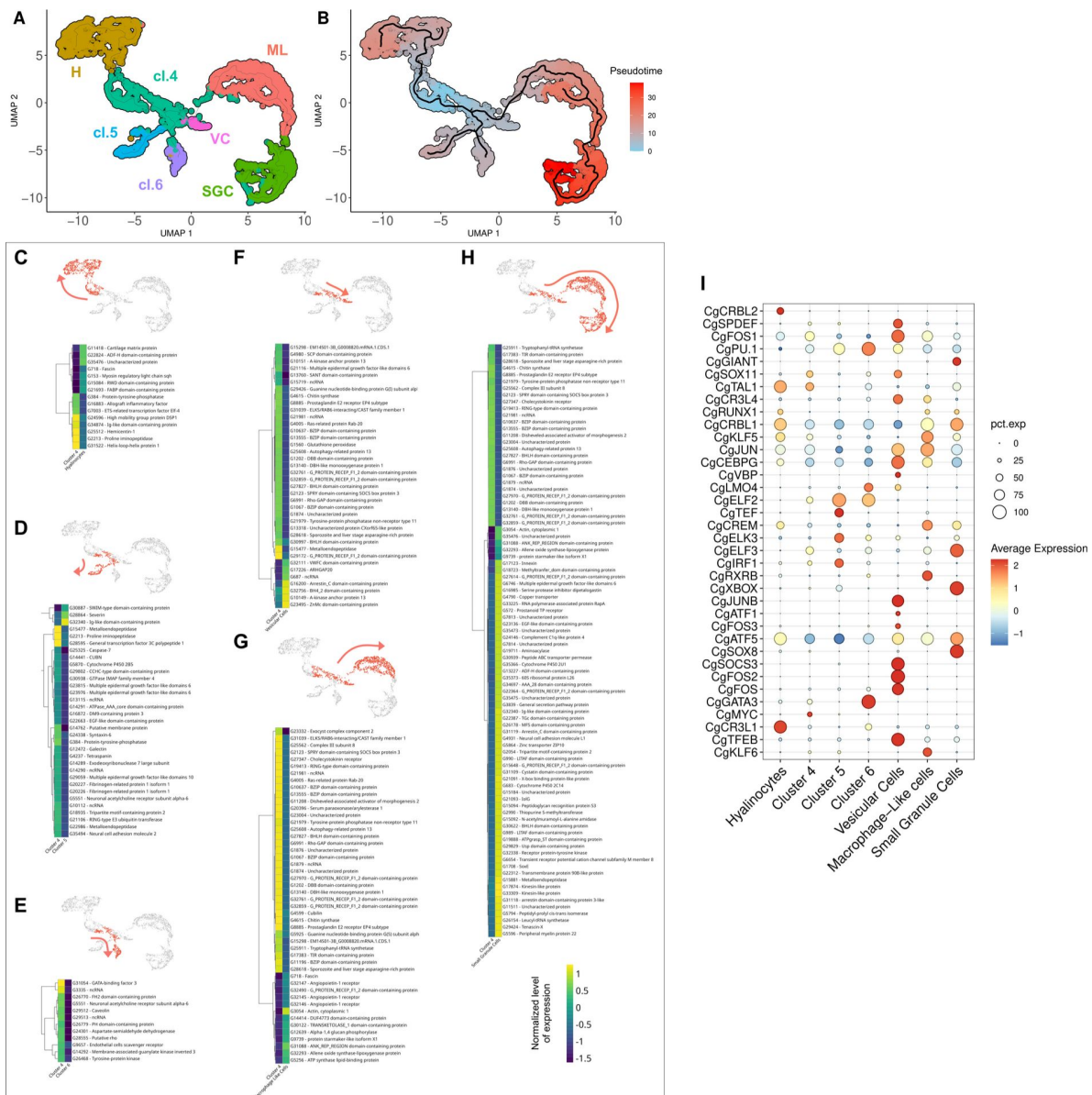


Figure 7.

Pseudotime ordering of cells revealed 6 potential differentiation pathways of hemocytes.

(A) UMAP plot of scRNA-seq analysis showing the 7 transcriptomic clusters used for pseudotime analysis. 4 clusters were identified cytologically (SGC for small granule cells - cluster 3, H for hyalinocytes - cluster 2, ML for Macrophage Like - cluster 1 and VC for vesicular cells - cluster 7), cl.4, cl.5, and cl.6 represent clusters 4, 5, and 6, respectively. (B) Graphical representation (UMAP projection) of the Monocle 3 pseudo-time order of the clustered cells. Cluster 4 (cl.4) was used as the origin for the pseudotime analysis. (C) (D) (E) (F) (G) and (H) show the gene expression level of selected marker genes obtained from the monocle3 trajectory analysis at the beginning and end of the modeled differentiation pathways (in red on the UMAP plot) from cluster 4 to hyalinocytes, to cluster 5 cells, to cluster 6 cells, to Vesicular Cells (VC), to Macrophage-Like cells (ML) and to Small Granule Cells (SGC) respectively. The color scale represents the normalized expression level of each gene. (I) Dot plot showing the average expression and the percentage of cells expressing identified transcripts encoding for transcription factors in the scRNA-seq dataset. Average expression is expressed in Log2FC.

their role in antibacterial and antiparasitic defenses, as they are found in the large majority of invertebrates (17 [↗](#)). Nevertheless, there has been considerable debate surrounding the cell types specialized for these critical immune functions in bivalves, particularly oysters. Moreover, the involvement of the different hemocyte subpopulations in immune functions is not yet fully understood.

Our findings reveal that the macrophage-like (ML) cells and small granule cells (SGC) are the sole hemocyte cell types that function as professional phagocytes, as demonstrated against Zymosan or *Vibrio*. These two distinct cell types could be distinguished functionally. First, two distinct transcriptomic clusters were identified for each cell type (Cluster 1 for ML / BGC and Cluster 3 for SGC). Secondly, only ML induced a measurable oxidative burst. Thirdly, only SGC accumulated intracellular copper in specific granules. The two types of professional phagocytes belong to the same granular cell lineage, as determined by pseudotime analysis. The notion that ML could serve as a precursor for SGC may seem counterintuitive. However, this is not an isolated phenomenon among invertebrate hemocytes. For instance, in *Drosophila* larvae (38 [↗](#)), some populations of professional phagocytes, the sessile plasmatocytes, give rise to crystal cells or lamellocytes that are morphologically and functionally distinct from plasmatocytes (39 [↗](#)). Similarly, the existence of multiple professional phagocytes in the oyster is reminiscent of vertebrate macrophages, polynuclear neutrophils and dendritic cells, which possess distinct functional specializations, including efferocytosis (40 [↗](#)), oxidative burst, NETosis, or antigen presentation (41 [↗](#)).

The characterization of professional phagocytes in oysters is of particular importance for a deeper understanding of oyster-*Vibrio* interactions during pathogenesis. Some of the most extensively studied oyster pathogens, including strains of *V. tasmaniensis* and *V. aestuarianus francensis*, harbor virulence traits that enable them to disrupt the phagocytic activity of hemocytes during pathogenesis. For instance, *V. tasmaniensis* behaves as a facultative intracellular pathogen with phagocytosis-dependent cytotoxicity (11 [↗](#)). Additionally, both *V. tasmaniensis* and *V. aestuarianus* display resistance to copper toxicity through CopA and CusA/B transporters. This trait is essential for the survival and virulence of these pathogens in oysters (9 [↗](#), 42 [↗](#)). Since SGCs have been demonstrated to be professional phagocytes with copper-rich granules, the cellular interactions between SGCs and these vibrios are likely to be critical during the antibacterial host response and pathogenesis. ML are phagocytes that possess a very potent NADPH-oxidase-dependent oxidative burst. The oxidative burst is a rapid and potent antimicrobial response observed in professional phagocytes, such as polynuclear neutrophils in mammals. It is worth noting that NADPH-dependent ETosis has been observed in *C. gigas* in a manner analogous to that observed in human neutrophils (43 [↗](#)). This cell death is characterized by the projection of DNA extracellular traps that capture and kill some pathogens like *Vibrio* (43 [↗](#)). Therefore, it is reasonable to hypothesize that ML may also be involved in ETosis.

The specialized functions of the two other types of granular cells, the BGCs and the VCs, remain unclear. Despite the difficulty in identifying a specific scRNAseq transcriptomic cluster for BGCs, the level of expression of laccase 24 was found to be higher in a particular subcluster among ML (Supp. Fig. S8 [↗](#)) and pseudo-time analysis highlighted the same subcluster as an alternative differentiation state among ML (Fig. 7B [↗](#)). The enrichment of transcripts involved in oxidation-reduction pathways, particularly laccase 24, aligns with their potential role in melanization and response to oxidative stress (44 [↗](#)). While melanin-like deposits have been observed in some cases of infestation by the parasites *Martelia* or *Bonamia* in the Sydney rock oyster (45 [↗](#)), this mechanism is not as robust as that described in arthropods, which perform melanization through a prophenoloxidase activation cascade (46 [↗](#), 47 [↗](#)). In many marine invertebrates (44 [↗](#)), a type of hemocyte known as Brown Cells could be related to the BGCs described here. When observed without any staining (as in Fig. 5F [↗](#)), their big granules with a yellow to dark brown content appear to align with the historical description of brown cells that often infiltrate tissues (like gills) in animals exposed to polluted waters (44 [↗](#)). It has therefore been theorized that these cells are involved in detoxification processes. Our pseudotime analysis indicates that they likely originate

from ML, with limited phagocytosis activity and a specialized role in melanization and potentially heavy metal detoxification, as evidenced by rhodanine staining showing copper accumulation in some of their granules (**Fig. 5G**). Further studies are recommended to clarify the role of BGCs, particularly in the context of parasitic infestation or exposure to toxic stresses. Lastly, VC are granular cells that remain to be functionally characterized. Their transcriptomic profile suggests strong intracellular vesicular trafficking and autophagy activity. However, our functional assays did not reveal any particular immune-related function. Their clear granules appear auto-fluorescent when illuminated with UV light (**Supp. Fig. S11**) but the biochemical nature of the content of these granules remains to be characterized. As they also appear as cell intermediates along the granular cell differentiation pathway in pseudotime analysis, they could represent functionally immature precursors of the other three granular cell types, much like promyelocytes which possess specific azurophilic granules but are functionally immature precursors during granulocyte differentiation in humans. However, the enrichment of autophagy-related transcripts in VC calls for further investigation into a potential role in antiviral immunity, as autophagy has been suggested to play a role in the response to OsHV-1 virus (48).

Hyalinocytes are a homogenous cell type, with only one morphotype matching one transcriptomic cluster. It can be deduced from pseudotime analysis that they originate from a specific and very different differentiation pathway than the granular cells. According to the literature, hyalinocytes are involved in the early stages of inflammation and can infiltrate wounds and interact with foreign particles. In the flat oyster *Ostrea edulis*, they contribute to shell production and wound healing (49). In the Sydney rock oyster they play a role in cell aggregation (50), while in *C. virginica* (51) they contribute to encapsulation, reminiscent of lamellocytes in *Drosophila*. Our results suggest an important role of AMP expression in the immune response. Indeed, *Cg-BigDefs*, which participate in the control of oyster microbiota (52), were found to be expressed in both hyalinocytes and blast-like cells. Moreover, hyalinocytes from the oyster *O. edulis* have been shown to express the AMP Myticin C (53), which lends further support to this immune function. Among the AMPs, we also found that *Cg-BPI* and *Cg-Defh*, are more expressed in BBL, ABL than in other cell types. These results are somewhat unexpected, given the prevailing assumption that AMPs are stored in granules of granular cells, rather than agranular cells (16). These results highlight the necessity to reassess the role of specific agranular cell types in the active production of humoral immune effectors. Our findings suggest that hyalinocytes and/or the blast-like cells may be a cellular target of the OsHV-1 virus, the causal agent of POMS, which dampens the expression of certain AMPs (7), thereby inducing bacterial dysbiosis. It is still unclear whether this is due to a decreased expression of AMPs and/or inhibition of immature blast cell differentiation involved in the renewal of agranular cell types.

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

In the animal kingdom, the innate immune system relies on specialized cells derived from pluripotent precursors through hematopoiesis. Transcription factors in particular have been found to exhibit a high degree of conservation throughout the animal kingdom, from invertebrates to vertebrates. However, the mechanisms underlying the functional differentiation of bivalves are only partially available or understood (54). Recent research has indicated the potential existence of hemocyte progenitors, also known as blast-like cells in several bivalve

species. These include clams, mussels, scallops, marine mussels, freshwater mussels, oysters, pearl oysters, and wing shells (54 [↗](#)). However, various models for hematopoiesis in bivalves have been proposed and extensively debated without a clear consensus or definitive proof. Single-cell RNA sequencing and pseudotime analysis have enabled us to propose a refined model of hematopoiesis at an unprecedented level of detail. In this model, the different types of hemocytes are likely produced through four differentiation pathways that originate from a common progenitor. One pathway results in the formation of agranular hyalinocytes, while another independent pathway gives rise to the granular cells, including VC, ML, BGC, and SGC. Furthermore, two additional differentiation pathways have been identified that may lead to terminally differentiated blast-like cells. The differentiation of mature hemocytes involves the establishment of lineage-specific gene expression profiles, which rely on transcription factors to modulate the expression of their target genes. Our study identified a combination of transcription factors that are differentially expressed during *C. gigas* hematopoiesis and are specific to differentiation stages and cell lineages, including many well-known hematopoietic transcription factors such as GATA, PU-1, TAL1 and SOX factors, which are positioned in this model of *C. gigas* hematopoiesis (55 [↗](#)–57 [↗](#)) (Fig. 8 [↗](#)).

This study significantly advances our understanding of bivalve immunity, especially in comparison to arthropods. By introducing a standardized reference hemocytogram for oysters using MCDH staining, defining cell type-specific markers and key transcription factors likely involved in cell fate determination, as well as clarifying the functions of different hemocytes, we have paved the way for future in-depth studies. This will facilitate further studies of the oyster's immune response to various biotic and abiotic stress at the cellular level. Improved comprehension of antiviral and antibacterial responses in bivalves, along with an enhanced understanding of immune priming and immune memory at the cellular level in bivalves, will benefit health and population management practices for sustainable aquaculture production. These findings will also contribute to the broader field of evolutionary immunology by enabling comparative studies and elucidating the diversification of immune cells and immunity-related genes in a protostome.

Materials and Methods

Conservation of oysters

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

Hemocyte collection and processing

Crassostrea gigas hemocytes were collected from live animals by puncture of the adductor muscle. The oyster shell was incised on the posterior side with forceps. Hemolymph was collected using a 23Gx1" needle mounted on a 5 mL syringe prefilled with 2 mL of ice-cold Alsever modified medium (20.8 g glucose – 8 g trisodium citrate - 22.5 g sodium chloride - 0.4 g BSA - pH=7.5). Samples were centrifuged for 4 minutes at 200 g – 4 °C and the supernatant was removed and replaced with 1 mL of fresh Alsever modified medium. Each hemocyte sample was thoroughly checked for quality, counted under a microscope using KOVA (Kova International, USA) slides, and stored on ice prior to processing. For scRNA-seq analysis, resuspended hemocytes were filtered on 30 µm filters, counted, the solution was adjusted to 1.10⁶ cells per mL and stored on ice prior to 10X genomic library preparation.

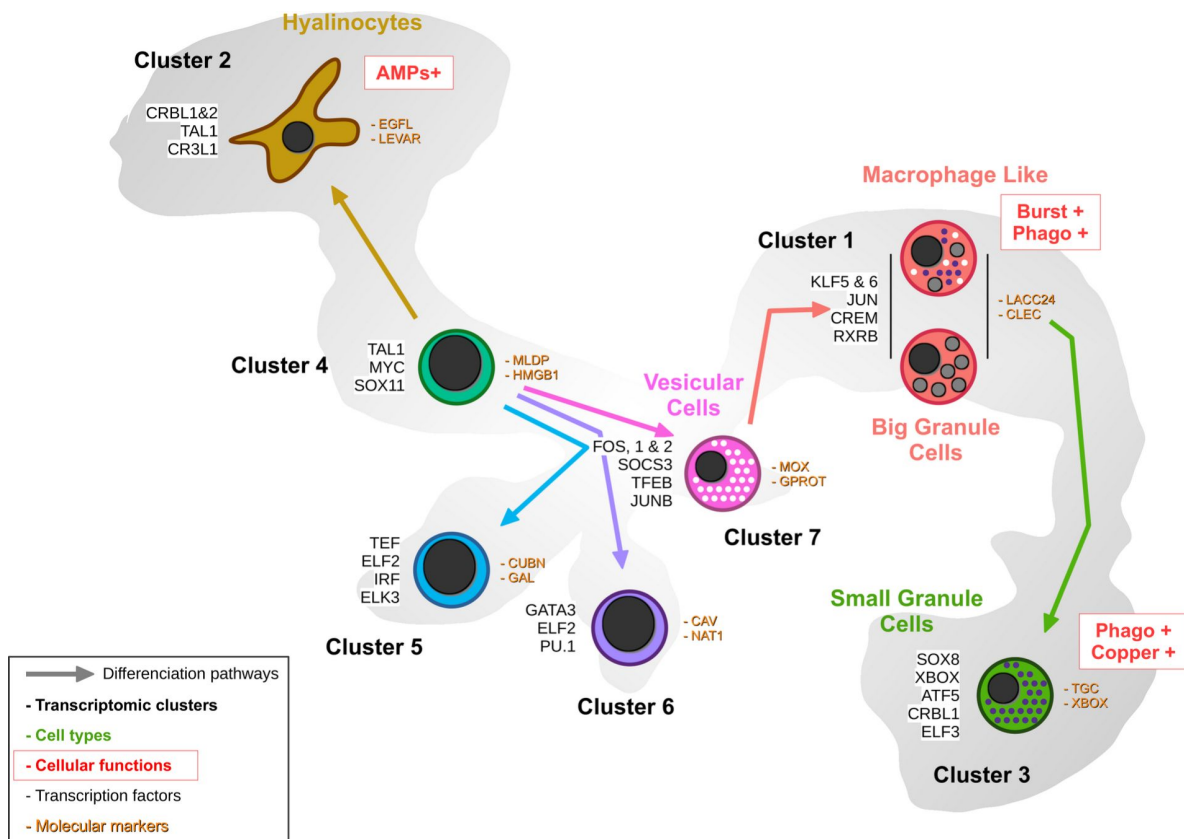


Figure 8.

Proposed hemocyte ontogeny in *Crassostrea gigas* based on the transcriptomic, cytological and functional results obtained.

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

C. gigas genome annotation

The *C. gigas* genome (Genbank GCA_905397895.1) (24) was used as a reference. Prior to annotation, the longest CDS sequences were extracted from the gff3 file. Annotation was realized using the ORSON script (<https://gitlab.ifremer.fr/bioinfo/workflows/orson>). ORSON combines cutting-edge tools for annotation processes within a Nextflow pipeline. The ORSON script performs a sequence similarity search with BLAST (58) against the Uniprot-Swissprot and Uniprot-trEMBL databases, and functional prediction with InterProScan (59) and eggNOG (60) orthogroup annotation. Interproscan analysis was performed against Pfam, Prosite, CDD, TIGR, SMART, SuperFamily, PRINTS and Hamap databases. Results were collected and processed using Blast2GO (61) for annotation mapping and validation.

Drop Seq-based scRNA-seq library generation

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

scRNA-seq analysis

Reads were aligned to the *C. gigas* reference genome (Genbank GCA_905397895.1) (27) using STAR solo software (v 2.7.10) (23). Unique molecular identifiers (UMIs) were extracted and counted for each cell, and an expression matrix was generated for further analysis. Single-cell RNA sequencing (scRNA-seq) data analysis was performed using the R programming language (version 4.2.1) (R Core Team, 2018) and the Seurat package (version 4.3.0) (62). The data were then pre-processed to remove unwanted sources of variation and to normalize gene expression. Cells with small library sizes and a high proportion of mitochondrial genes were excluded. Data normalization was performed using the SCTransform method. After normalization, highly variable genes were identified using the FindVariableFeatures function and the top variable genes were selected for downstream analyses. Dimensionality reduction was performed using Principal Component Analysis (PCA), followed by Uniform Manifold Approximation and Projection (UMAP) to visualize the data. Cell clustering was performed using the 'FindClusters' function, using the previously identified significant principal components (dims = 6) and a resolution parameter ($r = 0.1$) to define cluster granularity. Differential expression analysis was performed using the 'FindAllMarkers' function (pct.min = 0.25) to identify genes differentially expressed between clusters, with statistical significance determined using the Wilcoxon rank sum test. Functional enrichment analysis of differentially expressed genes was performed using gene set enrichment analysis (RBGOA) (27).

KEGG pathway analysis

KEGG analysis was performed using DAVID Bioinformatics Resources (NIAID/NIH) (26). Gene lists of specifically overexpressed genes in each cluster were obtained after scRNA-seq processing (genes with $\text{Log}_2\text{FC} > 0.25$ and significant p-value < 0.001) and used for KEGG annotation. The *C. gigas* reference genome from the DAVID bioinformatics resource was used for this analysis, with thresholds of 2 for counts and 0.05 for EASE value (p-value). KEGG annotation results were post-processed and presented as a heatmap showing the KEGG pathway, fold enrichment, p-value significance and number of positive terms.

Rank-Based Gene Ontology Analysis

RBGOA first clusters GOs according to their representative genes to identify the most significant GOs, and then ranks the identified biological processes according to their average expression levels (over all representative genes). Finally, biological processes, molecular functions and cellular components significantly enriched in DEGs are identified by a Mann-Whitney rank test with a strict FDR correction. The files generated by the GO-MWU scripts (https://github.com/z0on/GO_MWU) were post-processed to extract the category names and the fraction indicating the number of “good candidates” relative to the total number of genes belonging to that category. The “good candidates” were the genes that exceeded an arbitrary ‘absValue’ cutoff (defined as 0.001) in their significance measure. The results were presented as a heatmap.

Percoll Density Gradient Separation of Hemocytes

A concentration series of Percoll (Cytiva, Sweden) diluted in Alsever modified medium was prepared as follows: 10, 20, 30, 40, 50, 60 and 70 % (vol/vol). Discontinuous density gradients (from 10 % Percoll with a density of 1.0647 to 70 % Percoll with $d=1.1049$) were made using 1.5 mL of each concentration and loaded with 1 mL of the hemocyte suspension corresponding to approximately $2 \cdot 10^7$ cells. Centrifugation was performed (30 min, 800 g, 4 °C) in a swinging bucket on a Beckman Coulter JS-13.1 rotor (Beckman Coulter, USA). Hemocytes concentrated at each density interface were collected separately with a 70 mm long 20Gx2.75” needle mounted on a 1 mL syringe. The hemocytes were then washed from the Percoll by adding 10 mL of ice-cold Alsever modified medium, pelleted by centrifugation (10 min, 200 g, 4°C) and resuspended in Alsever modified medium or filtered seawater.

Cytological description of the hemocyte populations

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

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

Total RNA was extracted using the RNeasy kit (Qiagen, the Netherlands) and cDNA was synthesized from 1 µg total RNA using the Superscript IV kit (ThermoFisher Scientific, USA) with oligo(dT) primers. RT-PCR was performed on LightCycler© 480 thermocycler using the SYBR Green 1 Master kit (Roche, Switzerland). Primers were used at 200 nM. Primer sequences are listed in **Supplementary Table S3**. Expression was normalized to *Cg-rps6* reference gene. The standard cycling conditions were 95°C pre-incubation for 3 minutes followed by 40 cycles of 95°C for 10 seconds, 60°C for 20 seconds, and 72°C for 25 seconds. Each pair of primers was first tested and validated on a total hemocyte RNA preparation to control melting curves and establish calibration lines.

Oxidative Burst Assay

The production of reactive oxygen species was quantified by luminescence assay. Briefly, hemocytes from hemolymph puncture or Percoll density gradient were washed once with filtered sterile water. 50 µL of hemocytes were plated in triplicate on a 96-well plate ($3 \cdot 10^5$ cells/cm²). 50 µL of 40 mM luminol (Merck and Co, USA) was added to each well. After 45 minutes of incubation at room temperature, the oxidative burst was induced by adding 100 µL of zymosan (Merck and Co,

USA) at a multiplicity of infection (MOI) of 50:1. The plate was immediately placed in a Berthold Centro XS3 LB 960 luminescence microplate reader (Berthold GmbH, Germany) to measure luminescence emission every 2 minutes for 2 hours.

Phagocytosis assay

400 μL of filtered, sterile water-washed hemocytes were seeded in a 24-well plate at a concentration of 3.10^5 cells/ cm^2 . After 15 minutes, 50 μL of zymosan was added to the fractions at an MOI of 20:1. For LMG20012^T *Vibrio*, 50 μL of bacteria was added to a total hemolymph at an MOI of 5:1. After 1 hour of contact at room temperature, the cells were resuspended and 200 μL of the suspension was applied to a microscope slide using a Cytospin centrifuge. The slides were then stained with MCDH and observed under a LEICA DMR transmitted light microscope with a 40x magnification objective (Leica AG, Germany) to count phagocytic cells.

NitroBlueTetrazolium (NBT) staining

ROS production was measured using NitroBlue Tetrazolium reduction after zymosan stimulation. Briefly, 1 mL (1.10^6 cells) of hemocyte solution was mixed with 50 μL of filtered NBT solution (15 mg/mL in water). Zymosan was added at a 4:1 MOI and the mixture was incubated at room temperature for 10 minutes on a rocking shaker. Then, 50 μL of each sample was plated onto glass coverslips and observed under a transmitted light microscope LEICA DMR with a 40x magnification objective (Leica AG, Germany) to count NBT-positive cells. The positions of positive NBT cells were recorded prior to MCDH staining to identify hemocyte types.

Rhodanine copper staining of hemocytes

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

Pseudotemporal ordering of cells with Monocle3

Cells from the *C. gigas* dataset were analyzed using Monocle3 (<https://github.com/cole-trapnell-lab/monocle3>) (37). The Monocle3 analysis was performed on the Seurat object following the aforementioned processing steps. Clustering information (features, genes, partitions, clusters and UMAP coordinates) was transferred to a CDS object. The cell trajectory was calculated using the *learn_graph* function. The *choose_graph_segments* function was used to select three lineages. The gene expression along pseudotime data was extracted from the result. Then, the data were used to plot genes along pseudotime in three lineages using ggplot2 v3.4.4 R package and the heatmap was generated using the pheatmap v1.0.12 R package.

Statistical Analysis

To evaluate differences between samples, a statistical analysis was performed using (version 4.2.1) (R Core Team, 2018) and appropriate packages. All data were examined for normality, and statistical tests were selected accordingly. One-way analysis of variance (ANOVA) was used for

normally distributed data. Seven different hemolymph samples were used for cytological analysis. Oysters were provided by our local supplier) and approximately 300 hemocytes were counted per sample. Seven independent experiments were performed for Percoll density gradient separation. A Tukey test was used to evaluate the statistical difference between the proportions of hemocyte types. The phagocytic capacity of hemocytes was tested in three independent experiments and statistical differences were evaluated using the Tukey test for both the phagocytic capacity between hemocyte types and the number of particles per phagocyte. Finally, oxidative burst capacity was tested 3 times on Percoll-separated hemocytes. The Tukey test was also used to assess statistical differences between conditions. The null hypothesis was rejected at a significance level of $p = 0.05$.

Acknowledgements

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.

Supplementary Materials

Data S1. Single-cell RNA-seq analysis result file. CSV file containing the scRNA-seq analysis results with cluster number (<cluster>), gene number (<gene>), chromosome reference where the gene is located (<chromosome>), average expression level (<avg_log2FC>), percentage in the cluster (<pct1>) and in other clusters (<pct2>), pct1 / pct2 ratio (<pct. ratio>), adjusted p-val and p-val values (<p-val> and <p_val_adj>), and gene description (<description>).

Data S2. Annotation file of CDS extracted from the *Crassostrea gigas* genome file. CSV file containing the annotation results collected and processed using Blast2GO for annotation mapping and validation.

Data S3. Compilation file for RBGOA results. CSV file containing concatenated results of all RBGOA tests used to draw gene ontology analysis heatmaps with cluster number (<cluster_number>), GO term universe (<goterm_universe>), number of good gene candidates (<number_of_good_candidates>), total number of genes in this GO term category (<total_number_of_genes_of_this_category>), GO term name (<goterm_name>), adjusted p-value (<pval-adj>) and GO term name variation (<variation>).

Number of Reads	127,959,215
Reads With Valid Barcodes	97.03 %
Sequencing Saturation	75.63 %
Q30 Bases in CB+UMI	96.14 %
Q30 Bases in RNA read	93.62 %
Reads Mapped to Genome: Unique+Multiple	89.22 %
Reads Mapped to Genome: Unique	82.01 %
Reads Mapped to Gene: Unique+Multiple Gene	72.26 %
Reads Mapped to Gene: Unique Gene	72.26 %
Estimated Number of Cells	2937
Unique Reads in Cells Mapped to Gene	67,950,796
Fraction of Unique Reads in Cells	73.48 %
Mean Reads per Cell	23,136
Median Reads per Cell	18,145
UMIs in Cells	16,306,374
Mean UMI per Cell	5552
Median UMI per Cell	4412
Mean Gene per Cell	1578
Median Gene per Cell	1434
Total Gene Detected	23,841

Table S1.

STARsolo summary metrics report.

Metrics dashboard obtained after the STARsolo step, describing the quality of the sequencing and the various characteristics of the cells detected after aligning the reads to the *C. gigas* genome from the Roslin Institute.

Cluster	KEGG Prefix	KEGG Number	KEGG Pathway	Fold Enrichment	P-value (<0.05)	Genes
1	crg	04144	Endocytosis	3.75	1.77E-06	G32465, G28245, G29411, G28554, G23342, G27546, G4677, G4248, G23299, G4679, G29017, G27834, G23833, G20650, G1765, G2151, G15227, G11863, G17437, G11958, G17273
1	crg	00030	Pentose phosphate pathway	10.19	4.25E-05	G20983, G16716, COX2, COX1, ND1, CYTB, G2226, G3525, G3658, ND5
1	crg	00190	Oxidative phosphorylation	4.45	3.29E-04	G30122, G18442, G34645, G20982, G31268, G16805, G25073, G11958, G17273, G7839
1	crg	01200	Carbon metabolism	3.54	1.77E-03	G34645, G20982, G11958, G23780, G7839, G13028
1	crg	00010	Glycolysis / Gluconeogenesis	4.94	6.58E-03	G18442, G22426, G22427, G23780, G7839
1	crg	00620	Pyruvate metabolism	5.26	1.38E-02	G22034, G32778, G32842
1	crg	03266	Virion - Herpesvirus	8.74	4.39E-02	G5571, G25192, G5077, G3494, G33210, G26985, G26901, G23358, G1703, G1900, G1543, G2172, G12279, G5185, G26391, G10525, G4452, G11817, G25420, G2853, G4635, G3304, G25620, G5649, G3526, G740, G24712, G26976, G25626, G15236, G34618, G34615, G18702
2	crg	03010	Ribosome	11.87	5.45E-27	G32475, G24506, G4572, G3088, G1531, G15286, G1299, G153, G22991, G1471, G10110, G1470
2	crg	00980	Metabolism of xenobiotics by cytochrome P450	4.61	2.17E-02	G5106, G26659, G18623, G18624, G14458
3	crg	00190	Oxidative phosphorylation	8.03	1.47E-12	G31970, G3073, G5034, G3298, G5653, G27278, G25266, G650, G331, G23645, G1779, G16296, G23701, G13193, G27816, G10181, G15131, G21371, G2133, G34736
3	crg	01200	Carbon metabolism	5.42	4.82E-08	G26194, G5034, G4037, G3500, G20764, G16296, G23701, G14096, G22912, G15171, G27816, G21794, G19054, G15078, G16455, G34649, G9692
3	crg	00010	Glycolysis / Gluconeogenesis	7.41	5.01E-06	G26194, G14096, G22912, G21794, G19054, G15078, G4037, G3500, G34649, G12763
3	crg	01230	Biosynthesis of amino acids	5.77	1.41E-05	G26194, G14096, G22912, G21794, G4037, G25300, G34649, G7042, G25304, G9692, G20791
3	crg	04144	Endocytosis	3.2	3.22E-05	G10612, G34830, G4232, G31680, G28015, G28554, G2650, G35288, G23530, G22892, G3216, G2800, G22524, G25627, G21676, G20683, G1765, G11863
3	crg	00020	Citrate cycle (TCA cycle)	8.23	1.48E-04	G16296, G23701, G5034, G15171, G27816, G19054, G16455
3	crg	01100	Metabolic pathways	1.27	1.48E-02	G13605, G5034, G3012, G3298, G650, G331, G22961, G5758, G14096, G27816, G13000, G2033, G4037, G3500, G34394, G23387, G25763, G20764, G1779, G23701, G13193, G21533, G15171, G19054, G15131, G13175, G21371, G12861, G9692, G3073, G8883, G31970, G30520, G25177, G27276, G5653, G16296, G2133, G16455, G34649, G7042, G9342, G26194, G23482, G6556, G23164, G5106, G25300, G25266, G25304, G23645, G25904, G22912, G21794, G15078, G10181, G22919, G16488, G34736, G12763, G2280
3	crg	00250	Alanine, aspartate and glutamate metabolism	4.26	2.79E-02	G23164, G25300, G25763, G25304, G20764
3	crg	00430	Taurine and hypotaurine metabolism	5.05	4.23E-02	G23482, G34394, G25763, G9342
4	crg	03010	Ribosome	11.18	2.38E-62	G32202, G25192, G4582, G5077, G3494, G33210, G32255, G5115, G5555, G11948, G31071, G23694, G4502, G732, G1703, G1900, G21588, G18212, G1543, G28190, G12279, G13687, G28274, G10525, G7023, G27586, G3387, G11817, G26894, G4635, G3304, G2853, G24036, G5649, G24712, G26976, G1172, G12546, G27592, G11965, G5571, G4085, G5375, G5652, G23471, G32582, G26985, G26901, G23358, G18432, G16773, G21360, G10473, G2172, G35349, G5185, G26391, G4452, G4972, G3885, G25420, G25620, G3526, G740, G2636, G25626, G16242, G15236, G34618, G1350, G34615, G18702
4	crg	03050	Proteasome	9.09	3.18E-14	G11447, G11878, G9178, G28086, G6441, G27078, G6503, G27105, G4954, G21626, G1667, G2517, G11085, G13489, G32846, G1140, G29381, G12565, G30769, G34613
4	crg	03040	Spliceosome	4.6	1.62E-13	G30069, G23072, G3035, G27237, G4644, G67, G3357, G28826, G21225, G22448, G21463, G20152, G31757, G12256, G16599, G2099, G16114, G11265, G19821, G5284, G4177, G23063, G3388, G35242, G23100, G2896, G5803, G21555, G22119, G18742, G13575, G14401, G10466
4	crg	00190	Oxidative phosphorylation	3.14	3.88E-04	G31970, G3077, G3298, G26204, G34551, G11926, G27610, G650, G2799, G17080, G1779, G1500, ND3, G31989
4	crg	03015	mRNA surveillance pathway	3.18	1.10E-03	G21555, G1768, G21245, G23072, G3035, G1892, G27237, G22452, G12561, G11265, G19821, G15119
4	crg	03013	Nucleocytoplasmic transport	2.44	9.12E-03	G21555, G1768, G23072, G4818, G3035, G2485, G31807, G27237, G12561, G11265, G19821, G27008
4	crg	04141	Protein processing in endoplasmic reticulum	2.06	1.28E-02	G11932, G29430, G10911, G22082, G26491, G29962, G2451, G2694, G6732, G4985, G20713, G204, G20332, G2609, G10994
5	crg	03010	Ribosome	7.41	1.68E-12	G27592, G32202, G5185, G5375, G33210, G27586, G11948, G26985, G3885, G26894, G3304, G2853, G18212, G18432, G16773, G1543, G16242, G21360, G13687, G2172, G12279
5	crg	00190	Oxidative phosphorylation	4.59	6.31E-04	G3430, COX2, COX1, ND1, CYTB, ND3, ND2, ND5, ND4
5	crg	04137	Mitophagy - animal	4.65	8.54E-03	G30487, G25784, G29415, G3710, G22696, G12502
6	crg	03010	Ribosome	16.48	2.26E-55	G32202, G25192, G4582, G3494, G5115, G5555, G11948, G4502, G732, G1703, G21588, G18212, G1543, G28190, G12279, G13687, G10525, G7023, G27586, G3387, G11817, G26894, G4635, G3304, G2853, G24036, G5649, G24712, G26976, G27592, G5571, G4085, G5652, G26985, G26901, G23358, G18432, G21360, G10473, G2172, G5185, G26391, G3885, G25420, G25620, G740, G2636, G25626, G16242, G15236, G34618, G34615, G18702
7	crg	04144	Endocytosis	4.04	4.91E-09	G10612, G23299, G27538, G22524, G12273, G16951, G2151, G3185, G6211, G13317, G2450, G31680, G4232, G28015, G2650, G35288, G23342, G20567, G17254, G18762, G19412, G21991, G1774, G1993, G30759
7	crg	04137	Mitophagy - animal	4.99	3.50E-04	G18587, G25784, G11088, G2334, G30759, G3710, G27737, G30997, G12500
7	crg	04145	Phagosome	2.82	1.88E-03	G3185, G22181, G3430, G28015, G24310, G17795, G14056, G2026, G1780, G30759, G32778, G32842, G90
7	crg	04140	Autophagy - animal	2.97	3.47E-03	G20344, G5397, G25608, G25784, G11385, G2024, G11088, G30759, G10907, G21965, G12500
7	crg	04310	Wnt signaling pathway	2.9	6.82E-03	G24405, G23180, G11208, G35235, G16231, G2024, G14136, G2334, G1780, G2797
7	crg	04142	Lysosome	1.92	4.51E-02	G11645, G5397, G22181, G3430, G2026, G27524, G18705, G26324, G16202, G2976, G27526, G23337

Table S2.

Table presenting the result of KEGG analysis performed using DAVID Bioinformatics Resources.

GO term enrichment analysis was conducted on specifically overexpressed genes in each cluster obtained after scRNA-seq processing (genes with Log2FC > 0.25 and significant p-value < 0.001) to highlight the most relevant GO terms associated with a given gene list. The visualization of the different pathways can be obtained from the KEGG website using the KEGG prefix and KEGG number (<https://www.genome.jp/kegg/pathway.html>)

Target	Primer name	Sequence
Laccase 24	LACC24-F	CCT-TGA-TTC-TTC-TTG-CCA-TCC-G
	LACC24-R	AAA-GCT-TGC-GAT-CTT-TGG-CAA
C-type lectin domain-containing protein	CLEC-F	ATC-GGC-TTC-TAC-ATG-GAC-TGA-C
	CLEC-R	GTG-TCT-AAA-GCT-GCG-CCG-AT
Putative modulator of levamisole receptor-1	LEVAR-F	GTG-ACA-GAC-TTC-CCT-CAC-CCT
	LEVAR-R	GCA-CTG-AGT-CGA-GTC-GTA-TGT
EGF-like domain-containing protein 8	EGFL-F	GAG-TGT-TTG-ACA-GGA-CGA-AGC
	EGFL-R	CAT-CAT-CGT-TTC-CAA-CTG-AGG-C
X-Box binding protein-like domain	XBOX-F	GGG-TCA-ACA-GTG-CTA-GGC-AAT
	XBOX-R	GTA-AGC-CAC-CAT-CCC-TAC-CAC
Transglutaminase-like domain containing protein	TGC-F	CTA-CAA-GCT-GGA-CAC-CAC-CAA
	TGC-R	GCA-TTG-ACC-AGT-GAC-ACA-GTC
MD-2-related lipid-recognition domain containing protein	MLDP-F	CTT-GGA-CCT-CGT-TAT-CTT-CGC
	MLDP-R	CTC-CCT-CTG-GTC-CAC-AAA-CAA
High mobility group protein B1	HMGB1	GCC-CAC-GCT-GAA-CTA-TAC-AAG
	HMGB1	CAC-CTT-GTA-GTC-CCT-GAG-TGG
Galectin	GAL-F	CCA-CAG-TAT-CAA-CGA-CCC-TCC
	GAL-R	TCA-CTA-CCG-TCA-TAG-GGA-CCG
Cubilin	CUBN-F	TAA-GTT-CAC-TCT-GGC-CCA-AGG
	CUBN-R	GCT-CAT-GAT-CGT-AGT-GGT-GCT
Natterin-1	NAT1-F	CCG-TAC-GAT-GGT-GAG-GAG-AAA
	NAT1-R	CCG-TCC-CAC-ATA-CTT-GTC-GTT
Caveolin	CAV-F	TCC-AAA-TGA-CCA-TGA-CCC-AGA
	CAV-R	ACT-CTA-TTC-TTG-GTC-GCC-TGG
G Protein receptor F1-2 domain-containing protein	GPROT-F	CGA-ACG-CCT-GCT-TCT-GAT-ATG
	GPROT-R	TCC-ACA-TCG-AAT-GCT-CTG-TCT
DBH-like monooxygenase protein 1	MOX-F	CCT-CCG-CAG-CAA-GAA-GAA-GTA
	MOX-R	TTC-TGT-TTC-GTC-CTC-TCC-ACG
Ribosomal protein S6	RPS6-F	CAG-AAG-TGC-CAG-CTG-ACA-GTC
	RPS6-R	AGA-AGC-AAT-CTC-ACA-CGG-AC
Big Defensin 1&2	BigDef1-2-F	TTC-GCC-TGC-TTC-CAT-ACT-GG
	BigDef1-2-R	GTC-ATG-GTC-ACT-CCT-TAT-TC
Hemocyte defensin	HemDef-F	CTA-CCA-GTT-GTT-CAT-ACA-GAG
	HemDef-R	TCT-TGG-TCA-GAT-TCA-GTC-TGG
Bactericidal Permeability Increasing Protein	BPI	GGA-GGC-GGA-AAT-GGA-TTA-CT
	BPI	TGG-TTG-ACA-TCG-TTG-CTG-AC

Table S3.

Sequences of primers used in this study.

Name of the transcript targeted by the primer pair, name of the primer used in this study and nucleotide sequence of each primer.

Fraction	Cells	Average	Std.Dev.	Std.Error	Min	Max
1	H	46.43	9.47	4.73	37.50	59.63
1	ML	20.26	9.09	4.55	7.34	28.38
1	BBL	14.47	2.29	1.15	12.16	16.50
1	ABL	15.50	4.94	2.47	8.11	18.35
1	G	0.35	0.44	0.22	0.00	0.92
1	BGC	0.38	0.75	0.38	0.00	1.50
1	VC	2.62	1.97	0.98	0.92	5.41

Fraction	Cells	Average	Std.Dev.	Std.Error	Min	Max
2	H	20.17	5.26	2.63	14.10	26.56
2	ML	27.64	4.48	2.24	22.79	33.33
2	BBL	27.21	9.94	4.97	16.22	39.71
2	ABL	20.17	2.36	1.18	16.91	22.52
2	G	1.14	0.79	0.39	0.00	1.80
2	BGC	1.03	1.26	0.63	0.00	2.56
2	VC	2.64	1.56	0.78	0.74	4.50

Fraction	Cells	Average	Std.Dev.	Std.Error	Min	Max
3	H	12.49	3.78	1.89	10.08	18.04
3	ML	28.26	8.90	4.45	17.12	38.66
3	BBL	21.18	3.11	1.55	16.89	24.32
3	ABL	17.71	5.78	2.89	11.86	25.21
3	G	7.96	5.30	2.65	3.09	14.41
3	BGC	4.71	4.16	2.08	0.00	10.14
3	VC	7.68	2.87	1.43	4.20	10.81

Fraction	Cells	Average	Std.Dev.	Std.Error	Min	Max
4	H	7.82	5.97	2.99	2.20	14.29
4	ML	32.69	13.33	6.66	19.34	50.82
4	BBL	9.64	3.40	1.70	4.76	12.30
4	ABL	5.37	0.97	0.48	4.42	6.67
4	G	18.29	10.39	5.19	5.74	30.94
4	BGC	3.90	1.51	0.76	2.20	5.71
4	VC	22.29	8.25	4.12	11.48	30.77

Fraction	Cells	Average	Std.Dev.	Std.Error	Min	Max
5	H	4.29	4.41	2.21	0.76	10.45
5	ML	19.67	12.04	6.02	3.03	30.88
5	BBL	3.56	1.91	0.95	1.49	6.06
5	ABL	3.35	2.12	1.06	0.75	5.88
5	G	40.92	7.31	3.66	33.82	49.24
5	BGC	4.50	2.98	1.49	1.47	8.33
5	VC	23.71	4.48	2.24	17.91	28.79

Fraction	Cells	Average	Std.Dev.	Std.Error	Min	Max
6	H	3.57	3.82	1.91	0.65	9.09
6	ML	13.95	4.91	2.45	7.10	18.18
6	BBL	1.19	1.38	0.69	0.00	2.58
6	ABL	1.54	0.65	0.33	0.65	2.17
6	G	68.00	10.11	5.06	58.70	78.46
6	BGC	1.69	1.99	1.00	0.00	3.87
6	VC	10.06	5.30	2.65	3.08	15.94

Fraction	Cells	Average	Std.Dev.	Std.Error	Min	Max
7	H	2.43	3.04	1.52	0.00	6.31
7	ML	8.50	6.35	3.18	1.93	17.05
7	BBL	1.67	1.51	0.75	0.00	3.38
7	ABL	0.36	0.72	0.36	0.00	1.45
7	G	81.18	8.99	4.49	68.18	88.41
7	BGC	2.68	1.50	0.75	0.90	4.55
7	VC	3.18	2.92	1.46	0.00	6.82

Table S4.

Hemocyte composition of the 7 Percoll fractions used for qPCR analysis.

For each cell type in each fraction, the table presents the average percentage, standard deviation, standard error, minimum, maximum, and median count values.

Gene Number	Annotation Name	TF Name	Blast alignment results						Homo sapiens Homologs	
			Best Hit Blast Name	Specie	Accession number	Bit Score	Percentage Identity	Coverage	Name	Uniprot Number
G29966	Kruppel-like factor 15	CgKLF6	Kruppel-like factor 6	C.gigas	XP_011428991	659.448	100	100	HsKLF8	Q95600
G30997	BHLH domain-containing protein	CgTFEB	Transcription factor EC isoform X1	C.gigas	XP_011447558	1089.72	100	100	HsTFEB	P19484
G17147	BZIP domain-containing protein	CgCR3L1	cAMP-responsive element-binding protein 3-like protein 1 isoform 1	H.sapiens	NP_443036	1060.06	100	100	HsCR3L2	Q705Y1
G3043	Transcriptional regulator	CgMYC	Transcriptional regulator Myc-A-like	C.angulata	XP_052685275	548.128	100	100	HsMYC	P01106
G31054	GATA-binding factor 3	CgGATA3	Transcription factor GATA-3 isoform X1	C.gigas	XP_011412760	966.066	100	100	HsGATA2	P23769
G11196	BZIP domain-containing protein	CgFOS	Proto-oncogene c-Fos	C.gigas	XP_011446784	664.84	100	100	HsATF3	P18847
G10637	BZIP domain-containing protein	CgFOS2	Fos-related antigen 2	C.gigas	XP_011439871	271.937	100	100	HsATF4	P18848
G2123	SPRY domain-containing SOCS box protein 3	CgSOCS3	SPRY domain-containing SOCS box protein 3	C.gigas	XP_011413771	488.804	100	100	HsSPSB4	Q96444
G1708	SoxE	CgSOX8	Transcription factor Sox-8-like	C.gigas	NP_001295801	947.192	99.8	100	HsSOX9	P48436
G3506	BZIP domain-containing protein	CgATF5	cAMP-dependent transcription factor ATF-4	C.gigas	XP_011439092	566.614	85.5	100	HsATF5	Q9Y2D1
G11198	BZIP domain-containing protein	CgFOS3	Fos-related antigen 1	C.gigas	XP_034323853	456.062	100	99.62	none	none
G10636	BZIP domain-containing protein	CgATF1	Basic leucine zipper transcriptional factor ATF-like isoform X1	C.gigas	XP_011439867	270.011	100	99.41	none	none
G13555	BZIP domain-containing protein	CgJUNB	Transcription factor jun-B isoform X1	C.gigas	XP_011440361	224.557	100	99.53	none	none
G21091	X-box binding protein like protein	CgXBDX	X-box binding protein-like protein	C.arsakensis	AEF33390	375.491	80.5	99.12	none	none
G12164	Estrogen receptor	CgRXRB	Estrogen receptor beta	C.gigas	XP_011424805	1395.56	100	98.79	HsRXRA	P19793
G748	Interferon regulatory factor	CgIRF1	Interferon regulatory factor 1 isoform X1	C.gigas	XP_011449592	802.742	100	95.51	HsIRF1	P10914
G5204	ETS-related transcription factor Eif-3	CgELF3	ETS homologous factor isoform X1	C.gigas	XP_034310128	883.248	100	95.25	HsELF3	P76545
G26368	ETS domain-containing protein	CgELK3	ETS domain-containing protein Elk-3 isoform X1	C.gigas	XP_011421195	753.051	100	93.91	HsELK3	P41970
G35512	cAMP-responsive element modulator	CgCREM	cAMP-responsive element modulator isoform X2	C.gigas	XP_011448206	572.778	100	92.79	HsCREB1	P16220
G11013	BZIP domain-containing protein	CgTEF	Thyrotroph embryonic factor-like	C.virginica	XP_022332233	570.466	86.4	85.1	HsTEF	Q10587
G7003	ETS-related transcription factor Eif-4	CgELF2	ETS-related transcription factor Eif-2	C.gigas	XP_011447491	523.857	99.6	94.54	HsELF1	P32519
G27920	LIM domain transcription factor LMO4.1	CgLMO4	LIM domain transcription factor LMO4.1 isoform X2	C.gigas	XP_011443581	319.316	89.2	81.86	HsLMO4	P81968
G10665	BZIP domain-containing protein	CgVBP	Transcription factor VBP	C.gigas	XP_011457057	438.343	100	77.95	HsHLF	Q16534
G36439	BZIP domain-containing protein	CgCEBPG	CCAAT/enhancer-binding protein beta	C.gigas	XP_011452433	442.195	100	75.61	HsCEBPG	P35367
G2334	Proto-oncogene c-Jun	CgJUN	Transcription factor AP-1	C.gigas	XP_034313954	429.098	100	58.31	HsJUN	P05627
G35127	Kruppel-like factor 5	CgKLF5	Kruppel-like factor 5 isoform X1	C.gigas	XP_011421316	905.975	100	57.28	HsKLF7	Q75840
G20682	cAMP-responsive element-binding protein-like 2	CgCRBL1	cAMP responsive element binding protein-like, partial	C.gigas	AALB3879	319.546	99.4	56.83	HsCRBL2	O60519
G17284	Runt-related transcription factor 1	CgRUNX1	Runt-related transcription factor 1-like isoform X5	C.angulata	XP_052711215	1085.86	100	56.68	HsRUNX3	Q13761
G10077	BZIP domain-containing protein	CgCR3L4	Clumping factor A isoform X1	C.gigas	XP_011425092	775.393	100	56.11	HsCR3L3	Q68CJ9
G2021	BHLH domain-containing protein	CgTAL1	T-cell acute lymphocytic leukemia protein 1-like isoform X1	C.gigas	XP_011439324	394.045	100	49.23	HsTAL2	Q16599
G2933	Putative transcription factor SOX-14	CgSOX11	Transcription factor Sox-11	C.gigas	XP_011445203	424.091	100	43.97	HsSOX12	O15370
G11358	BZIP domain-containing protein	CgGIANT	Protein giant	C.gigas	XP_011417277	486.108	100	38.41	HsTEF	Q10587
G6354	ETS domain-containing protein	CgPU.1	ETS-related transcription factor Eif-3	C.gigas	XP_034320185	189.119	100	36	HsPU.1	P17947
G10633	BZIP domain-containing protein	CgFOS1	Fos-related antigen 1	C.gigas	XP_019918193	371.719	99.5	31.4	HsFOSL1	P15407
G6394	ETS domain-containing protein	CgSPDEF	SAM pointed domain-containing Ets transcription factor-like isoform X2	C.gigas	XP_052897855	356.681	99.4	30.12	HsSPDEF	O95238
G32716	cAMP-responsive element-binding protein-like 2	CgCRBL2	cAMP-responsive element-binding protein-like 2 isoform X1	C.gigas	XP_034309665	297.745	83.4	26.7	HsCRBL2	O60519

Table S5.

Transcription factors identified in the scRNA-seq dataset of *Crassostrea gigas* hemocytes.

Transcription factors identified in the scRNA-seq dataset and their homology with human proteins, as indicated by the blast alignment results. Each entry includes the gene identifier, the protein it represents in *C.gigas*, and its human counterpart. The bit score and coverage indicate the strength and extent of the alignment, respectively.

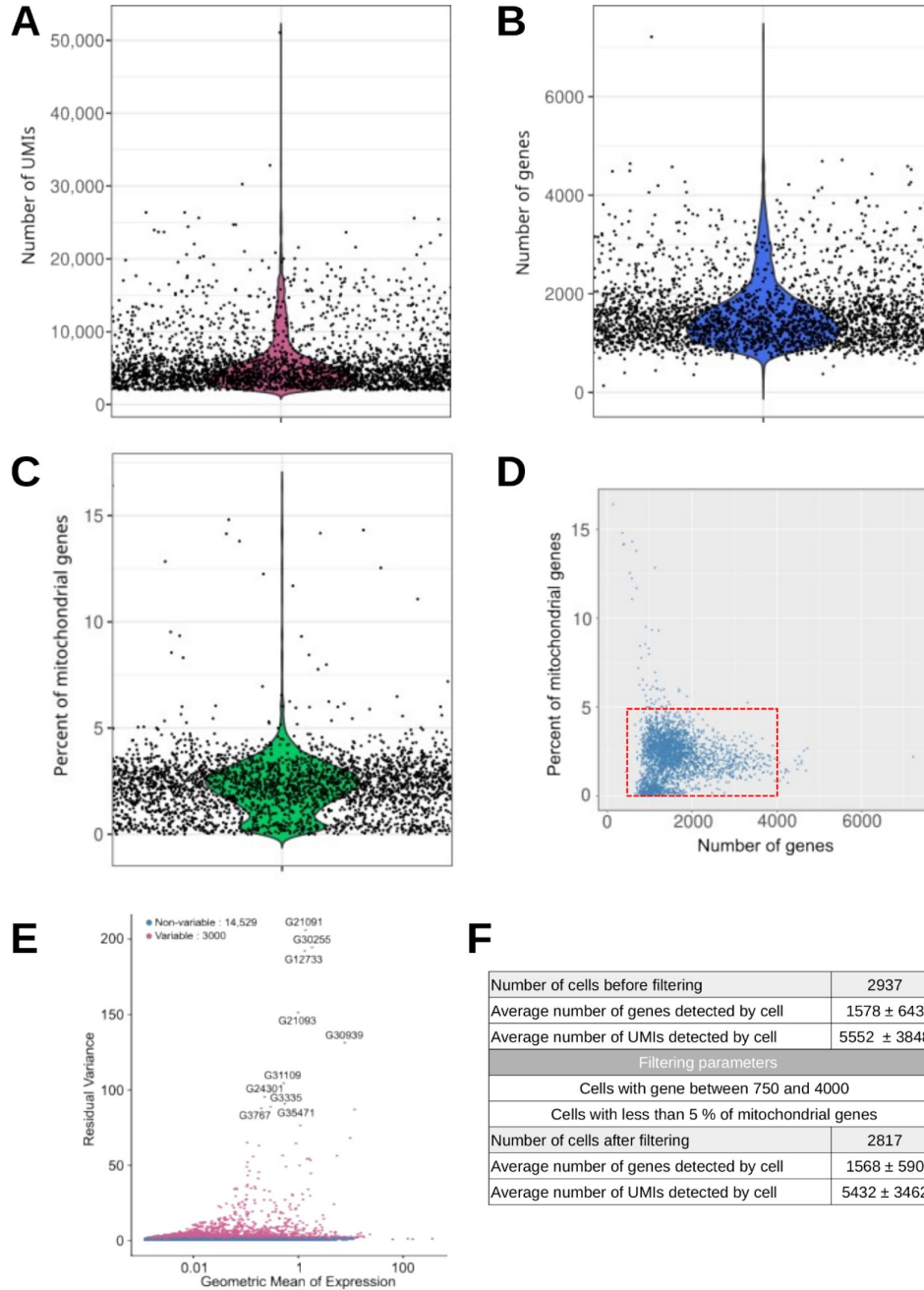


Figure S1.

ScrRNA-seq quality control metrics.

Distribution of **A**) unique molecular identifiers (UMIs), **B**) genes and **C**) percentage of mitochondrial genes detected per cell. Each dot represents one cell. **D**) Plot of the percentage of mitochondrial genes versus the number of genes detected in each cell. The red box represents the cells selected for further analysis (number of genes detected between 750 and 4000 and with a percentage of mitochondrial genes less than 5%) **E**) The *FindVariableFeatures()* function was used to identify features with high cell-to-cell variation in the dataset to highlight the biological signal in the single cell dataset. **F**) Table summarizing some quality control metrics. The table shows the thresholds to remove poor quality cells (doublets or empty droplets). The number of cells, UMIs and genes before and after filtering are shown. **G**) Uniform Manifold Approximation and Projection (UMAP) plot of the cells with the number of expressed genes. **H**) UMAP plot of the number of reads per cell. **I**) UMAP plot of the percentage of mitochondrial genes in each cell. **J**) UMAP graph of the percentage of ribosomal protein transcripts in each cell.

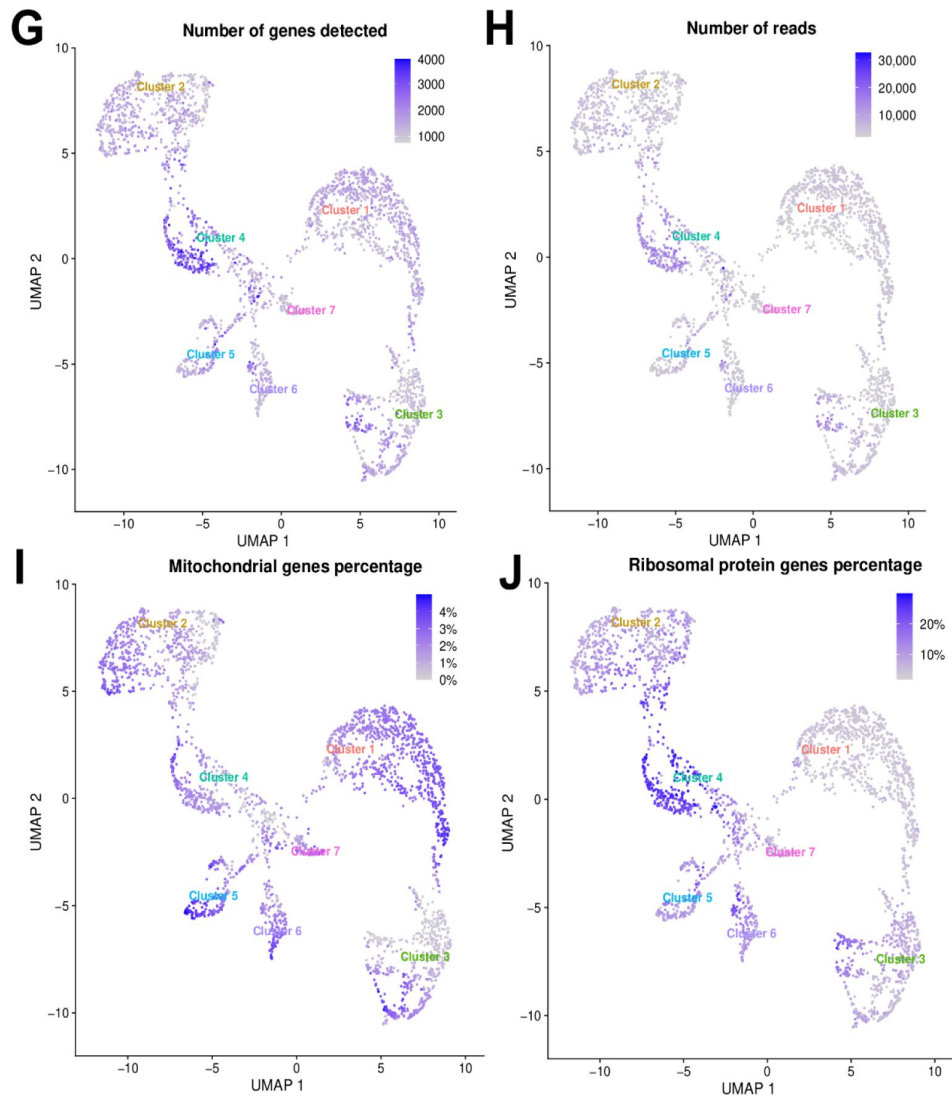
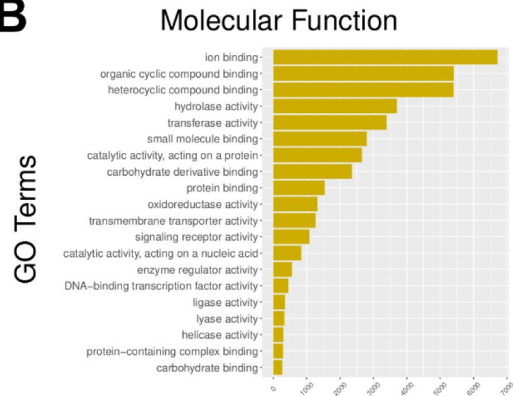
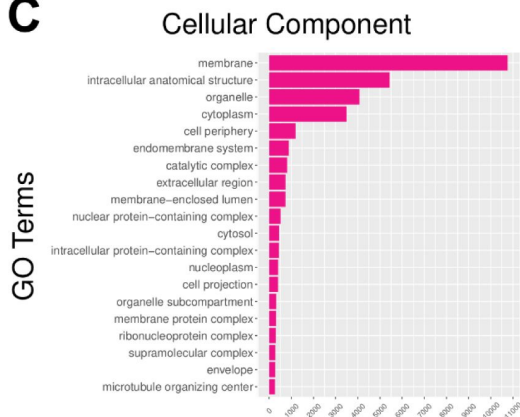
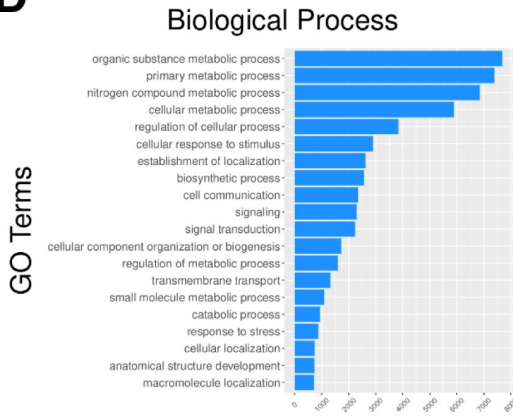


Figure S1. (continued)

A

Stage	Number of CDS	% of CDS
Input	30,724	100
Database querying		
Blast annotated	30,511	99.3
Interpro annotated	30,718	99.9
EggNogg annotated	12,630	41.1
Blast2Go		
Mapped	28,570	92.9
Annotated	22,462	73.1

B**C****D****Figure S2.****Results of the *C. gigas* genome re-annotation.**

A) The number and percentage of Coding DNA Sequences (CDS) with valid annotation after each annotation step are shown. A BLAST query was performed against the TrEMBL/Uniprot database and InterproScan annotation against Pfam, PrositeSiteProfiles, CDD, TIGRFAM, PRINTS, SMART, SUPERFAMILY and Hamap databases. Blast and InterProScan results were compiled and processed using Blast2Go. A first mapping step was used to enrich the Blast result with GO-terms, and the annotation step was used to optimize and validate the GO-terms annotations. **B) C) and D)** show the distribution of the various categories of GO-terms across the three primary domains of Gene Ontology : Molecular Function, Cellular Component and Biological Process, respectively.

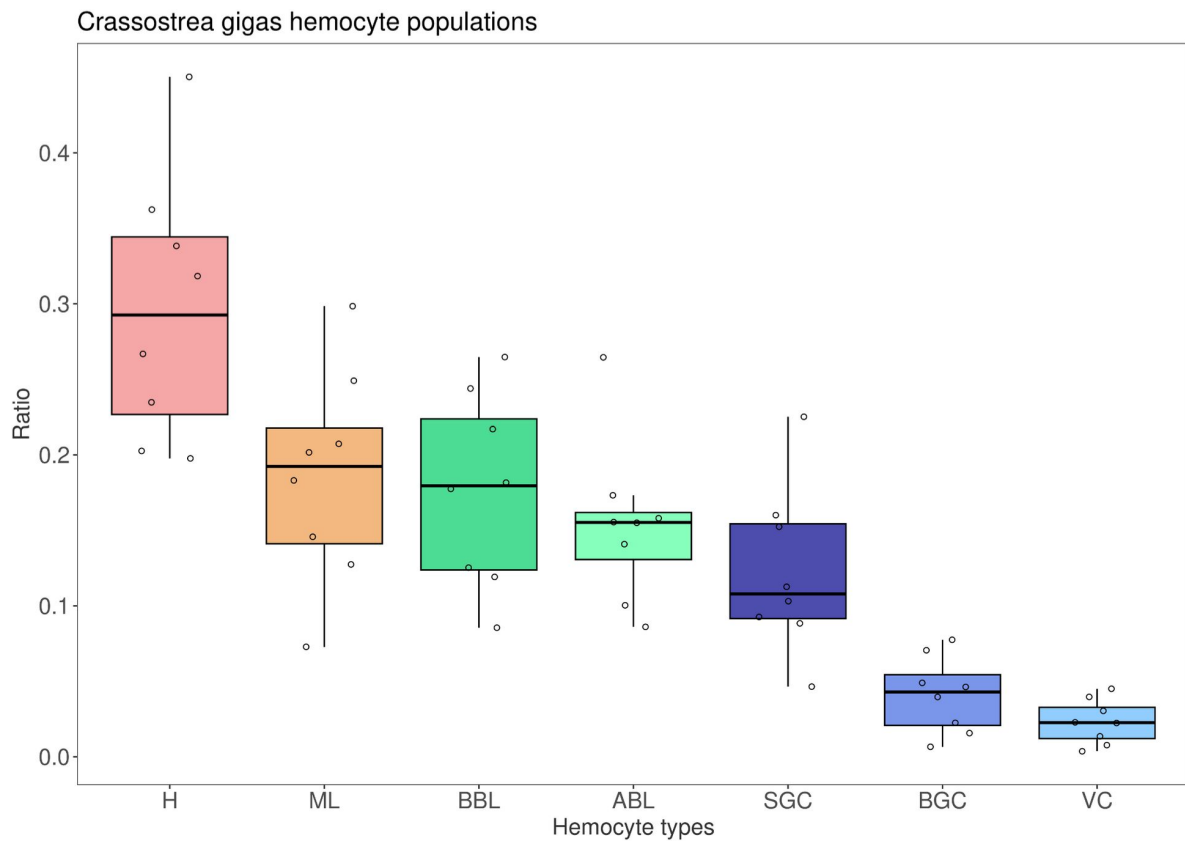


Figure S3.

Whisker plot showing the distribution of counts of different hemocyte populations.

Hemolymph was collected from eight different oysters and hemocytes were plated on a slide using cytopspin centrifugation. The proportion of each hemocyte type in these hemolymphs was calculated after MCDH staining. **H** : Hyalinocytes, **ML** : Macrophage-Like cells, **BBL** : Basophilic Blast-Like cells, **ABL** : Acidophilic Blast-Like cells, **SGC** : Small Granule Cells, **BGC** : Big Granule Cells, and **VC** : Vesicular Cells. The calculation of the proportion of each cell type in these hemolymphs reveals a heterogeneous composition of hemocytes between individuals as well as between cell types.

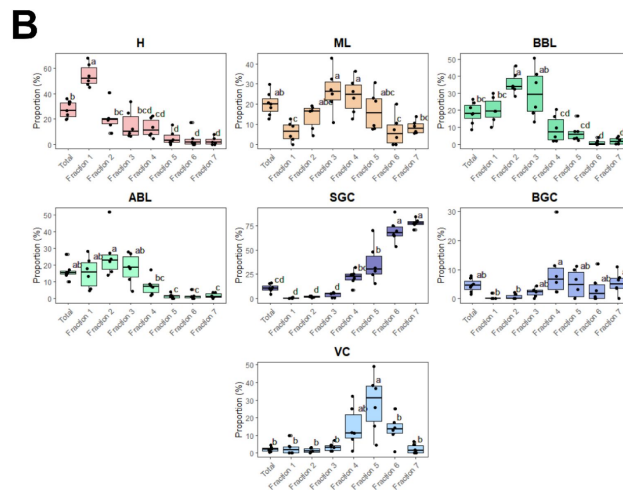
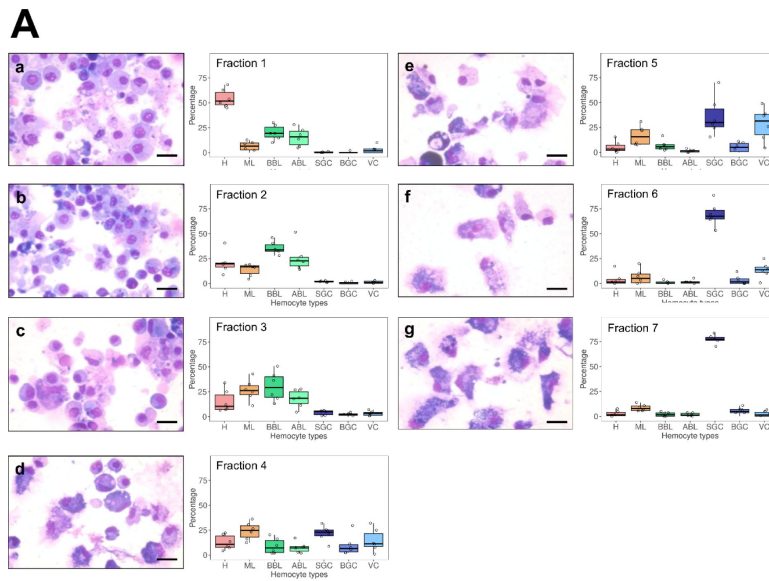


Figure S4.

Statistical significance of enrichment in different hemocyte types in Percoll gradient fractions.

(A) **a,b,c,d,e,f, and g** : MCDH staining of cells from the 7 fractions isolated from the percoll gradient in (B). Scale bar : 10 μ m. **Fraction 1 to 7** : Quantification of the different types of hemocytes found in each of the 7 fractions from 5 independent fractionation experiments. (B) Statistical significance of enrichment of different hemocyte types in Percoll gradient fractions. Results are from six independent experiments. Statistical significance is indicated by letters, as different letters indicate a significant difference between enrichments of cell types within the Percoll density gradient fractions (ANOVA, Tukey's test, p-value <0.05). Hyalinocytes (**H**) were significantly enriched in the first fraction compared to the other fractions and compared to unsorted hemocytes. However, they were significantly depleted in fractions 4, 5, 6 and 7 compared to unsorted hemocytes. Macrophage-like cells (**ML**) were significantly enriched in fractions 3 and 4 compared to fractions 1, 6 and 7. They were depleted in fraction 1 compared to unsorted hemocytes. Acidophilic blasts (**ABL**) were significantly depleted in fractions 4, 5, 6, and 7 compared to unsorted hemocytes. Basophilic blasts (**BBL**) were significantly enriched in fractions 2 and 3 compared to fractions 4, 5, 6, and 7 and in fraction 1 compared to fractions 6 and 7. Compared to unsorted hemocytes, basophilic blasts (**BBL**) were significantly enriched in fraction 2 and depleted in fractions 6 to 7. Small granule cells (**SGC**) were significantly depleted in fractions 1, 2, and 3 compared to fractions 4, 5, 6, and 7, and also significantly depleted in fractions 4 and 5 compared to fractions 6 and 7. In addition, small granule cells (**SGC**) were significantly enriched in fractions 5, 6, and 7 compared to unsorted hemocytes. The distribution of the big granule cells (**BGC**) showed a significant enrichment in fraction 4, compared to fractions 1 and 2, but no significant changes were observed with unsorted hemocytes. Vesicular cells (**VC**) were enriched in fraction 5 compared to fractions 1, 2, 3, 6, 7 and unsorted hemocytes in fractions 1, 2, and 3 compared to fraction 5 and enriched in fraction 5 compared to fraction 7 and unsorted hemocytes.

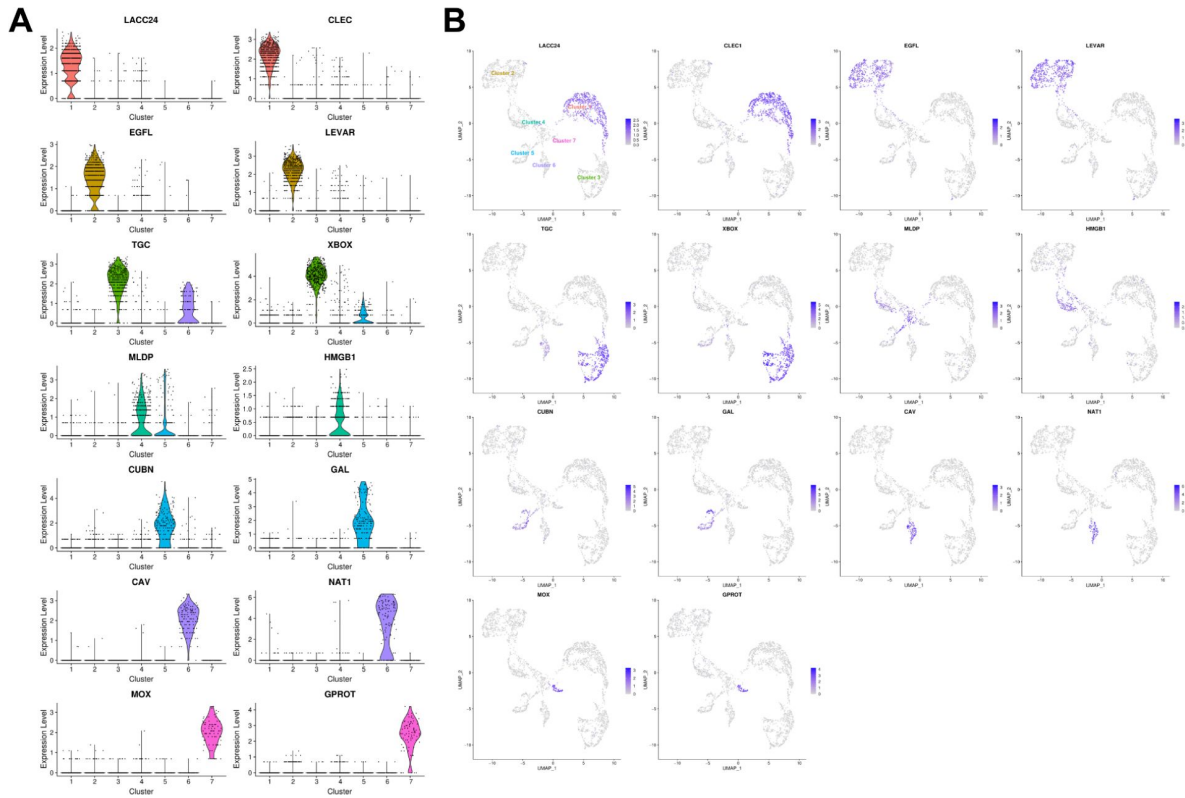


Figure S5.

Cluster specificity and expression level of the 14 selected cluster markers

A) Violin graph showing the average expression level (Log₂FC) of the 14 selected marker transcripts specific to the different scRNA-seq clusters. **B)** Identification of cells expressing the selected markers on the UMAP plot. Positive cells are colored purple according to the Log₂FC value. LACC24 and CLEC are specific of cluster 1, EGFL and LEVAR of cluster 2, TGC and XBOX of cluster 3, MLDP and HMGB1 of cluster 4, CUBN and GAL of cluster 5, CAV and NAT1 of cluster 6 and MOX and GPROT of cluster 7. **LACC24** : Laccase 24, **CLEC** : C-type lectin domain-containing protein, **EGFL** : EGF-like domain-containing protein 8, **LEVAR** : Putative regulator of levamisole receptor-1, **TGC** : Tgc domain-containing protein, **XBOX** : X-box binding protein-like protein, **MLDP** : ML domain-containing protein, **HMGB1** : High mobility group protein B1, **CUBN** : Cubilin, **GAL** : Galectin, **CAV** : Caveolin, **NAT1** : Natterin-1, **MOX** : DBH-like monooxygenase protein 1, **GPROT** : G protein receptor F1-2 domain-containing protein.

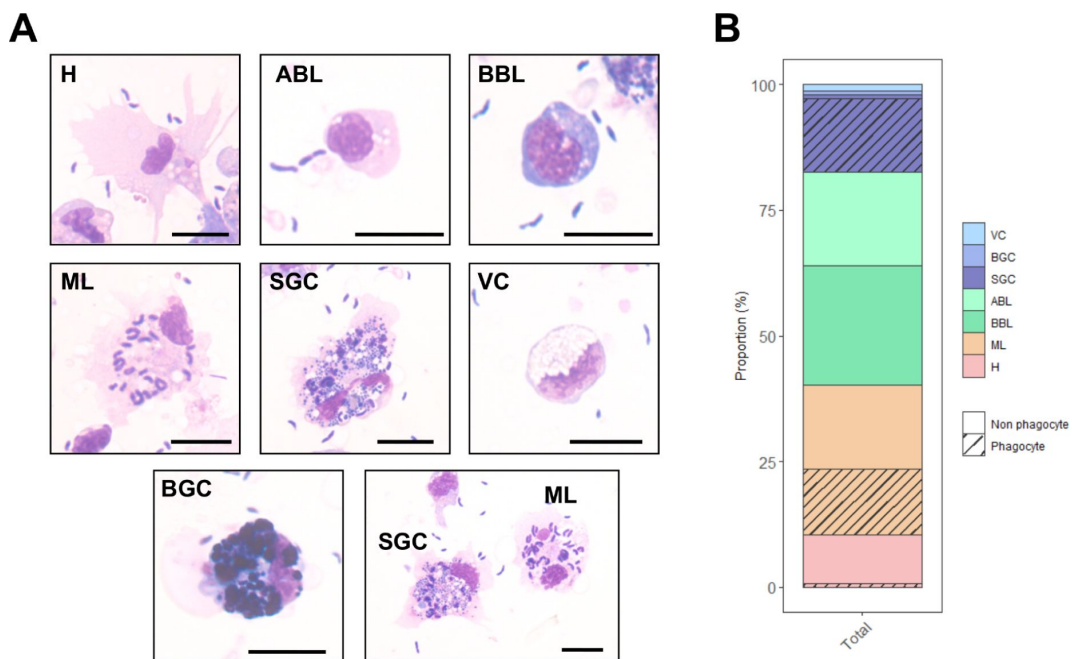


Figure S6.

Measurement of the ability of *C.gigas* hemocytes to phagocytose *Vibrio tasmaniensis* LMG20012^T.

Oyster hemocytes were challenged with non-pathogenic *Vibrio tasmaniensis* LMG20012^T and phagocytosis was measured by observing intracellular bacteria after MCDH staining. **(A)** MCDH staining of hemocytes after phagocytosis assay. Scale bar : 10µm. **(B)** Bar plot showing the proportion of each cell type and the proportion of phagocytic cells. **H** : Hyalinocytes, **ABL** : Acidophilic Blast-Like cells, **BBL** : Basophilic Blast-Like cells, **ML** : Macrophage-Like cells, **SGC** : Small Granule Cells, **VC** : Vesicular Cells and **BGC** : Big Granule Cells.

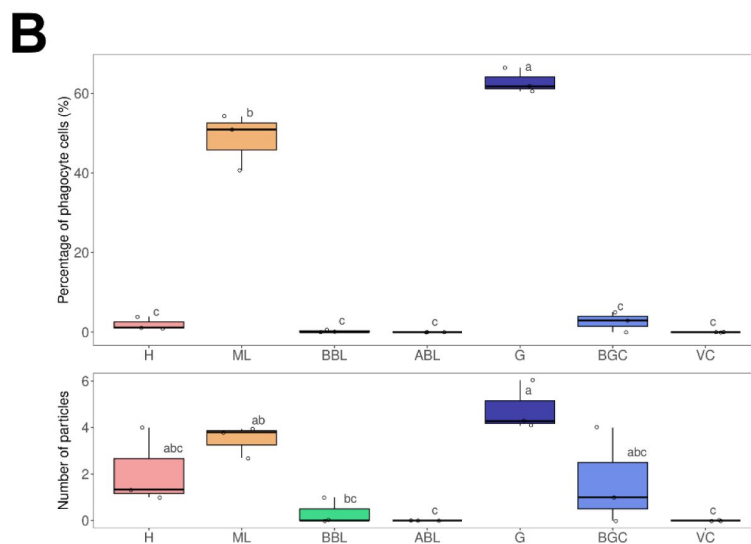
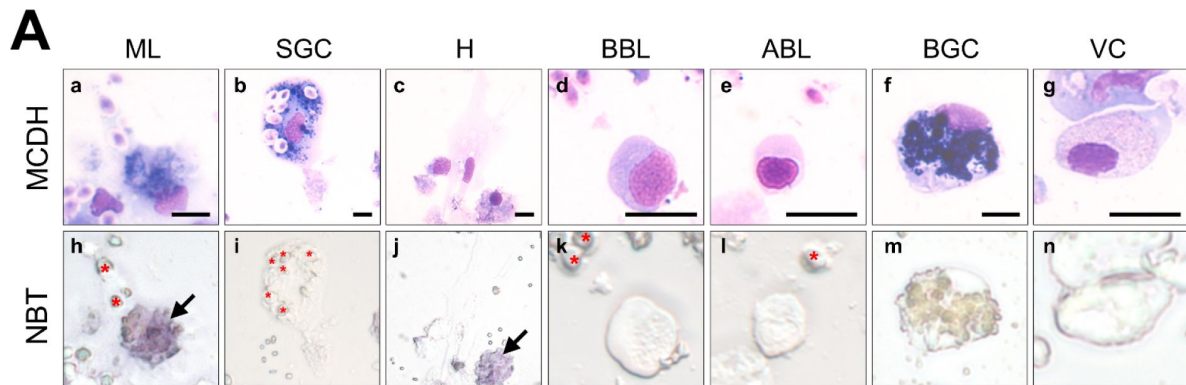


Figure S7.

NBT (NitroBlueTetrazolium) staining of oyster hemolymph exposed to zymosan particles.

(A) Hemocyte morphology after MCDH staining : Macrophage Like (a), Small Granule Cells (b), Hyalinocyte (c), Basophilic (d) and Acidophilic (e) Blast cells, Big Granule Cells (f) and Vesicular Cells (g). NBT staining of the different hemocyte types (h-n). Red stars show zymosan and bacteria particles. Black arrows identify Macrophage-Like cells. Scale bar : 10 μ m. (B) Quantification of the phagocytic activity of each cell type for zymosan particles from 3 independent experiments. Results of quantification of the phagocytic activity of each cell type and number of zymosan particles per cell type. The graph shows the result of 3 independent experiments.

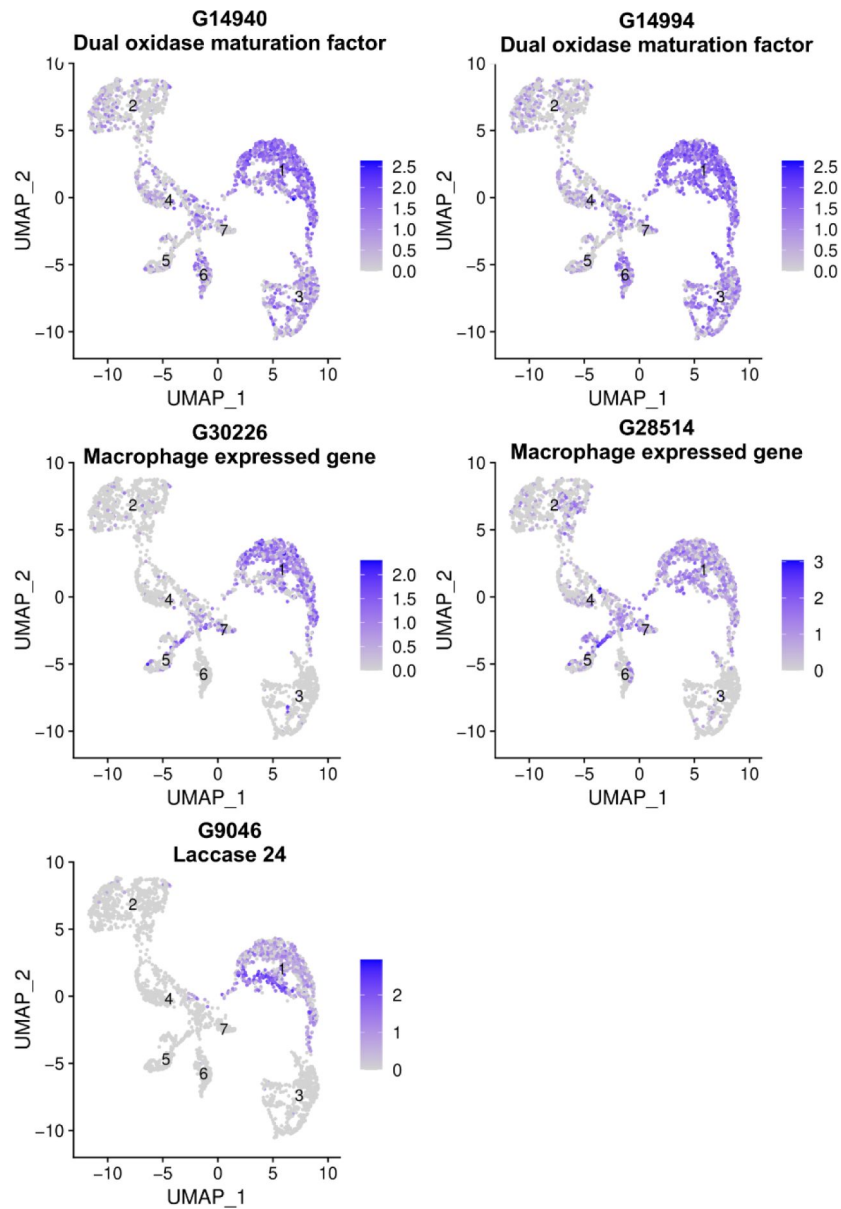


Figure S8.

Uniform Manifold Approximation and Projection (UMAP) plots of cells expressing Macrophage-Like markers.

Cluster numbers are indicated on each cluster. Each point in the UMAP plot represents a single hemocyte, and the clustering of these points reveals the distinct transcriptional profiles of macrophage-like specific markers within the hemocyte population.

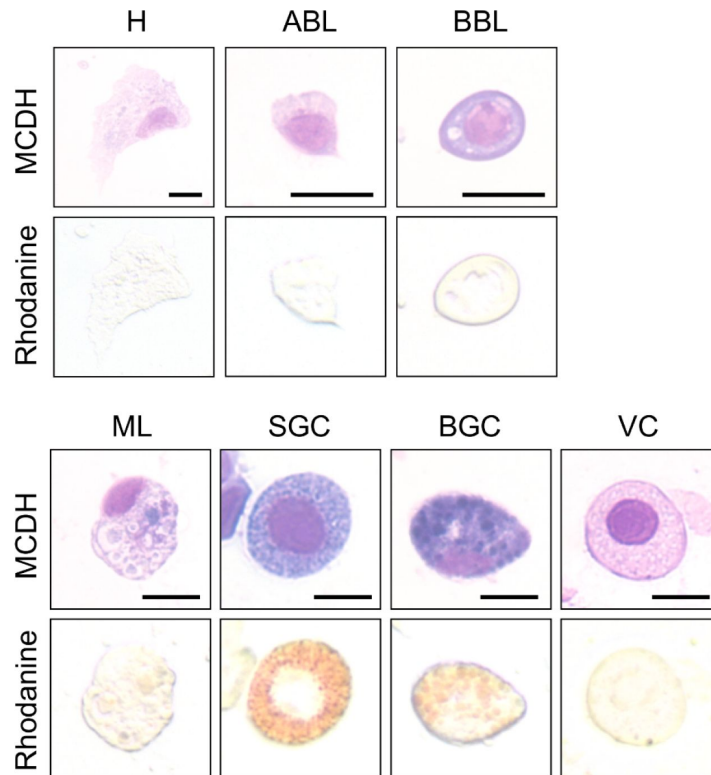


Figure S9.

Labeling of intracellular copper stores in *C.gigas* hemocytes.

MCDH (upper panels) and rhodanine (lower panels) staining of oyster hemocytes to reveal copper accumulation. Cells were first processed for copper staining and then stained according to MCDH protocol. **H** : Hyalinocytes, **ABL** : Acidophilic Blast-Like cells, **BBL** : Basophilic Blast-Like cells, **ML** : Macrophage-Like cells, **SGC** : Small Granule Cells, **VC** : Vesicular Cells and **BGC** : Big Granule Cells. Bar : 10µm.

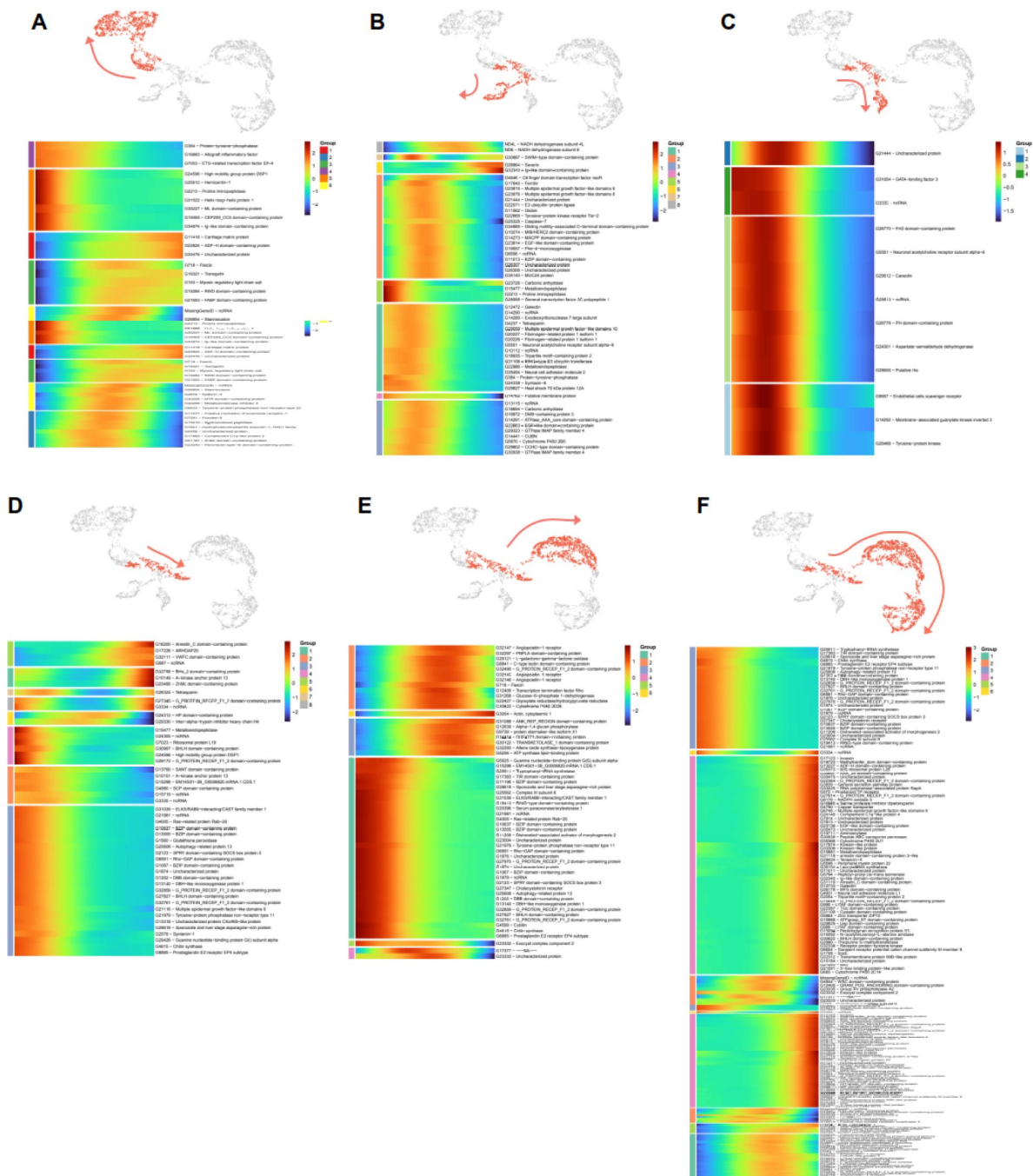


Figure S10.

Results of scRNA-seq trajectory analysis using Monocle3.

Analysis was performed using cluster 4 as the zero pseudotime. **(A)** Lineage from immature cells to hyalinocytes, **(B)** to cluster 5 cells, **(C)** to cluster 6 cells, **(D)** to vesicular cells, **(E)** to macrophage-like cells and **(F)** to small granule cells. For each lineage, cell trajectories are shown in red and heat maps of pseudo time-dependent genes are shown. Blue indicates low expression, and red indicates high expression. Pseudotime flows from right to left.

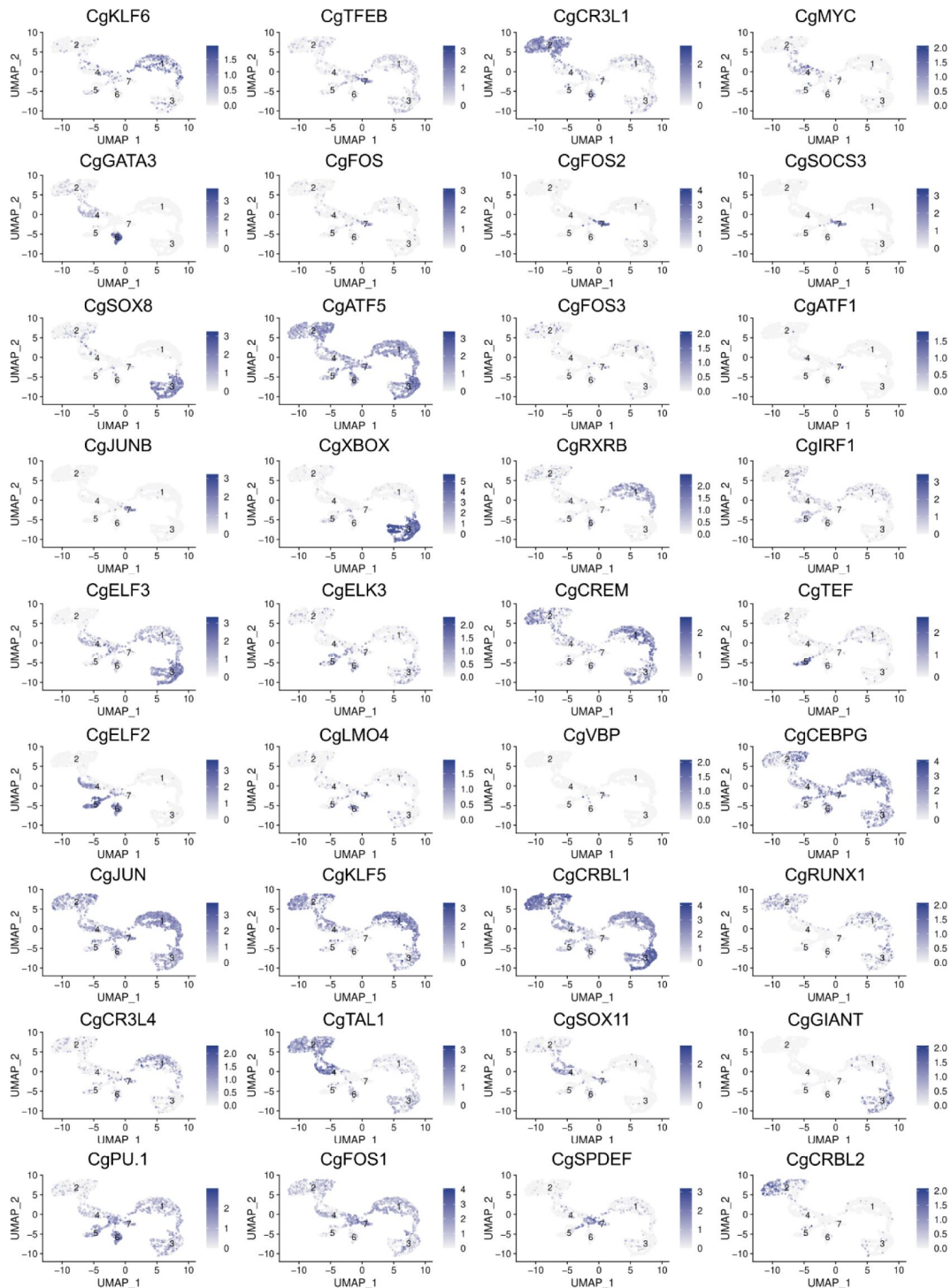


Figure S11.

Uniform Manifold Approximation and Projection (UMAP) plots of cells expressing transcription factors.

28 UMAP representation for the transcription factors identified in the scRNA-seq dataset. Each UMAP plot shows cells expressing the transcription factor in purple. Log2FC expression level is also reported.

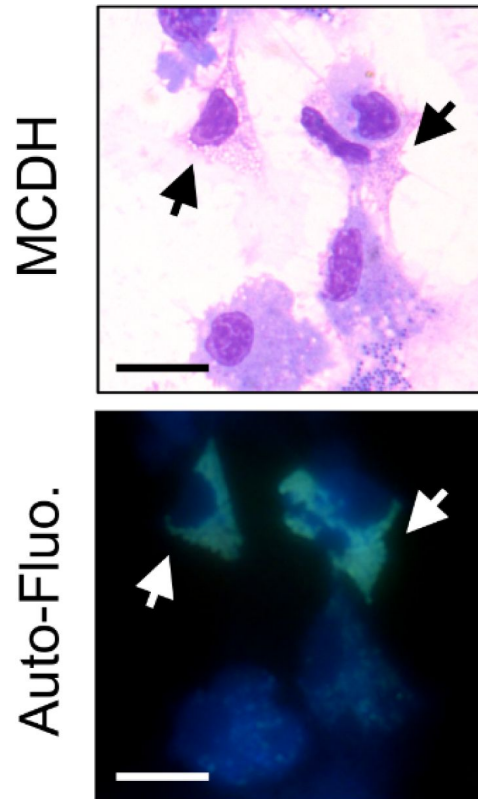


Figure S12.

Observation of autofluorescence of vesicular cells in hemolymph.

Freshly punctured total hemolymph was cytopun, directly observed under a microscope using a DAPI filter set (DAPI Blue ex : 350/50 nm, DC : 400 nm and em : 460/50 nm) and then processed for MCDH staining. Arrows indicate autofluorescent cells.
Scale bar : 10 μ m

References

1. Bar-On Y. M., Phillips R., Milo R. (2018) **The biomass distribution on Earth** *Proceedings of the National Academy of Sciences* **115**:6506–6511
2. Leicester Bayne B. (2017) **Biology of Oysters** Academic Press
3. Zhang G., Li L., Meng J., Qi H., Qu T., Xu F., Zhang L. (2016) **Molecular Basis for Adaptation of Oysters to Stressful Marine Intertidal Environments** *Annu. Rev. Anim. Biosci* **4**:357–381
4. FAO (2023) **Global aquaculture production** *Fisheries and aquaculture division*
5. Pernet F., Lupo C., Bacher C., Whittington R. J. (2016) **Infectious diseases in oyster aquaculture require a new integrated approach** *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**
6. Alfjorden A. *et al.* (2017) **New Trends in Important Diseases Affecting the Culture of Fish and Molluscs in the ICES Area 2002– 2015** *ICES cooperative research report* <https://doi.org/10.17895/ices.pub.2800>
7. de Lorgeril J. *et al.* (2018) **Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters** *Nat Commun* **9**
8. Coyle N. M. *et al.* (2023) **Vibrio aestuarianus clade A and clade B isolates are associated with Pacific oyster (*Magallana gigas*) disease outbreaks across Ireland** *Microbial Genomics* **9**
9. Mesnil A. *et al.* (2023) **Emergence and clonal expansion of *Vibrio aestuarianus* lineages pathogenic for oysters in Europe** *Molecular Ecology* **32**:2869–2883
10. Oyanedel D. *et al.* (2023) **Cooperation and cheating orchestrate *Vibrio* assemblages and polymicrobial synergy in oysters infected with OsHV-1 virus** *Proceedings of the National Academy of Sciences* **120**
11. Rubio T. *et al.* (2019) **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
12. Labreuche Y., Le Roux F., Henry J., Zatylny C., Huvet A., Lambert C., Soudant P., Mazel D., Nicolas J.-L. (2010) ***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
13. Lafont M. *et al.* (2020) **A Sustained Immune Response Supports Long-Term Antiviral Immune Priming in the Pacific Oyster, *Crassostrea gigas*** *mBio* **11** <https://doi.org/10.1128/mbio.02777-19>
14. Fallet M. *et al.* (2022) **Early life microbial exposures shape the *Crassostrea gigas* immune system for lifelong and intergenerational disease protection** *Microbiome* **10**

15. de la Ballina N. R., Maresca F., Cao A., Villalba A. (2022) **Bivalve Haemocyte Subpopulations: A Review** *Front. Immunol* **13**
16. Bachère E., Rosa R. D., Schmitt P., Poirier A. C., Merou N., Charrière G. M., Destoumieux-Garzón D. (2015) **The new insights into the oyster antimicrobial defense: Cellular, molecular and genetic view** *Fish & Shellfish Immunology* **46**:50–64
17. Escoubas J.-M., Gourbal B., Duval D., Green T. J., Charrière G. M., Destoumieux-Garzón D., Montagnani C., Ratcliffe M. J. H. (2016) **Immunity in Molluscs** *Encyclopedia of Immunobiology* Oxford: Academic Press :417–436
18. Mount A. S., Wheeler A. P., Paradkar R. P., Snider D. (2004) **Hemocyte-Mediated Shell Mineralization in the Eastern Oyster** *Science* **304**:297–300
19. Fischer W. S. S. (1988) **Environmental influence on bivalve hemocyte function** *Am Fish Soc Symp* **18**:225–237
20. Nguyen T. V., Alfaro A. C., Merien F. (2019) **Omics approaches to investigate host–pathogen interactions in mass mortality outbreaks of *Crassostrea gigas*** *Reviews in Aquaculture* **11**:1308–1324
21. Meng J., Zhang G., Wang W.-X. (2022) **Functional heterogeneity of immune defenses in molluscan oysters *Crassostrea hongkongensis* revealed by high-throughput single-cell transcriptome** *Fish & Shellfish Immunology* **120**:202–213
22. Rosa R. D., Alonso P., Santini A., Vergnes A., Bachère E. (2015) **High polymorphism in big defensin gene expression reveals presence–absence gene variability (PAV) in the oyster *Crassostrea gigas*** *Developmental & Comparative Immunology* **49**:231–238
23. Kaminow B., Yunusov D., Dobin A. (2021) **STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data** *bioRxiv* <https://doi.org/10.1101/2021.05.05.442755>
24. Peñaloza C., Gutierrez A. P., Eöry L., Wang S., Guo X., Archibald A. L., Bean T. P., Houston R. D. (2021) **A chromosome-level genome assembly for the Pacific oyster *Crassostrea gigas*** *GigaScience* **10**
25. Satija R., Farrell J. A., Gennert D., Schier A. F., Regev A. (2015) **Spatial reconstruction of single-cell gene expression data** *Nat Biotechnol* **33**:495–502
26. Sherman B. T., Hao M., Qiu J., Jiao X., Baseler M. W., Lane H. C., Imamichi T., Chang W. (2022) **DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update)** *Nucleic Acids Research* **50**:W216–W221
27. Wright R. M., Aglyamova G. V., Meyer E., Matz M. V. (2015) **Gene expression associated with white syndromes in a reef building coral, *Acropora hyacinthus*** *BMC Genomics* **16**
28. Bachère E., Chagot D., Grizel H. (1988) **Separation of *Crassostrea gigas* hemocytes by density gradient centrifugation and counterflow centrifugal elutriation** *Developmental & Comparative Immunology* **12**:549–559
29. Wang W., Li M., Wang L., Chen H., Liu Z., Jia Z., Qiu L., Song L. (2017) **The granulocytes are the main immunocompetent hemocytes in *Crassostrea gigas*** *Developmental & Comparative Immunology* **67**:221–228

30. Lambert C., Soudant P., Choquet G., Paillard C. (2003) **Measurement of *Crassostrea gigas* hemocyte oxidative metabolism by flow cytometry and the inhibiting capacity of pathogenic vibrios** *Fish & Shellfish Immunology* **15**:225–240
31. Schmitt P., Rosa R., Duperthuy M., de Lorgeril J., Bachère E., Destoumieux-Garzon D. (2012) **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**
32. Vogt G. (2012) **Hidden Treasures in Stem Cells of Indeterminately Growing Bilaterian Invertebrates** *Stem Cell Rev and Rep* **8**:305–317
33. Jemaà M., Morin N., Cavelier P., Cau J., Strub J.-M., Delsert C. (2014) **Adult somatic progenitor cells and hematopoiesis in oyster** *Journal of Experimental Biology*, jeb **106575**
34. Fahl S. P., Wang M., Zhang Y., Duc A.-C. E., Wiest D. L. (2015) **Regulatory Roles of Rpl22 in Hematopoiesis: An Old Dog with New Tricks** *Crit Rev Immunol* **35**:379–400
35. Rehn M. *et al.* (2022) **PTBP1 promotes hematopoietic stem cell maintenance and red blood cell development by ensuring sufficient availability of ribosomal constituents** *Cell Reports* **39**
36. Athanasiadis E. I., Botthof J. G., Andres H., Ferreira L., Lio P., Cvejic A. (2017) **Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis** *Nat Commun* **8**
37. Cao J. *et al.* (2019) **The single-cell transcriptional landscape of mammalian organogenesis** *Nature* **566**:496–502
38. Hultmark D., Andó I. (2022) **Hematopoietic plasticity mapped in *Drosophila* and other insects** *eLife* **11**
39. Banerjee U., Girard J. R., Goins L. M., Spratford C. M. (2019) ***Drosophila* as a Genetic Model for Hematopoiesis** *Genetics* **211**:367–417
40. Boada-Romero E., Martinez J., Heckmann B. L., Green D. R. (2020) **The clearance of dead cells by efferocytosis** *Nat Rev Mol Cell Biol* **21**:398–414
41. Dale D. C., Boxer L., Liles W. C. (2008) **The phagocytes: neutrophils and monocytes** *Blood* **112**:935–945
42. Vanhove A. S. *et al.* (2016) **Copper homeostasis at the host vibrio interface: lessons from intracellular vibrio transcriptomics** *Environmental Microbiology* **18**:875–888
43. Poirier A. C., Schmitt P., Rosa R. D., Vanhove A. S., Kieffer-Jaquinod S., Rubio T. P., Charrière G. M., Destoumieux-Garzon D. (2014) **Antimicrobial Histones and DNA Traps in Invertebrate Immunity** *Journal of Biological Chemistry* **289**:24821–24831
44. Zaroogian G., Yevich P. (1993) **Cytology and biochemistry of brown cells in *Crassostrea virginica* collected at clean and contaminated stations** *Environmental Pollution* **79**:191–197
45. Allam B., Raftos D. (2015) **Immune responses to infectious diseases in bivalves** *Journal of Invertebrate Pathology* **131**:121–136

46. Nakhleh J., El Moussawi L., Osta M. A., Ligoxygakis P. (2017) **Chapter Three - The Melanization Response in Insect Immunity** *Advances in Insect Physiology* Academic Press :83–109
47. Amparyup P., Charoensapsri W., Tassanakajon A. (2013) **Prophenoloxidase system and its role in shrimp immune responses against major pathogens** *Fish & Shellfish Immunology* **34**:990–1001
48. Picot S., Faury N., Pelletier C., Arzul I., Chollet B., Dégremont L., Renault T., Morga B. (2022) **Monitoring Autophagy at Cellular and Molecular Level in Crassostrea gigas During an Experimental Ostreid Herpesvirus 1 (OsHV-1) Infection** *Front. Cell. Infect. Microbiol* **12**
49. de la Ballina N. R., Villalba A., Cao A. (2020) **Differences in proteomic profile between two haemocyte types, granulocytes and hyalinocytes, of the flat oyster Ostrea edulis** *Fish & Shellfish Immunology* **100**:456–466
50. Aladaileh S., Nair S. V., Birch D., Raftos D. A. (2007) **Sydney rock oyster (Saccostrea glomerata) hemocytes: Morphology and function** *Journal of Invertebrate Pathology* **96**:48–63
51. Rifkin E., Cheng T. C., Hohl H. R. (1969) **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
52. De San Nicolas N. *et al.* (2022) **Functional Diversification of Oyster Big Defensins Generates Antimicrobial Specificity and Synergy against Members of the Microbiota** *Marine Drugs* **20**
53. de la Ballina N. R., Villalba A., Cao A. (2021) **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
54. Pila E. A., Sullivan J. T., Wu X. Z., Fang J., Rudko S. P., Gordy M. A., Hanington P. C. (2016) **Haematopoiesis in molluscs: A review of haemocyte development and function in gastropods, cephalopods and bivalves** *Developmental & Comparative Immunology* **58**:119–128
55. de Bruijn M., Dzierzak E. (2017) **Runx transcription factors in the development and function of the definitive hematopoietic system** *Blood* **129**:2061–2069
56. Scott E. W., Simon M. C., Anastasi J., Singh H. (1994) **Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages** *Science* **265**:1573–1577
57. Bergiers I. *et al.* (2018) **Single-cell transcriptomics reveals a new dynamical function of transcription factors during embryonic hematopoiesis** *eLife* **7**
58. Altschul S. F., Gish W., Miller W., Myers E. W., Lipman D. J. (1990) **Basic local alignment search tool** *Journal of Molecular Biology* **215**:403–410
59. Jones P. *et al.* (2014) **InterProScan 5: genome-scale protein function classification** *Bioinformatics* **30**:1236–1240
60. Cantalapiedra C. P., Hernández-Plaza A., Letunic I., Bork P., Huerta-Cepas J. (2021) **eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale** *Molecular Biology and Evolution* **38**:5825–5829

61. Götz S., García-Gómez J. M., Terol J., Williams T. D., Nagaraj S. H., Nueda M. J., Robles M., Talón M., Dopazo J., Conesa A. (2008) **High-throughput functional annotation and data mining with the Blast2GO suite** *Nucleic Acids Research* **36**:3420–3435
62. Hao Y. *et al.* (2021) **Integrated analysis of multimodal single-cell data** *Cell* **184**:3573–3587

Author information

Sébastien de La Forest Divonne

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

Juliette Pouzadoux

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

Océane Romatif

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

Caroline Montagnani

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

Guillaume Mitta

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

Delphine Destoumieux-Garzon

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

Benjamin Gourbal

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Perpignan, France
ORCID iD: [0000-0003-2097-2563](https://orcid.org/0000-0003-2097-2563)

Guillaume M Charrière[#]

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

[#]EV and GC co-supervised the work.

Emmanuel Vignal[#]

IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France
ORCID iD: [0000-0002-2585-119X](https://orcid.org/0000-0002-2585-119X)

For correspondence: emmanuel.vignal@umontpellier.fr

[#]EV and GC co-supervised the work.

Editors

Reviewing Editor

Bruno Lemaitre

École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Senior Editor

Satyajit Rath

Indian Institute of Science Education and Research (IISER), Pune, India

Reviewer #1 (Public review):

Summary:

In this manuscript, De La Forest Divonne et al. build a repertory of hemocytes from adult Pacific oysters combining scRNAseq data with cytologic and biochemical analyses. Three categories of hemocytes were described previously in this species (i.e. blast, hyalinocyte, and granulocytes). Based on scRNAseq data, the authors identified 7 hemocyte clusters presenting distinct transcriptional signatures. Using Kegg pathway enrichment and RBGOA, the authors determined the main molecular features of the clusters. In parallel, using cytologic markers, the authors classified 7 populations of hemocytes (i.e. ML, H, BBL, ABL, SGC, BGC, and VC) presenting distinct sizes, nucleus sizes, acidophilic/basophilic, presence of pseudopods, cytoplasm/nucleus ratio and presence of granules. Then, the authors compared the phenotypic features with potential transcriptional signatures seen in the scRNAseq. The hemocytes were separated in a density gradient to enrich for specific subpopulations. The cell composition of each cell fraction was determined using cytologic markers and the cell fractions were analysed by quantitative PCR targeting major cluster markers (two per cluster). With this approach, the authors could assign cluster 7 to VC, cluster 2 to H, and cluster 3 to SGC. The other clusters did not show a clear association with this experimental approach. Using phagocytic assays, ROS, and copper monitoring, the authors showed that ML and SGC are phagocytic, ML produces ROS, and SGC and BGC accumulate copper. Then with the density gradient/qPCR approach, the authors identified the populations expressing anti-microbial peptides (ABL, BBL, and H). At last, the authors used Monocle to predict differentiation trajectories for each subgroup of hemocytes using cluster 4 as the progenitor subpopulation.

The manuscript provides a comprehensive characterisation of the diversity of circulating immune cells found in Pacific oysters.

Strengths:

The combination of the two approaches offers a more integrative view.

Hemocytes represent a very plastic cell population that has key roles in homeostatic and challenged conditions. Grasping the molecular features of these cells at the single-cell level will help understand their biology.

This type of study may help elucidate the diversification of immune cells in comparative studies and evolutionary immunology.

Weaknesses:

The study should be more cautious about the conclusions, include further analyses, and inscribe the work in a more general framework.

<https://doi.org/10.7554/eLife.102622.1.sa2>

Reviewer #2 (Public review):

Summary:

This work provides a comprehensive understanding of cellular immunity in bivalves. To precisely describe the hemocytes of the oyster *C. gigas*, the authors morphologically characterized seven distinct cell groups, which they then correlated with single-cell RNA sequencing analysis, also resulting in seven transcriptional profiles. They employed multiple strategies to establish relationships between each morphotype and the scRNAseq profile. The authors correlated the presence of marker genes from each cluster identified in scRNAseq with hemolymph fractions enriched for different hemocyte morphotypes. This approach allowed them to correlate three of the seven cell types, namely hyalinocytes (H), small granule cells (SGC), and vesicular cells (VC). A macrophage-like (ML) cell type was correlated through the expression of macrophage-specific genes and its capacity to produce reactive oxygen species. Three other cell types correspond to blast-like cells, including an immature blast cell type from which distinct hematopoietic lineages originate to give rise to H, SGC, VC, and ML cells. Additionally, ML cells and SGCs demonstrated phagocytic properties, with SGCs also involved in metal homeostasis. On the other hand, H cells, non-granular cells, and blast cells expressed antimicrobial peptides. This study thus provides a complete landscape of oyster hemocytes with functional validation linked to immune activities. This resource will be valuable for studying the impact of bacterial or viral infections in oysters.

Strengths:

The main strength of this study lies in its comprehensive and integrative approach, combining single-cell RNA sequencing, cytological analysis, cell fractionation, and functional assays to provide a robust characterization of hemocyte populations in *Crassostrea gigas*.

- (1) The innovative use of marker genes, quantifying their expression within specific cell fractions, allows for precise annotation of different cellular clusters, bridging the gap between morphological observations and transcriptional profiles.
- (2) The study provides detailed insights into the immune functions of different hemocyte types, including the identification of professional phagocytes, ROS-producing cells, and cells expressing antimicrobial peptides.
- (3) The identification and analysis of transcription factors specific to different hemocyte types and lineages offer crucial insights into cell fate determination and differentiation processes in oyster immune cells.
- (4) The authors significantly advance the understanding of oyster immune cell diversity by identifying and characterizing seven distinct hemocyte transcriptomic clusters and morphotypes.

These strengths collectively make this study a significant contribution to the field of invertebrate immunology, providing a comprehensive framework for understanding oyster hemocyte diversity and function.

Weaknesses:

- (1) The authors performed scRNAseq/lineage analysis and cytological analysis on oysters from two different sources. The methodology of the study raises concerns about the consistency of the sample and the variability of the results. The specific post-processing of hemocytes for scRNAseq, such as cell filtering, might also affect cell populations or gene expression profiles. It's unclear if the seven hemocyte types and their proportions were consistent across both samples. This inconsistency may affect the correlation between morphological and transcriptomic data.

(2) The authors claim to use pathogen-free adult oysters (lines 95 and 119), but no supporting data is provided. It's unclear if the oysters were tested for bacterial and viral contaminations, particularly *Vibrio* and OsHV-1 μ Var herpesvirus.

(3) The KEGG and Gene Ontology analyses, while informative, are very descriptive and lack interpretation. The use of heatmaps with dendrograms for grouping cell clusters and GO terms is not discussed in the results, missing an opportunity to explore cell-type relationships. The changing order of cell clusters across panels B, C, and D in Figure 2 makes it challenging to correlate with panel A and to compare across different GO term categories. The dendrograms suggest proximity between certain clusters (e.g., 4 and 1) across different GO term types, implying similarity in cell processes, but this is not discussed. Grouping GO terms as in Figure 2A, rather than by dendrogram, might provide a clearer visualization of main pathways. Lastly, a more integrated discussion linking GO term and KEGG pathway analyses could offer a more comprehensive view of cell type characteristics. The presentation of scRNAseq results lacks depth in interpretation, particularly regarding the potential roles of different cell types based on their transcriptional profiles and marker genes. Additionally, some figures (2B, C, D, and 7C to H) suffer from information overload and small size, further hampering readability and interpretation.

(4) The pseudotime analysis presented in the study provides modest additional information to what is already manifest from the clustering and UMAP visualization. The central and intermediate transcriptomic profile of cluster 4 relative to other clusters is apparent from the UMAP and the expression of shared marker genes across clusters (as shown in Figure 1D). The statement by the authors that 'the two types of professional phagocytes belong to the same granular cell lineage' (lines 594-596) should be formulated with more caution. While the pseudotime trajectory links macrophage-like (ML) and small granule-like (SGC) cells, this doesn't definitively establish a direct lineage relationship. Such trajectories can result from similarities in gene expression induced by factors other than lineage relationships, such as responses to environmental stimuli or cell cycle states. To conclusively establish this lineage relationship, additional experiments like cell lineage tracing would be necessary, if such tools are available for *C. gigas*.

(6) Given the mention of herpesvirus as a major oyster pathogen, the lack of discussion on genes associated with antiviral immunity is a notable omission. While KEGG pathway analysis associated herpesvirus with cluster 1, the specific genes involved are not elaborated upon.

(7) The discussion misses an opportunity for comparative analysis with related species. Specifically, a comparison of gene markers and cell populations with *Crassostrea hongkongensis*, could highlight similarities and differences across systems.

Conclusion:

The authors largely achieved their primary objective of providing a comprehensive characterization of oyster immune cells. They successfully integrated multiple approaches to identify and describe distinct hemocyte types. The correlation of these cell types with specific immune functions represents a significant advancement in understanding oyster immunity. However, certain aspects of their objectives have not been fully achieved. The lineage relationships proposed on the basis of pseudotime analysis, while interesting, require further experimental validation. The potential of antiviral defense mechanisms, an important aspect of oyster immunity, has not been discussed in depth.

This study is likely to have a significant impact on the field of invertebrate immunology, particularly in bivalve research. It provides a new standard for comprehensive immune cell characterization in invertebrates. The identification of specific markers for different hemocyte types will facilitate future research on oyster immunity. The proposed model of

hemocyte lineages, while requiring further validation, offers a framework for studying hematopoiesis in bivalves.

<https://doi.org/10.7554/eLife.102622.1.sa1>

Reviewer #3 (Public review):

The paper addresses pivotal questions concerning the multifaceted functions of oyster hemocytes by integrating single-cell RNA sequencing (scRNA-seq) data with analyses of cell morphology, transcriptional profiles, and immune functions. In addition to investigating granulocyte cells, the study delves into the potential roles of blast and hyalinocyte cells. A key discovery highlighted in this research is the identification of cell types engaged in antimicrobial activities, encompassing processes such as phagocytosis, intracellular copper accumulation, oxidative bursts, and antimicrobial peptide synthesis.

A particularly intriguing aspect of the study lies in the exploration of hemocyte lineages, warranting further investigation, such as employing scRNA-seq on embryos at various developmental stages.

In the opinion of this reviewer, the discussion should compare and contrast the transcriptome characteristics of hemocytes, particularly granule cells, across the three species of bivalves, aligning with the published scRNA-seq studies in this field to elucidate the uniformities and variances in bivalve hemocytes.

<https://doi.org/10.7554/eLife.102622.1.sa0>