

## OPEN ACCESS

EDITED BY  
Linsheng Song,  
Dalian Ocean University, China

REVIEWED BY  
Yueling Zhang,  
Shantou University, China  
Fan Wang,  
Shantou University, China

\*CORRESPONDENCE  
Xiaoqian Tang  
tangxq@ouc.edu.cn

SPECIALTY SECTION  
This article was submitted to  
Comparative Immunology,  
a section of the journal  
Frontiers in Immunology

RECEIVED 28 June 2022  
ACCEPTED 22 August 2022  
PUBLISHED 13 September 2022

CITATION  
Cui C, Tang X, Xing J, Sheng X, Chi H  
and Zhan W (2022) Single-cell RNA-  
seq uncovered hemocyte functional  
subtypes and their differentional  
characteristics and connectivity with  
morphological subpopulations in  
*Litopenaeus vannamei*.  
*Front. Immunol.* 13:980021.  
doi: 10.3389/fimmu.2022.980021

COPYRIGHT  
© 2022 Cui, Tang, Xing, Sheng, Chi and  
Zhan. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](#). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Single-cell RNA-seq uncovered hemocyte functional subtypes and their differentional characteristics and connectivity with morphological subpopulations in *Litopenaeus vannamei*

Chuang Cui<sup>1</sup>, Xiaoqian Tang<sup>1,2\*</sup>, Jing Xing<sup>1,2</sup>, Xiuzhen Sheng<sup>1</sup>, Heng Chi<sup>1,2</sup> and Wenbin Zhan<sup>1,2</sup>

<sup>1</sup>Laboratory of Pathology and Immunology of Aquatic Animals, The Key Laboratory of Mariculture, Ministry of Education (KLMME), Ocean University of China, Qingdao, China, <sup>2</sup>Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

Hemocytes play central roles in shrimp immune system, whereas whose subclasses have not yet been completely defined. At present, the morphological classification of hemocytes is inadequate to classify the complete hemocyte repertoire and elucidate the functions and differentiation and maturation processes. Based on single-cell RNA sequencing (scRNA-seq) of hemocytes in healthy *Litopenaeus vannamei*, combined with RNA-FISH and flow cytometric sorting, we identified three hemocyte clusters including TGase<sup>+</sup> cells, CTL<sup>+</sup> cells and Crustin<sup>+</sup> cells, and further determined their functional properties, potential differentiation trajectory and correspondence with morphological subpopulations. The TGase<sup>+</sup> cells were mainly responsible for the coagulation, exhibiting distinguishable characteristics of hyalinocyte, and appeared to be developmentally arrested at an early stage of hemocyte differentiation. The CTL<sup>+</sup> cells and Crustin<sup>+</sup> cells arrested at terminal stages of differentiation mainly participated in recognizing foreign pathogens and initiating immune defense responses, owning distinctive features of granule-containing hemocytes. Furthermore, we have revealed the functional sub-clusters of three hemocyte clusters and their potential differentiation pathways according to the expression of genes involved in cell cycle, cell differentiation and immune response, and the successive differentiation and maturation of hyalinocytes to granule-containing hemocytes have also mapped. The results revealed the diversity of shrimp hemocytes and provide new theoretical rationale for hemocyte classification, which also facilitate systematic research on crustacean immunity.

## KEYWORDS

single-cell RNA-seq, hemocytes, *Litopenaeus vannamei*, functional cluster, differentiation trajectory

## Introduction

Shrimp have been considered to the most heavily traded fish products and shrimp culture is one of the fastest growing industries for producing animal proteins in the world (1, 2). The *Litopenaeus vannamei* is the most important cultured shrimp worldwide, accounting for approximately 80% of total cultured shrimp production (3). Over the past decades, disease still represent a major impediment to the development of shrimp culture, and disease outbreaks result in large economic losses to the industry (4). As a typical aquatic invertebrate, shrimps rely on innate immune system for defense against pathogenic infection (5, 6), in which hemocytes perform central roles *via* multiple mechanisms. The hemocytes participate in humoral immune responses by synthesis and release of antimicrobial peptides, lectins, reactive oxygen species and inflammatory cytokines, and also are involved in cellular immunity *via* recognition of foreign pathogens, phagocytosis, melanization and apoptosis (7–10). Therefore, it is vital to characterize hemocytes systemically and comprehensively for deepening understanding of their immune roles in shrimp, which would benefit designing novel and efficient strategies for disease prevention and control.

Due to the lack of marker genes available to define cell types, the classification of crustacean hemocytes is still performed on the basis of morphological characteristics (11–13). In shrimp, hemocytes are also classified into granulocytes, semi-granulocytes and hyalinocytes depending on their morphology and intracellular granule staining characteristics (9, 14). However, the functional properties and differentiation relationships of the three hemocyte subpopulations have not been fully revealed. Moreover, there is inconsistent finding regarding hemocytes classification, some researches have classified shrimp hemocytes into four types by ultrastructural and microscopic observation (15), five types by iodixanol density gradient centrifugation (16), two or three types by lectin staining (17, 18), and even much more types based on differential binding affinities of hemocytes towards monoclonal antibodies (19, 20). Nevertheless, none of these cell subpopulations have been characterized at the molecular level and their transcriptional profiles remain unknown, and the correspondence between subpopulations and differentiation stages are still also ambiguous and vague (21, 22). In addition, the classification based on low-resolution of morphological and staining characteristics is difficult to identify rare cell types and might hinder distinguishing the cells in transient state (9, 23, 24). Hence, it is important to thoroughly unravel molecular signatures of hemocytes, which enable elucidation of functional and differentiation characteristics of hemocytes subpopulation and efficient identification of marker genes.

In recent years, the emergence and advancement of single-cell RNA sequencing (scRNA-seq) technologies has made it possible

to comprehensively characterize complex tissues cells at the molecular level (25, 26). The scRNA-seq provides access to the dynamic gene expression patterns of single cells and according to the large amount of transcript information obtained, it is possible to accurately characterize cell type, annotate the unclassified cells, explore marker genes and map the developmental trajectories of cellular lineages, thus providing a thorough analysis of cell heterogeneity and their states (27–29). Single cell transcriptome technology has been widely used in the research on the properties of the organ and tissue cells in plants and animals (30–32). Meanwhile, the successful applications of scRNA-seq on invertebrate hemocytes from *Drosophila*, mosquitoes and oysters have established the basis for scRNA-seq studies on shrimp and further revealed the significance of characterizing the functional properties of hemocytes at the single-cell transcriptional level (33–35).

Here, we performed scRNA-seq of hemocytes from healthy *L. vannamei* to distinguish the hemocytes types and explore their functional properties. As a result, thousands of hemocytes transcripts were obtained, which were utilized to identify the main cell types and their putative marker genes, characterize their function and reveal the differentiation relationships among different cell types. Moreover, we sorted morphological subpopulations of hemocytes using flow cytometry and organically aligned them with the above cell types. In summary, our scRNA-seq analysis revealed the diversity of the shrimp hemocyte populations and provided a deeper understanding of the hemocytes subpopulations.

## Materials and methods

### Shrimp and hemocytes preparation

The healthy male adult shrimp (*Litopenaeus vannamei*) of 30–40 g were obtained from an aquaculture farm in Qingdao, Shandong Province. The shrimp were temporarily maintained at 23°C in a recirculating system containing aerated filtered seawater for one week and then used for experimental studies. The studies were carried out in agreement with the International Guiding Principles for Biomedical Research Involving Animals documented by Guide for the Use of Experimental Animals and the Committee of the Ethics on Animal Care and Experiments at Ocean University of China (permit number: 20180101).

Using sterile modified cold Alsever solution (27 mM Na citrate, 336 mM NaCl, 115 mM glucose, 9 mM EDTA, pH 7.2, AS) as anticoagulant, hemolymph was withdrawn from the pericardial cavity of three randomly selected healthy prawns and pooled in a sterile centrifuge tube (7). The hemocytes suspension was then centrifuged at 400g for 5 min at 4°C and washed three times with sterile prawn homoiosmotic phosphate buffered saline (377 mM NaCl, 2.70 mM KCl, 8.09 mM

Na<sub>2</sub>HPO<sub>4</sub>, 1.47 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4, 780 mOsm/L, PHPBS). After washing, the prepared hemocytes suspension was passed through a 40 µm cell strainer (Corning, USA) to remove cell aggregates and counted on a cell counting plate. Cellular viability was measured using a Trypan Blue Staining kit (Solarbio) according to the operating instructions.

## Single hemocyte encapsulation and sequencing

Single cells encapsulation, library synthesis and RNA sequencing were performed by the Gene Denovo (Guangzhou, China). In brief, the concentration of hemocytes was adjusted to 1500–2000 cells/µl. Each single cell was encapsulated in a droplet to generate a Gel Beads-In-Emulsions (GEMs) using 10X Genomics GemCode Technology according to the manufacturer's instructions. Every cell was barcode-labeled, and every transcript was labeled with a unique molecular identifier (UMI). The cDNA library was then amplified by PCR with the sequencing primer R1 and P5 arm. The concentration of cDNA library was measured by Qubit (Invitrogen) and qualitative analysis of cDNA library was performed by an Agilent 2100 Bioanalyzer. The qualified cDNA libraries were sequenced on the Illumina 10X Genomics Chromium platform.

## Quality control and data processing

The obtained raw single-cell RNA sequencing dataset was submitted for data quality statistics, reads aligning and generating gene-cell matrices by the official 10X Genomics analysis software Cell Ranger. Read1 and Read2 were obtained by Illumina double-end sequencing, with Read1 containing a 16 bp size Gemcode for distinguishing different cells and a 10 bp size UMI for labeling each transcript, while Read2 contained mainly cDNA sequence fragments. The Spliced Transcripts Alignment to a Reference alignment software (STAR) of Cell Ranger was used to compare Read2 with the reference genome of *L. vannamei* (ASM378908v1) (36, 37). The reads located in exons were aligned with transcripts containing annotation information, and if the two were in the same direction, they were considered to be transcriptome and annotated as transcriptome reads. In transcriptome reads, if the reads matched only one gene, it was considered to be un-mapped. UMI counting was only available using uni-mapped reads. Subsequently, quality control analysis was performed to filter and correct for barcodes and UMIs. The QC relies on the following criteria: (1) 500 < gene counts < 4000 per cell. (2) UMI counts < 8000 per cell. (3) the percentage of mitochondrial genes < 10%. UMIs were not allowed to be single oligo-chains and were not able to contain bases with a mass value of less than

10. Meanwhile, UMI counting could only be conducted by valid comments with validate barcode and UMI. Following removal of the UMI repeats, the number of unique UMIs was calculated as the expression level of the cellular gene. The filtered and quantified raw data was generated by Cell Ranger as gene-cell matrices and used for further analysis.

## Data dimensionality reduction and cell clustering

The eligible gene-cell matrices were loaded into the Seurat package of R software for cell clustering and dimensionality reduction analysis. The data were first normalized by 'LogNormalize' function and the variations between cells were regressed out by counting the number of target molecules in each cell (UMI). Then, Principal Component Analysis (PCA) was performed on the scaled data for dimensionality reduction, and the low *P*-values genes were strong enrichment. The distribution of *P*-values was compared by the JackStrawPlot function, and the most significant principal component (PC, *p* < 10<sup>-5</sup>) was selected in the PCA results for downstream clustering and cluster analysis (38). Based on the previously determined PCs, the cells were clustered and grouped using the 'FindClusters' function. The data structures and cell trajectories were separately visualized and explored by *t*-SNE and UMAP (39, 40).

## Marker gene identification

For the cell clusters obtained above, marker genes (cluster-enriched genes) were identified and tested by Seurat function 'FindAllMarkers' and 'roc', respectively. Following this analysis, cluster-specific marker genes were further screened for genes expressed in more than 50% cells of a cluster and an average natural log-fold change greater than 0.5. Multiple replicate analyses were performed using the same or different parameters to verify the accuracy of the marker genes.

## Differential expression genes analysis

To further characterize the transcriptional regulatory patterns of each hemocyte cluster, likelihood ratio-tests were used to find the DEGs of individual cell clusters compared to all other hemocyte clusters (41). Subsequently, differential gene expression analysis was performed on different cell populations using the Seurat 'bimod' function. The up-regulated genes in each hemocyte cluster were screened for functional analysis based on the conditions |logFC| ≥ 0.25, *p*-value ≤ 0.05, and the percentages of cells in specific cluster for which the gene was detected to be > 25%. Based on the differentially upregulated genes expression profiles in each of

the cell clusters, we first used gene ontology (GO) enrichment to analyze the biological function in each cell type and the important differentially expressed transcripts (42). The differentially expressed transcript was mapped to GO terms in Gene Ontology database (<http://www.geneontology.org/>), and a list of transcripts and transcriptions with a certain GO functions were obtained by calculating the number of genes per term. Then, hypergeometric tests were applied to define GO terms that were significantly enriched in peak-associated genes relative to the genomic background and to identify significant biological functions of differentially expressed transcripts. Meanwhile, KEGG pathway analysis was used to unravel the coordinated interactions between different transcripts and obtain the enriched pathways in different cell clusters (43). Using the KEGG pathway as a unit, the pathways significantly enriched in differentially expressed transcripts were extracted by hypergeometric testing to further analyze the biological functions of differentially transcribed genes in each cluster.

### Pseudotemporal ordering of cells using Monocle 3

The Monocle (v.3.0) package was utilized for in-depth analysis of the cell differentiation process and cell fate. Based on the previously identified cell clusters and differentially expressed genes, the above data were transferred to Monocle package, then the differentiation and cell fate related genes were selected to define the cell differentiation process. Monocle compressed the data dimensionality to one with two dimensions and arranges all cells in an ordered manner. Genes with similar expression trends and might share common biological functions were grouped together to calculate cell trajectories using the 'learning-graph' function.

### Sub-clustering analysis

We applied the 'FindClusters' function to each of the obtained cell clusters to perform sub-cluster analysis with a view to clarifying the subpopulation composition and functional relationships of the cell clusters. The transcript expression patterns of each secondary subpopulation were clarified using gene differential analysis, and all sub-clusters obtained were reconstructed by Monocle3 package for the Pseudotime analysis.

### Fluorescent *in situ* hybridization

The FISH procedure was carried out strictly according to the manufacturer's protocol of FISH probes kit (GenePharam). Three healthy shrimp hemocytes were drawn, washed and fixed on slides with 4% paraformaldehyde. The samples were

sequentially dehydrated with 100%, 95%, 85% and 75% ethanol for 2 min. After pepsin digestion for 5 min, the samples were washed with hybridization solution. The nucleic acid probes with fluorescent labels were then incubated overnight at 37°C. The sensing probe was used as a negative control. Following washing, the samples were observed by fluorescence microscopy. The nucleic acid probes used for the experiments were synthesized by GenePharam and Tsingke (seen Table 1).

### Cell sorting of hemocytes

To investigate the expression characteristics of critical genes in hemocyte morphological subpopulations, we first analyzed and sorted total hemocytes using FACSARIA Fusion flow cytometry (BD, USA). The hemocytes of 10 shrimp were drawn and analyzed by flow cytometry. The side-scatter parameters and forward-scatter parameters (SSC and FSC) were used to distinguish different morphological subpopulations of hemocytes based on their size and granularity. In the FSC/SSC two-dimensional space, two main populations were detected, the Region1 (R1)-small/granular simple population defined as agranulocytes and the Region2 (R2)-large/granular complex population defined as granulocytes. After sorting the total hemocytes according to R1 and R2 regions, the different subpopulations were stained with May-Grunwald and Giemsa staining solutions (Solarbio) in order to observe the composition of the cells.

### QRT-PCR and FISH of selected genes from sorted hemocyte subpopulations

The total RNA of sorted hemocyte subpopulations were separately extracted using TRIzol Reagent<sup>®</sup> reagent (Invitrogen, USA), and the quality and quantity were detected by a Nanodrop

TABLE 1 Probes used in FISH of this study.

Genes Names (Gene ID)	Sequences
TGase1 (ncbi_113823934)	AGGGAGATATG+TGCCGTCATGT+TCATT
TGase (ncbi_113823945)	ACGGCGT+TGTAAGGCTGGCTGGAAGT
CTL-2 (ncbi_113804553)	GCTCG+TAGTGG+TAGCAGGCCGAAGT
SVC5 (ncbi_113809352)	GGCAGCAACCAGGG+TAAACCAAGGAC
CrustinL (ncbi_113816267)	GCACCTACCGGCT+TGCAGCAG
Crustin (ncbi_113801825)	GTAGTCGT+TGGAACAGGT+TGTGGGG
Notch1 (ncbi_113814237)	GGGCAACACCT+TCCCGTCAACCGA
Ast2 (ncbi_113822052)	ATGTA+TCCGTAAGGACAA +TCTGAAGGCC
CHF (ncbi_113802754)	CAATA+TCCGCAGTAATCAACA +TCAACCC
Negative control	TGCTTTGCACGGTAACGCCTGTTTT

8000 spectrophotometer (Thermo Scientific, MA, USA). The single-strand cDNA was synthesized by M-MLV reverse transcriptase reagent Kit (Promega). Then the expression profiles of related genes in hemocyte subpopulations were detected by qRT-PCR with specific primers (shown in Table 2). The 18S rRNA was used as the reference gene. The qRT-PCR was performed in triplicate using ChamQ Universal SYBR qPCR (Vazyme) in a LightCycler<sup>®</sup> 480 II Real Time PCR System (Roche, Basel, Switzerland), following the manufacturer's protocol. The relative gene expression levels were calculated by the  $\Delta\Delta\text{CT}$  method against 18S rRNA. For FISH, the obtained hemocyte were fixed and settled onto slides according to the previous procedure, and nucleic acid probes of selected genes were incubated. Then the slides were observed by fluorescence microscopy.

## Statistical analysis

Statistical analysis was performed using GraphPad Prism. The data was expressed as the mean  $\pm$  SD,  $n=3$ . Unpaired Student's  $t$ -test was used to determine statistically significant differences, and  $p < 0.05$  was considered significant. Data analysis and graph plotting were conducted with the OmicShare online tool (<https://www.omicshare.com/>).

TABLE 2 Primers used in qPCR of this study, related to Figure 6.

Genes Names (Gene ID)	Primers	Sequences (5'-3')
TGase1 (ncbi_113823934)	F	GTGACCCGAAGGAACCAGCC G
	R	AACTGCAGCGAGAACATACCTCCTC
TGase (ncbi_113823945)	F	CGAGATGTACGAGCTGGTCGAG
	R	GGACGAGGACCGAAGGAGAAG
CTL-2 (ncbi_113804553)	F	CCACTACTTCTACATTAGCGACGAGG
	R	TCGAAAAGCATCAGGTGCTCAC
SVC5 (ncbi_113809352)	F	GCGCCGACAAGAAGGCAG
	R	AGGCCGAACGGAGAACAGG
CrustinL (ncbi_113816267)	F	CGGTGGCGTCTTCCCAG
	R	TGCAGTAATTGCAGTTGAATCCGC
Crustin (ncbi_113801825)	F	TCTTGGAAGTGGCACAAGCGAC
	R	AACGTGGGCATGTGGGAC
Notch1 (ncbi_113814237)	F	GTTGCCAAGATGCTCCTCC
	R	CCTTGCCCTTGTGCGTTG
Ast2 (ncbi_113822052)	F	CCTCCAGTCAGAGGTTTGTATTGCA
	R	AACTTCAGCGAACTTCCTCACAG
CHF (ncbi_113802754)	F	GTTGGGTCTATCTGAGGTGGACAATG
	R	CCTTCCACTTTTCTCAATAGCTGGAG
18S rRNA (AY438005)	F	ACAATGGCTATCAGGGTAACG
	R	CTGCTGCCTTCCTTAGATGTGGTA

F, Forward primer; R, Reverse primer.

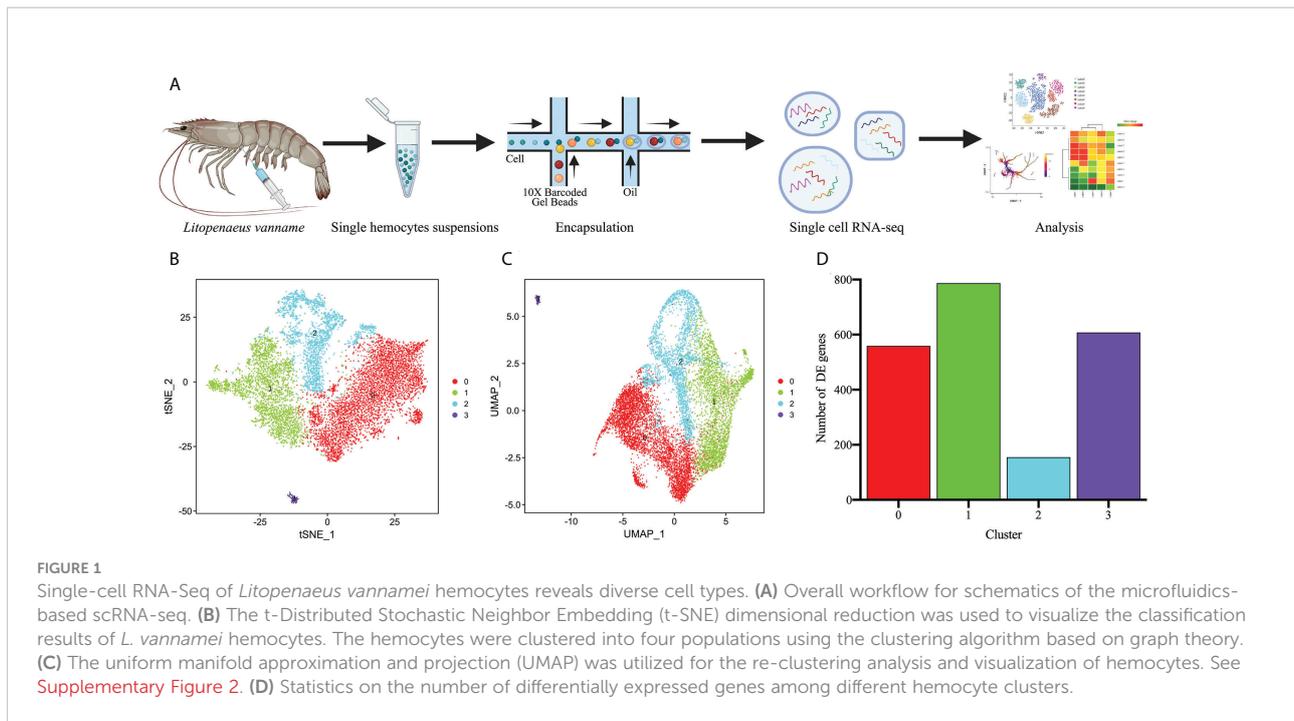
## Results

### scRNA-seq of hemocytes from healthy *L. vannamei*

The shrimps used for hemocytes preparation were free of specific pathogenic infections by microscopic observation and bacteriological examination. To perform scRNA-seq of *L. vannamei* hemocytes, the hemocytes of three shrimp were harvested and mixed as one sample, and then the library construction and sequencing analysis were conducted using 10x Genomics platform and Illumina NovaSeq sequencer (Figure 1A). By filtering out ineffective cells and low-quality genes, we finally obtained 9896 valid hemocytes for subsequent analysis, each with a median unique molecular identifier (UMI) of 5733.5 and a median gene counts of 1199 (Supplementary Figure 1; Supplementary File 1). The filtered datasets were visualized and analyzed using principal component linear dimensional reduction and the  $t$ -distributed stochastic neighborhood embedding (t-SNE) tool (44, 45). The unsupervised cluster detection algorithm (SEURAT) grouped hemocytes into four cell clusters based on the similar gene expression profiles (46) (Figures 1B, C). Details of the four cell clusters obtained were as follows: Cluster 0, 4706 cells (47.55%), with a median gene value of 952 and a median UMI of 4066.5 per cell; Cluster 1, 2818 cells (28.48%), with a median gene value of 1552.5 and a median UMI of 8675 per cell, Cluster 2, 2253 cells (22.77%) with a median gene value of 1237 per cell and a median UMI of 10988, and Cluster 3, 119 cells with a median gene value of 1107 per cell and a median UMI of 3251 (Figure 1D; Supplementary Figure 2).

### Identification of hemocyte clusters

To identify the four types of hemocyte clusters obtained, we first used the Seurat 'FindMarkers' tool to predict the TOP20 genes for each cluster (Supplementary Figure 3), which were further screened to define highly specific marker genes for each cell cluster, and ultimately a number of potential marker genes were identified for the four hemocyte clusters (Supplementary Figure 4). The specific marker genes for Cluster 0 included hemocyte transglutaminase (TGase1), hemocyte transglutaminase (TGase) and neurogenic locus notch homolog protein 1 (Notch1) (Figures 2A, B). TGases are the key enzymes in organisms that play critical roles in coagulation and other biochemical processes involving post-translational protein remodeling (47, 48). In crustaceans, TGases act as hemocyte coagulants by cross-linking coagulation proteins, and they also could regulate the expression of immune-related genes such as lysozyme (49, 50). Two types of TGase have been discovered in *L. vannamei* and studies have shown that they were both involved in hemocyte coagulation (49, 51). It was also reported that TGase

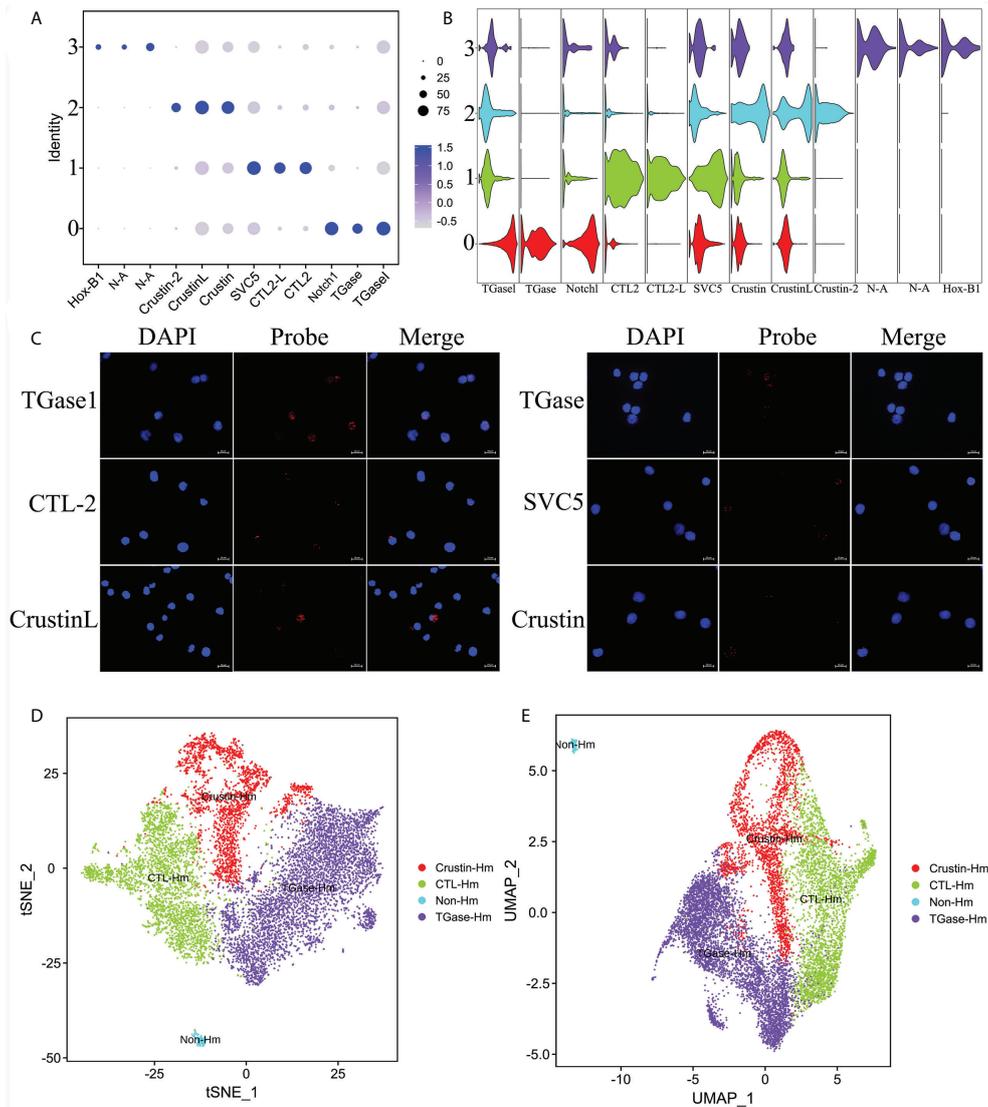


was one of the immature hemocyte markers in crayfish, and once the TGase around immature hemocytes is digested, hemocytes begin to migrate outward and differentiate into mature hemocytes (52, 53). Notch1, one of the major functional proteins in the notch signaling pathway, is widely associated with cell proliferation, differentiation, activation and innate immunity (54–56). Notch1 mediates the differentiation of *Drosophila* crystal cell clusters and also participates in the coagulation function performed by crystal cells (57–59). In shrimp, it is hypothesized that Notch1 might mediate the differentiation of Cluster 0 cells and also perform a certain function in hemocyte coagulation. The high expression profiles of TGase and Notch1 implied that the Cluster 0 is likely to be the main cellular type for hemocyte clotting and is in the early stage of hemocyte differentiation. C-type lectin 2 (CTL-2), C-type lectin 2 like (CTL-2L), single VWC domain protein 5 (SVC5) were identified as the marker genes for Cluster 1 (Figures 2A, B). Lectins are considered the primary candidates for pattern recognition receptors (PRRs) in innate immunity because of their ability to bind specific carbohydrates on the surface of microorganisms (60–62). Calcium-dependent lectins (C-type lectin) are members of the typical lectin family with two or three pairs of disulfide-bonded characteristic carbohydrate recognition structural domains (CRDs) that exert pathogen recognition and clearance (62). A variety of C-type lectins with broad antibacterial and antiviral activity have been characterized in invertebrates, such as *Penaeus japonicus* and *L. vannamei* (63–65). SCV5 is a member of the von Willebrand factor type C (VWC) domains protein family, which play an important role in response to pathogen infestation (66, 67). Based on the analysis of key genes described above, we speculated that Cluster 1 might be the primary cell type that

could recognize pathogens. Crustins (Crustin like, Crustin, Crustin 2) are specific marker genes of Cluster 2 (Figures 2A, B). As a family of antimicrobial peptides (AMP) rich in cysteines and with several structural domains of whey acidic proteins (WAP), Crustin is widespread in decapod crustaceans (68–70). Due to its important antimicrobial role, Crustin is also known as a key component of the innate immune system of shrimp (71, 72). Thus, Cluster 2 might be the hemocyte type that perform the primary immune functions. Cluster 3 is the most specific kind of hemocytes, with the lowest number of cells. In the downscaling analysis, Cluster 3 is the most distant from the other cell clusters and its genes are largely unannotated, most likely being non-hemocyte (Figures 2A, B). The non-hemocyte has also been reported in previous single cell sequencing study of *Drosophila* hemocytes (34). The results of RNA-FISH further confirmed the distribution of the above-mentioned marker genes in hemocytes (Figure 2C). In summary, we defined the above four hemocyte clusters as TGase<sup>+</sup> cells, CTL<sup>+</sup> cells, Crustin<sup>+</sup> cells and non-hemocyte, respectively (Figures 2D, E).

## Functional prediction of hemocyte clusters

To further explore the functional characteristics of the above hemocyte clusters, we first performed GO functional annotation and KEGG pathway enrichment analysis of the differential genes among the clusters. According to the GO functional annotation, the differential genes of the TGase<sup>+</sup> cell cluster were mainly



**FIGURE 2** Identification of highly specific cell clusters marker genes and definition of cell clusters. **(A)** Dot plots showing the expression of three specific cell type marker genes across all clusters. Color gradient of the dot represents the expression level, while the size represents percentage of cells expressing any gene per cluster. **(B)** Stacked violin plots showing the expression of 10 commonly used cell type marker genes across all clusters. See [Supplementary Figure 3](#), [Supplementary File 3](#). **(C)** Validation of the specific marker genes expression in hemocytes by FISH. **(D)** The t-SNE dimensional reduction was used to visualize the classification results of hemocytes. **(E)** The UMAP was utilized for the re-clustering analysis and visualization of hemocytes.

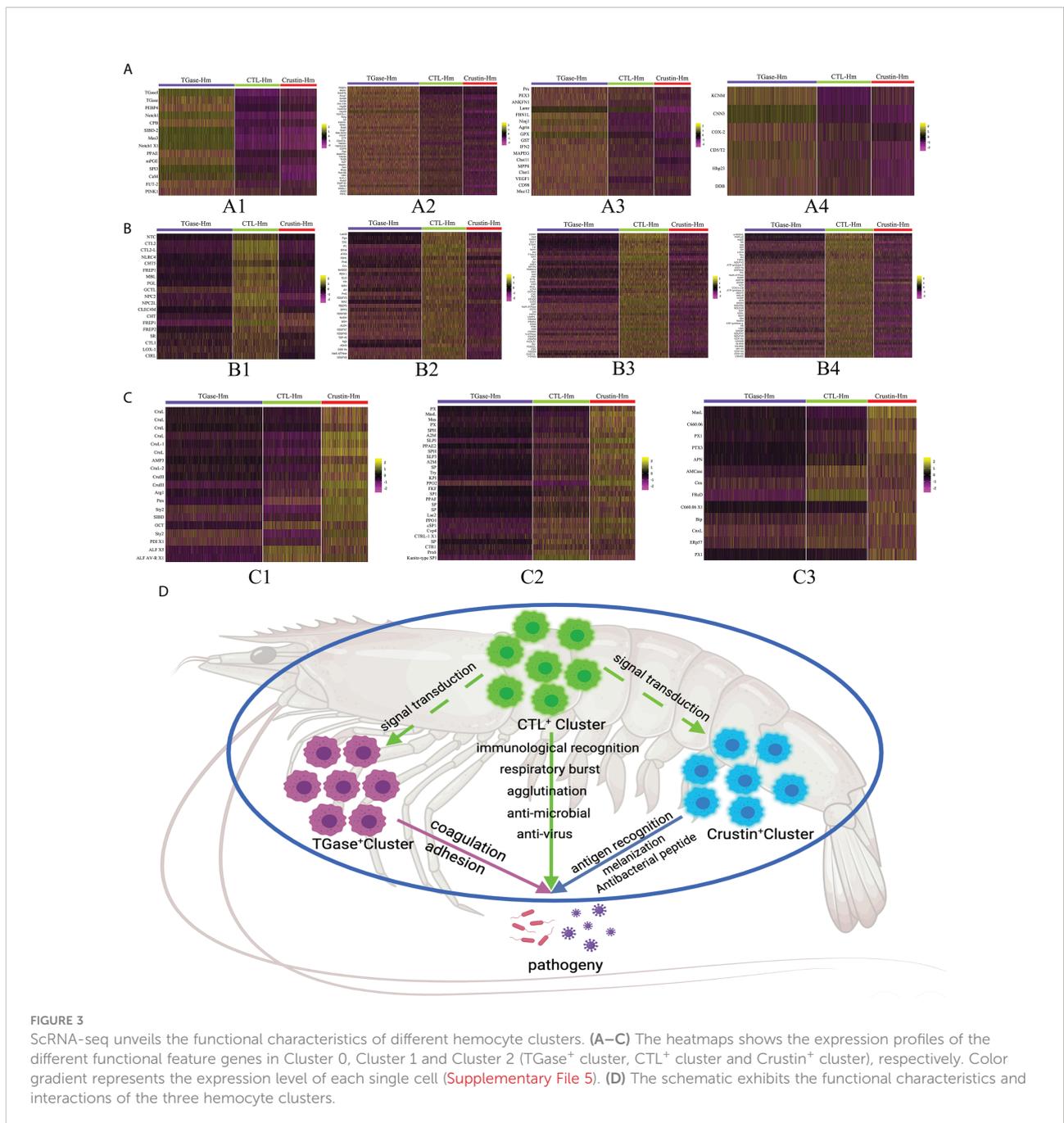
assigned to 13 categories of biological processes, six categories of molecular functions and 21 categories of cellular components. The KEGG pathway enrichment showed that the differential genes of the TGase<sup>+</sup> cell cluster were significantly enriched in 21 biological pathways such as ribosome (ko03001), spliceosome (ko03040) and VEGF signaling pathway (ko04370). The differential genes of the CTL<sup>+</sup> cell cluster were mainly enriched in GO annotations to 14 categories of biological processes, 20 categories of molecular function and 14 categories of cellular component and were assigned to 18

biological pathways including oxidative phosphorylation (ko03050) in the KEGG functional enrichment. The differential genes of the Crustin<sup>+</sup> cell cluster were enriched to a total of 15 classes of biological processes and nine classes of molecular function, and the cellular components were not enriched. The KEGG enrichment result of the Crustin<sup>+</sup> cell cluster showed that its differential genes were mainly enriched in seven pathways including antigen treatment (ko0462) ([Supplementary Figure 5](#); [Supplementary File 2](#)). The GO term and KEGG pathways enriched for non-hemocyte differential genes was too complex,

again illustrating the specificity of this cluster. As previously reported, GO and KEGG functional annotation did not provide specific functions of the cell clusters, probably due to the lack of information on the annotation of the relevant genes (73).

Then, we performed more detailed analysis of the differential genes that were significantly upregulated in each cluster to comprehensively reveal the functional characteristics of the four types of cell clusters (Supplementary File 3). Based on the scRNA-seq analysis, we found that TGase<sup>+</sup> cells significantly overexpressed genes associated with coagulation, such as

phosphatidylethanolamine-binding protein 4-like (PEBP4), carboxypeptidase B-like (CBP) and single insulin-like growth factor-binding domain protein-2 (SIBD-2) (Figure 3A). These molecules could act on TGase to promote its coagulation function or participate in hemocyte coagulation by regulating the cross-linking of other coagulation proteins (74–77). TGase has been reported to be secreted by the exosome pathway to exocytosis for its function. Significantly high expression of genes involved in the exosome pathway was also found in TGase<sup>+</sup> cells, suggesting that this cluster is prominently functioned by the TGase (Figure 3A)



(78–81). Meanwhile, TGase<sup>+</sup> cells also showed high expression of genes related to cell adhesion, which is capable of aiding hemocyte coagulation and the wound healing process (Figure 3A) (82–84). Numbers of genes involved in blood coagulation of higher animals were also highly expressed in the cluster, but the roles of their auxiliary coagulation function remain to be investigated (Figure 3A) (85, 86). Based on the results of these analyses, we ventured to speculate that TGase<sup>+</sup> hemocyte primarily perform the clotting function.

The analysis of highly expressed genes showed significantly upregulated expressions of lectin-like genes in CTL<sup>+</sup> cells, including C-type lectin, fibrinogen and serine protease-like proteins (Figure 3B). Lectins are the most thoroughly studied pattern recognition receptors in shrimp, sensing pathogen invasion and activating the immune response. Lectins are also involved in other biological processes, such as intercellular interactions, protein synthesis and signal transduction (87, 88). Moreover, oxidative stress genes were found to be highly expressed in CTL<sup>+</sup> cells along with certain immune-related genes (Figure 3B). Oxidative stress is the process of excess reactive oxygen species (ROS) production resulting from the imbalance between cellular oxidative and antioxidant reactions. ROS is produced as an intermediate product of oxidative respiration in mitochondria as well as many metabolic processes, which is involved in the defense of immune cells against invading pathogens (89, 90). The presence of significantly high expression of immune-related genes, such as lysosomal enzyme-related molecules and anti-lipopolysaccharide factors (ALFs), indicated that CTL<sup>+</sup> cell cluster play the anti-infective function and generate immune responses. Further analysis revealed that the expression of cellular metabolism-related molecules, especially oxidative phosphorylation-related genes, were significantly higher in the CTL<sup>+</sup> cell than in other clusters (Figure 3B). Oxidative phosphorylation is the key component of mitochondrial respiration and is intimately linked to ROS production, again suggesting that this cluster is active in oxidative stress processes (91). Combined with the significantly upregulated genes described above, the CTL<sup>+</sup> cell cluster is likely to be the key cell type that recognizes pathogens and produces ROS and immune molecules for anti-infective functions, while the high expression of its signal transduction-related genes allows the cluster cells to transmit triggering signals to stimulate other cell types to perform different immune functions (Figure 3B).

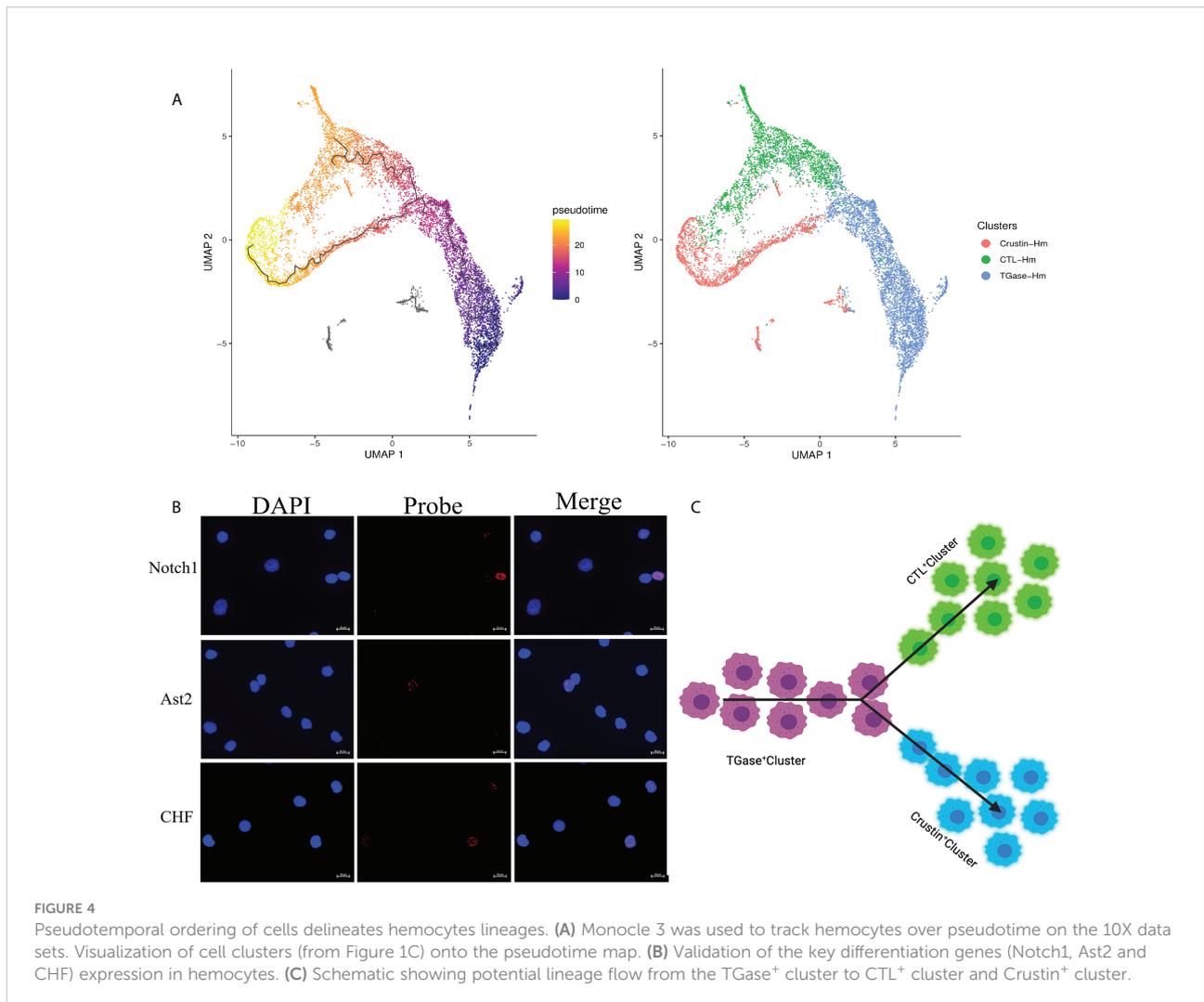
The Crustin<sup>+</sup> cell cluster had the lowest number of significantly up-regulated genes compared to the two previous hemocyte clusters. Analysis revealed that the genes significantly overexpressed in this cluster consisted mainly of various antimicrobial peptide (AMPs), and molecules associated with the phenoloxidase system (Figure 3C). AMPs, as the main molecules of the innate immune response in crustaceans, are able to fight against various pathogenic invasions (92). The phenoloxidase system is essential for the production of melanin in shrimp hemocytes, preventing further infection by

pathogens (93). The high expression of these genes implied that the Crustin<sup>+</sup> cell cluster is the main cell type exerting immune responses. Also, we found that some of the classical pattern recognition receptors, such as Toll-like receptors and F-type lectins, were highly expressed in Crustin<sup>+</sup> cells, which indicated that this cell cluster also functions in immune recognition (Figure 3C). The high expression of immune recognition genes in this cluster might be required for AMP molecules production and activation of the phenoloxidase system (94, 95). Based on the functional characteristics of these highly expressed genes, we considered that the Crustin<sup>+</sup> cell cluster is the main immune response cell and plays the critical anti-disease function.

In summary, we hypothesized that TGase<sup>+</sup> cells are the main cell type in hemolymph that function as hemocyte coagulation and cell adhesion. CTL<sup>+</sup> cells are the main immune recognition cells that could recognize foreign pathogens and transmit immune signals to other cell clusters, as well as acting in certain immune response roles. And Crustin<sup>+</sup> cells are the main immune response cells that play essential roles in fighting pathogens (Figure 3D). Taken together, the three types of cells form the shrimp hemocyte immune system that defends against pathogenic invasion.

## Pseudotemporal ordering and trajectory reconstitution of hemocyte lineages

To investigate the kinetics of hemocyte differentiation and to reveal the cell lineage relationships among different hemocyte clusters, we used Monocle3 functions to perform pseudotime analysis and cell lineage-tree reconstruction of the obtained hemocyte clusters. The pseudotime analysis assigned TGase<sup>+</sup> cluster as the starting point of the pseudotime intervals and showed that two lineages (CTL<sup>+</sup> cluster and Crustin<sup>+</sup> cluster) emerge from the start point (Figure 4A). Meanwhile, the pseudotime analysis showed distinct intermediate states of cells among the three types of hemocyte clusters, suggesting that the differentiation process of hemocytes was continuous, which was consistent with the hematopoietic mechanism of crustacean (53). We analyzed the cell differentiation and cell cycle related highly expressed genes of each cell cluster and the results further confirmed the accuracy of the pseudotime analysis. It has been previously clarified that high expression level of TGase in TGase<sup>+</sup> cells could be served as a marker for immature cells in crayfish (52, 96), suggesting that it would be logical to consider TGase<sup>+</sup> cluster as a starting point for pseudotemporal analysis. The Notch signaling pathway functions as a crucial pathway for *Drosophila* hemocytes differentiation (97), and the high expression of differentiation-related genes in this cluster suggest that TGase<sup>+</sup> cells are likely to act as pre-differentiated cells with differentiating and maturing in response to the differentiation-related signaling pathways. However, there is no clear evidence of whether the Notch signaling pathway contributes to cell development in crustaceans, particularly in



shrimp (98). We also found a number of cell cycle-related genes exhibited highly expressions in TGase<sup>+</sup> cells, such as AKB and GATA6 (Supplementary Figure 6; Supplementary File 4). These genes are extremely important in regulating the cell cycle by stabilizing chromosomes, promoting DNA replication and controlling the process of chromosome and cytoplasm splitting during cell division (99–101). These results indicated that the TGase<sup>+</sup> cell cluster has the ability to divide continuously, which is in line with the inference that the cluster is in the early stage of differentiation and confirms the plausibility of this cluster as the presumed starting point for cell differentiation. Based on the pseudotemporal differentiation analysis, the CTL<sup>+</sup> cluster and Crustin<sup>+</sup> cluster were at the end of the differentiation spectrum as terminally differentiated hemocytes. In the differential genes analysis, a lower number of differentiation-related genes was found in these two types of clusters. However, pivotal genes associated with maturation of crustacean hemocytes were detected both in two cell clusters (Supplementary Figure 6; Supplementary File 4). Astakine2 (Ast2), a homolog of vertebrate

hematopoietic factors, is highly expressed in the CTL<sup>+</sup> cluster and plays a key role in the maturation of shrimp hemocytes, while Ast2 has been shown to be essential for the differentiation of the granulosa cell lineage in crayfish (102). VEGF-3 and Sulf, which have been shown to be involved in *Drosophila* hematopoietic cell differentiation and maturation (103, 104), are also highly expressed in CTL<sup>+</sup> cluster. In addition, several molecules implicated in precursor cell differentiation and maturation in vertebrates such as BMPER, Elk-3 and HDAC were also significantly upregulated in the CTL<sup>+</sup> cluster (105, 106). Crustacean hematopoietic factor (CHF) and its analogue were highly expressed in the Crustin<sup>+</sup> cluster, which are critical for the maturation of hematopoietic tissue cells and is an important regulatory molecule in the development of the crayfish granulocyte lineage (Supplementary Figure 6; Supplementary File 4) (107). The presence of these maturation-related molecules implied the existence of cells in the intermediate state of differentiation and suggested that CTL<sup>+</sup> cluster and Crustin<sup>+</sup> clusters are indeed the end-state cells of the shrimp hemocytes lineage development. Collectively, the results of the pseudotime

analysis identified TGase<sup>+</sup> cluster arrested at an early developmental state with the ability to differentiate into CTL<sup>+</sup> clusters and Crustin<sup>+</sup> clusters, and that this differentiation process was continuous (Figures 4B, C).

## Diversity of hemocyte sub-clusters and their transcriptional dynamics

The previous pseudotime analysis found continuous differentiation among the obtained three hemocyte clusters, along with intermediate state differentiated hemocytes types. In order to further obtain the cell types of intermediate state, we subclassified TGase<sup>+</sup> cluster, CTL<sup>+</sup> cluster and Crustin<sup>+</sup> cluster, acquired 10 classes of sub-clusters and resolved their gene expression profiles (Supplementary File 5) and differentiation relationships.

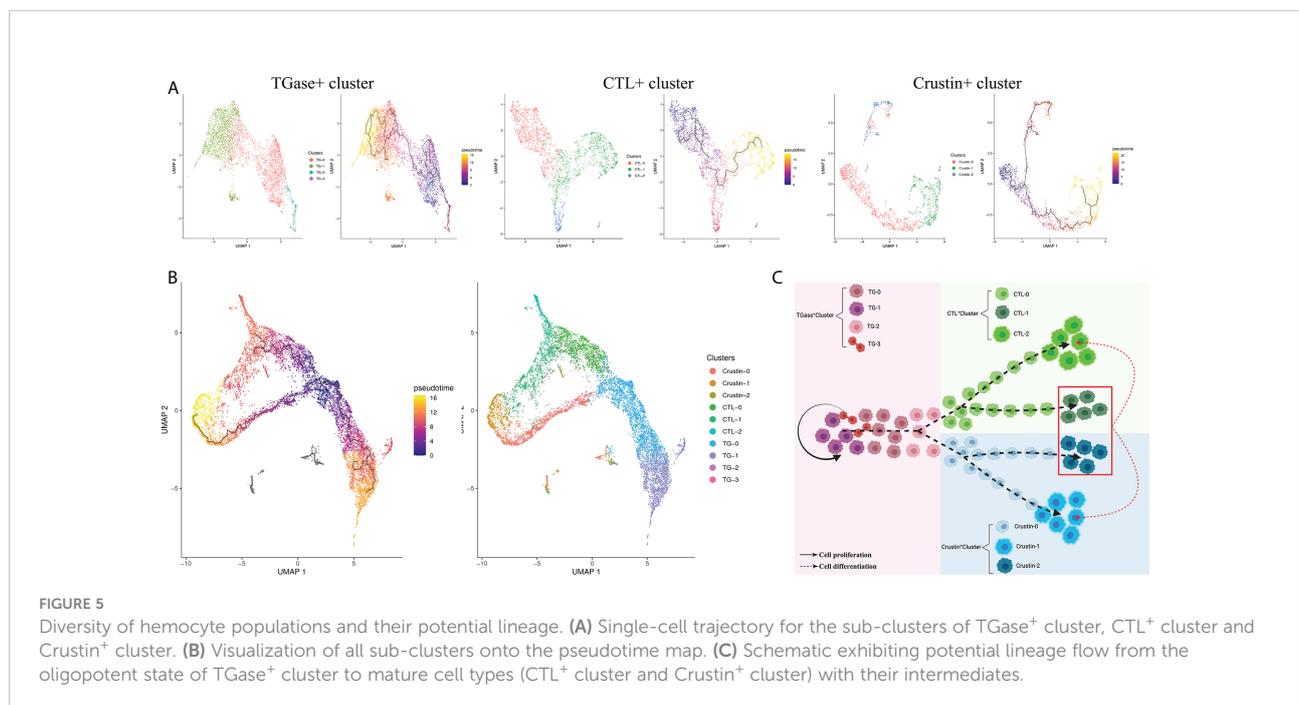
### TGase<sup>+</sup> cluster

Sub-clustering of the TGase<sup>+</sup> cluster obtained a total of four subclusters, named TG-0, TG-1, TG-2 and TG-3 (Supplementary Figure 7). The differential genes analysis of the four subclusters revealed that TG-0 was highly expressed the genes involved in cell cycle and differentiation, such as ATRX, TTF1 and LIG1 (Supplementary Figure 8) (108–110). In parallel, TG-0 was assigned as the presumed starting point for differentiation in a pseudotime differentiation analysis of the TGase<sup>+</sup> cell subcluster, suggesting that TG-0 might be the differentiated precursor cells for the TGase<sup>+</sup> cluster with the ability to continue differentiation

(Figure 5A). Compared to other subclusters, TG-1 has high in expressions of hemocyte coagulation-related molecules, especially TGase, which indicated that this subcluster might be the main sub-type that exerts coagulation function. Some genes that are involved in the regulation of cell differentiation and division also appeared to be hyper-expressed, and TG-1 should be able to regulate the differentiation or division state of certain cells. TG-2 showed a different differentiation direction from TG-1 in the reconstruction of differentiation trajectories and was enriched for genes that promote cell migration to the correct location during cell differentiation (Supplementary Figure 8). In addition, TG-2 showed upregulated expression of several functional genes of CTL<sup>+</sup> cluster and Crustin<sup>+</sup> clusters, implying that it is likely to be an intermediate state cell in the differentiation stage from TGase<sup>+</sup> clusters to CTL<sup>+</sup> or Crustin<sup>+</sup> clusters (Supplementary Figure 8). We determined TG-3 as the cell cycling or self-renewing state of TGase<sup>+</sup> cluster based on the expression of genes which are indispensable marker of the cell cycle such as CCNA1, zmc3 and INCENP (111–113).

### CTL<sup>+</sup> cluster

The CTL<sup>+</sup> clusters were re-clustered to obtain CTL-0, CTL-1 and CTL-2 subclusters, and the three subclusters were analyzed for pseudotime differentiation (Figure 5A; Supplementary Figure 7). The results show that CTL-0 was assigned as the starting point for pseudo-temporal differentiation. And molecules acting on invertebrate hemocyte differentiation such as Notch1 and VEGF3 were highly expressed in CTL-0 (97, 103), which confirmed the plausibility of the pseudotime differentiation for the CTL<sup>+</sup> cluster



(Supplementary Figure 8). This subcluster also showed high expression levels of genes related to mitochondrial function. In the CTL-1 subcluster, we found high expressions of immune defense genes, especially AMPs and molecules of regulating the phenoloxidase system (Supplementary Figure 8). We speculated that this subcluster is probably analogous to Crustin<sup>+</sup> cluster with immune function. The high expression of CHF in CTL-1 confirmed our speculation that the subcluster is likely to be an intermediate state cell type which differentiates into distinct cells in response to various differentiation stimulating molecules (Supplementary Figure 8) (107). The CTL-2 showed relatively simple gene expression profiles with signature genes largely engaged in immune recognition as well as signal transduction, which supported the view that the subcluster is the main functional cells of CTL<sup>+</sup> cluster (Supplementary Figure 8).

## Crustin<sup>+</sup> cluster

The Crustin<sup>+</sup> cluster was divided into three subclusters, Crustin-0, Crustin-1 and Crustin-2 (Supplementary Figure 7). The cell differentiation-related molecules, particularly CHF, were significantly highly expressed in Crustin-0, suggesting that the subcluster might be used as putative starting point for differentiation of Crustin<sup>+</sup> cluster (Figure 5A). Crustin-1 was probably the effector cell type of the Crustin<sup>+</sup> cluster, and its high expression profiles of AMPs and phenoloxidase-related molecules lent support to this conclusion (Supplementary Figure 8). Interestingly for Crustin-2, the subcluster had both cellular differentiation genes such as Notch1, VEGF3 and functional genes in other clusters like TGase and CTL. Crustin-2 seemed to be analogous with CTL-1, and the subcluster had divergent differentiation directions in response to stimulation by differentiation factors (Supplementary Figure 8).

Based on the above analysis, we performed a pseudo-temporal differentiation analysis of the 10 secondary subpopulations here, and the results were presumably in line with what we had hypothesized (Figure 5B). TG-0, as the putative starting point for differentiation, could differentiate into TG-1 for hemocyte coagulation on the one hand, and TG-2 for migration and differentiation into other cell clusters on the other. TG-3 acts as self-renewing cells for TGase<sup>+</sup> cluster and is capable of dividing and proliferating. TG-2 is progressively differentiated into CTL-0 in response to stimulation by Notch-related molecules, and CTL-0 further differentiate into mature CTL-2 for immune recognition and signal transduction. Crustin-0 is also supposed to differentiate and migrate from TG-2, which differentiates into mature Crustin-1 under the involvement of CHF for performing immune defense functions. Crustin-2 and CTL-1 are similar in that both might differentiate into different subclusters in response to stimulation by differentiation factors (Figure 5C).

## The relationship between hemocyte clusters and morphology

To correlate the obtained hemocyte clusters with morphological subpopulations, we sorted the hemocytes using flow cytometry. Two types of hemocyte subpopulations were defined and sorted according to relative cell size (forward scatter plot, FSC) and subcellular organelle characteristics/relative cell granularity (lateral scatter plot, SSC) (Figure 6A). The FCM showed that the hemocytes in region 1 are small in relative cell size and have the low intracellular complexity (R1, 31.7% ± 5.3%). Region 2 (R2) contained larger cells with high intracellular complexity (49.2% ± 4.6%). The two types of subpopulations obtained by sorting were stained to visualize histological features. The hemocytes in R1 were relatively small, with a large nucleoplasmic ratio, and the interior of the cell was basically free of granular material, similar to the morphological characteristics of hyalinocytes (Figure 6B) (73, 114). The hemocytes in R2 were slightly larger, with a smaller nucleoplasmic ratio and more granules inside the cells, resembling the morphological characteristics of granule-containing hemocytes (Figure 6B) (73, 114).

In order to determine the relationship between the two hemocyte subpopulations obtained by flow cytometry sorting and the hemocyte clusters gained by scRNA-seq, qRT-PCR was performed to explore the expression characteristics of marker genes and crucially differentiated genes. The results showed that marker genes of TGase<sup>+</sup> cell cluster such as TGase1 and Notch1 were mainly expressed in hyalinocytes (R1), while CTL<sup>+</sup> as well as Crustin<sup>+</sup> cell clusters marker genes like CTL-2, SVC5, and Crustin were primarily expressed in granule-containing hemocytes (R2) (Figure 6C; Supplementary File 6). Moreover, the distribution of important genes in different subpopulations were further determined by RNA-FISH. TGase1 and TGase were predominantly found in hyalinocytes. CTL-2, SVC5, CrustinL and Crustin were mainly observed in granule-containing hemocytes (Figure 6D). Based on these data, we can reasonably conclude that the TGase<sup>+</sup> cells are morphologically characterized as hyalinocytes and the CTL<sup>+</sup> cells and Crustin<sup>+</sup> cells are likely to be the granule-containing hemocytes.

## Discussion

The hemocytes are the vital immune cells in shrimp, thus the classification and identification of hemocytes are essential for better elucidating the innate immune mechanisms of shrimp. Studies on the classification of shrimp hemocytes began as early as the 19th century (14, 115). Although the currently agreed classification system for hemocytes is based primarily on morphological characteristics, the different criteria researchers have adopted to describe and identify hemocyte has led to a

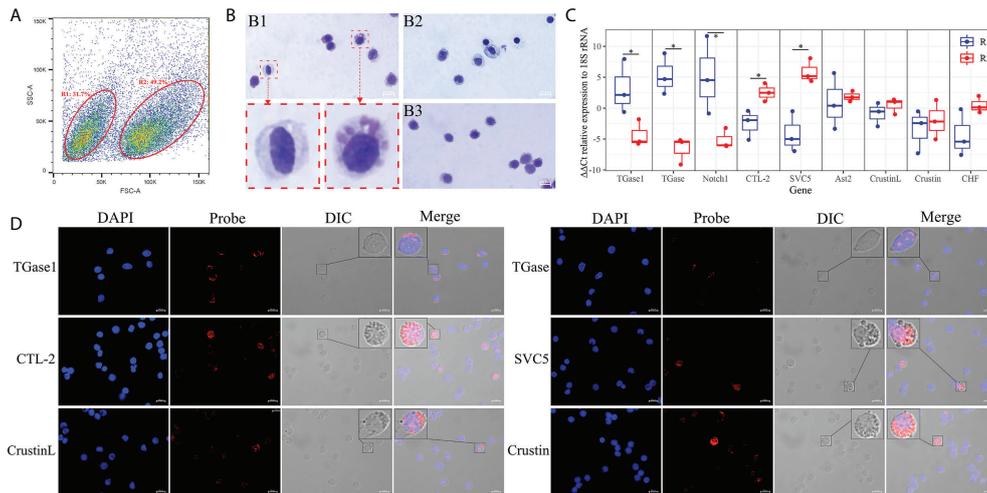


FIGURE 6

Morphological analysis of hemocytes clusters. (A) FACS analysis of hemocytes. Based on the FSC and SSC, two regions (R1 and R2) were sorted, respectively. (B) Hemocytes and cell subpopulations were stained by May-Grunwald and Giemsa dye. B1: hemocytes. B2: R1 sorted hemocytes. B3: R2 sorted hemocytes. (C) Differential expression analysis of highly specific cell clusters gene between R1 and R2 of hemocytes sorted using FACS. (D) Validation of the highly specific cell clusters gene among hemocyte subpopulations by FISH. The asterisks on the bars represent the statistical difference of the hemocyte subpopulations.

highly confusing classification system for hemocytes (11, 116–118). Subsequent researchers have proposed interpretation of hemocyte types through multiple perspectives on the basis of cell morphology combined with cytochemistry. Some studies have also established new means for hemocyte classification according to biochemical markers on the cell surface, with reference to the classification criteria adopted for mammalian immune cells (19, 20). Yet, there is still no consistent agreement on the classification of crustacean hemocytes. In addition, there is much controversy about the differences in immunological functions performed by different hemocyte subpopulations, such as cytotoxicity (11, 117, 119). The developmental characteristics of circulating hemocyte lineages have not been fully elucidated (21, 22). In the present work, the gene expression characteristics of single hemocyte were obtained by high-throughput sequencing with the aid of scRNA-seq technology, and the classification features of hemocytes from healthy *Litopenaeus vannamei* was comprehensively analyzed in terms of molecular features, with a view to elucidating the functions of different hemocyte clusters and their differentiation relationships.

Through scRNA-seq analysis, four hemocyte clusters were identified based on their differential gene expression profiles. Due to the extremely small cell number of Cluster 3 and the fact that most of the highly expressed genes in the cluster have not been annotated, we have tentatively defined it as a non-hematopoietic cluster. Similar non-hemocytes cluster have been also found in *Drosophila* blood and might be impurity cells (34). We made functional definitions of the other three hemocyte clusters using specific marker genes and functional

genes, identified the cell types and found that they do exhibit heterogeneous immune functions (120–123). Different numbers of cellular clusters have also been identified in previous studies on the scRNA-seq of shrimp hemocytes. Six cell clusters were identified in healthy *Marsupenaeus japonicus* that had not been subjected to any treatment (73). Under ammonia stress, hemocytes of *Penaeus monodon* were divided into seven main cell groups (124). Following stimulation by hemocyte activating factors, hemocytes of *L. vannamei* were classified into prohemocytes, granulocytes, monocytic hemocytes, transitional cells and germ-like cells (125). The difference in hemocyte classification results of scRNA-seq might be due to species specificity. In addition, our results did not find cell cluster with significantly higher expression of antiviral molecules such as Vago5 in healthy *L. vannamei* hemocytes, nor were the cluster reported in *P. monodon*, but it appeared in the second subpopulation of healthy *M. japonicus* hemocytes. Differences in hemocyte clustering might also be the result of various treatments, which could lead to significant upregulation of key genes in a cluster, resulting in different classifying outcomes, which is also common in scRNA-seq studies of other species, as reported in the case of hemocytes in *Drosophila* and mosquito (33, 34, 126, 127). It is also possible that the cell clustering differences are attributable to the fact that shrimp hemocytes are themselves a dynamic and changing tissue that are not fully mature, and that shrimp growth status, sampling time, sampling methods and analytical parameters might all contribute to the differences. Nevertheless, all scRNA-seq studies of shrimp hemocytes have emphasized the existence of cell clusters that

perform coagulation cascades, kill pathogens, and other distinct immune functions. Given our analysis, it is assumed that CTL<sup>+</sup> cells could recognize invading pathogens and initiate partial immune functions, while transmitting immune signals to other clusters. Crustin<sup>+</sup> cells recognize pathogens and receive immune signals, initiating immune response pathways, releasing molecules such as AMP and activating phenoloxidase system to directly kill pathogens. During this process, TGase<sup>+</sup> cells are stimulated by the immune signals and begin to release cellular agglutination factors that drive coagulation. The three types of hemocytes complement each other and form the shrimp hemocytes immune defense system. Interestingly, our results did not reveal differences in the expression profiles of phagocytosis-related genes among hemocyte clusters, indicating that all types of hemocytes have phagocytic potential (120, 128). In healthy *M. japonicus* hemocytes, the cell type that mainly exerts phagocytosis was also not clearly identified (73). Nevertheless, the hemocytes from treated *L. vannamei* and *P. monodon* exhibited cell clusters that exerted phagocytosis (124, 125), suggesting that the phagocytic cell types might be differentiated from a particular type of hemocytes, which also indicated that different treatments would lead to the differences of cell clustering.

For the purpose of clarifying the functional characteristics of hemocyte morphological subpopulations, we further developed the connectivity between the hemocyte clusters obtained by scRNA-seq with morphological subpopulations. Two hemocyte subpopulations, hyalinocytes and granule-containing hemocytes, were sorted out by flow cytometry, whereas semi-granulocytes and granulocytes did not appear to be identified by flow cytometry (73). May-Grunwald and Giemsa staining of the sorted granule-containing hemocytes was still able to reveal the presence of two classes of granulocytes with different granularity. With the aim of functionally validating the two morphologically distinct subpopulations, we detected the distribution of key functional genes by FISH and qPCR. The results showed that hyalinocytes shared similar gene expression profile with TGase<sup>+</sup> cluster, suggesting its main role in hemocyte coagulation and hemolymph adhesion. Granule-containing hemocytes highly expressed the genes involved in immune recognition, ROS burst, AMPs and phenoloxidase, indicating they executed immune response functions as CTL<sup>+</sup> and Crustin<sup>+</sup> cells. This conclusion is supported by previous studies that found TGase is mainly expressed in hyalinocytes, and semi-granulocytes take the lead in responding to foreign pathogens and AMPs are synthesized and secreted in granulocytes (36, 52, 129). Moreover, based on previously reported marker genes for different morphological subpopulations, recent scRNA-seq analysis of shrimp hemocyte also inferred that hyalinocytes are the main cell type that exert cellular agglutination and granule-containing hemocytes mainly perform the immune defense functions against pathogens (73, 124, 125). However, we still

lack direct evidence to develop the connectivity between CLT<sup>+</sup> cluster and Crustin<sup>+</sup> cluster with semi-granulocytes and granulocytes. We hypothesized that when pathogens invade, granule-containing hemocytes of shrimp are able to recognize foreign invaders and perform direct immune functions to kill them by releasing lectins, AMPs and other immune functional molecules and activate the phenoloxidase system. simultaneously, the granule-containing hemocytes could transmit immune signals and stimulate hyalinocytes to secrete cellular coagulation factors, thus inducing hemocytes coagulation, encapsulation and nodulation (Supplementary Figure 9A).

Subsequently, we mapped pseudotime differentiation trajectories of the obtained hemocyte clusters and their sub-clusters using the Monocle 3 function. By analyzing highly expressed differentiation-related genes, we identified TGase<sup>+</sup> cluster as being at an early stage of differentiation, capable of migrating and differentiating into CTL<sup>+</sup> cluster and Crustin<sup>+</sup> cluster. FISH and qPCR results further found that the TGase served as a marker for immature hemocytes in crustaceans is highly expressed in hyalinocytes (52), indicating that hyalinocytes are arrested at the early stage of hematopoietic differentiation. In hyalinocytes, a significant enrichment of molecules was found to be associated with the Notch signaling pathway that is the regulatory lynchpin of invertebrate hemocytes lineage development (130, 131), implicating that the Notch signaling pathway probably holds central roles in the induction of shrimp hyalinocytes differentiation. The CTL<sup>+</sup> and Crustin<sup>+</sup> clusters at the terminal states of the pseudotime differentiation trajectory showed high expressions of Ast2 and CHF respectively, which are two important regulators of granulocyte lineage maturation in shrimp (102, 132). On the basis of these results, we speculated that TGase<sup>+</sup> cluster (hyalinocytes) are first released in hematopoietic tissue (96) and begin to differentiate with the action of the Notch signaling pathway, followed by gradual development into mature CTL<sup>+</sup> cluster and Crustin<sup>+</sup> cluster (granule-containing hemocytes) under the effects of molecules such as Ast and CHF (Supplementary Figure 9B). This seems to be verified in shrimp hematopoietic tissues, where the precursor cells of granule-containing hemocytes exhibit characteristics of hyalinocytes (22, 96). This is also consistent with the studies on hemocytes of *M. japonicus* and *P. monodon*, in which cells highly expressing TGase were classified as early differentiating cell types, while cells highly expressing molecules of AMP and proPO were considered to be mature cell types (73, 124). Yang et al. defined that the prohemocytes of *L. vannamei* hemocytes exhibiting high expression of TGase were classified as early differentiated cell types, while they defined monocytic hemocyte and granulocytes with high expression of molecules such as AMP were classified as late differentiated cell types (125). Taken together, these evidence suggested that there are two developmental lineages of shrimp hemocytes that differentiate and mature from hyalinocytes into different granule-containing hemocytes, and semi-granulocytes

and granulocytes still show some differentiation association (96, 102, 107, 125). However, the differentiation relationship between granulocytes and semi-granulocytes still cannot be determined, since we are unable to arbitrarily judge how the CTL<sup>+</sup> cluster and Crustin<sup>+</sup> cluster correspond to granulocytes and semi-granulocytes.

To further characterize the presence of oligopotent/initial state cells, intermediate state cells and terminally differentiated effector cells in the hemocytes, we performed sub-clustering analysis of the three hemocyte clusters. The results showed that less than 1% of the TGase<sup>+</sup> cluster (hyalinocytes) are in the cytokinesis cycle (TG-3). This subcluster was in an oligopotent state and might be critical for the self-renewal of TGase<sup>+</sup> cluster. Previous studies have also discovered a small number of hemocytes undergoing the cell cycle in crustacean hemolymph (9, 125, 133). In parallel, we identified subclusters of hemocyte clusters in transient intermediate states. TG-0 highly expressed the differentiation related genes and was at the putative starting point for pseudotime differentiation analysis, which could be considered the point of initiation of differentiation. TG-2 showed the expression of functional genes of other hemocyte clusters (CTL<sup>+</sup> cells and Crustin<sup>+</sup> cells), suggesting the onset of branching points of differentiation. The CTL-0 and Crustin-0 subclusters contiguous to TG-2 actively express the genes that are involved in cell differentiation and migration, indicating that these hemocytes were gradually maturing at this juncture. Furthermore, three effector cell subclusters (TG-1, CTL-2, Crustin-1) were found to exert their specific immune functions. To be noted, the expression of some cell cycle-related genes could still be observed in TG-1, but not in CTL-2 and Crustin-1. Since terminally differentiated cells might require downregulation of cell cycle gene expression to arrest cell division, we hypothesized that CTL-2 and Crustin-1 are terminally differentiated cells whereas TG-1 is not (134, 135). The CTL-1 and Crustin-2 were the most complicated cell, and we assumed that these two types of subclusters are likely to have bidirectional differentiation potential in response to different differentiation factors (21, 136). According to the above analysis, hyalinocytes (TGase<sup>+</sup> cluster) are most probably dynamically changing hemocytes that do not have sub-clusters of terminally differentiated cells, which acts as differentiation precursor cells and gradually differentiates into mature granule-containing hemocytes with specific functions (CTL<sup>+</sup> cluster and Crustin<sup>+</sup> cluster) (Supplementary Figure 9C).

This study revealed the functional subtypes and differentiation characteristics of shrimp hemocytes by scRNA-seq and explored the association information between morphological cell subpopulations and functional clusters for the first time, which provided abundant data for the classification, identification and developmental lineages of shrimp hemocyte subpopulations and established a solid foundation for further investigations on the immune functions of hemocyte subpopulations.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI, accession ID: SRR18899521.

## Ethics statement

The animal study was reviewed and approved by the International Guiding Principles for Biomedical Research Involving Animals documented by Guide for the Use of Experimental Animals and the Committee of the Ethics on Animal Care and Experiments at Ocean University of China.

## Author contributions

CC and XT contributed to the conception and design of this study, performed most of experiments and statistical analysis, drafted and revised the manuscript; JX, XS, and HC participated in the design of the study, helped analyzed experiments and data; XT and WZ designed the study, edited the manuscript, and provided reagents and experiment space. All authors contributed to the article and approved the submitted version.

## Funding

This research was financially supported by the National Key Research and Development Program of China (2018YFD0900504, 2019YFD0900101), Qingdao National Laboratory for Marine Science and Technology (QNL2016ORP0307) and the Taishan Scholar Program of Shandong Province.

## Acknowledgments

We are grateful to Guangzhou Genedenovo Biotechnology Co., Ltd for assisting in sequencing and/or bioinformatics analysis. The schematics were created with [BioRender.com](https://BioRender.com).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.980021/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

Sequencing data quality control and expression quantification. **(A)** Identification of recovered cells. The abscissa is the number of barcode sequences, and the ordinate is the number of UMIs. The barcode corresponding to the green line is the effective cell, and the gray line is the background noise. **(B, C)** Distribution of basic information (the number of genes, the total number of UMI and mitochondrial gene expression percentage) of each sample cell before and after filtering. **(B)** Distribution of basic information of each sample cell before filtering; **(C)** Distribution of basic information of each sample cell after filtering. **(D–G)** Scatter plot of basic information of each sample cell before and after filtering. **(D, F)** scatter plot before filtering. **(E, G)** scatter plot after filtering. **(D, E)** Relationship between nUMI and nGene. **(F, G)** Relationship between nUMI and pMito. The X-axis is the number of UMIs, and the Y-axis is the number of genes/mitochondria of mitochondria. The numbers above the graph are Pearson correlation coefficients for the number of uMI and the number of genes of mitochondria.

### SUPPLEMENTARY FIGURE 2

Statistics on the number and percentage of different hemocyte clusters in cluster analysis.

### SUPPLEMENTARY FIGURE 3

Expression profiles of TOP20 genes in each hemocytes cluster. **(A)** The heatmap of 20 up-regulated expression genes from four cell populations as a marker gene. **(B)** The dot plot of 20 up-regulated expression genes from four cell populations as a marker gene.

### SUPPLEMENTARY FIGURE 4

Expression profiles of marker genes in each hemocytes cluster. **(A)** The heatmap of three marker genes from four cell clusters. **(B)** Expression pattern of the marker genes across cell clusters.

### SUPPLEMENTARY FIGURE 5

GO functional annotation and KEGG pathway enrichment analysis of the differential genes among the hemocyte clusters. **(A)** Cluster 0/TGase<sup>+</sup> cluster, **(B)** Cluster 1/CTL<sup>+</sup> cluster, **(C)** Cluster 2/Crustin<sup>+</sup> cluster, **(D)**

Cluster 3/non-hemocytes. **(A1–D1)** Biological processes in GO functional annotation of the differential genes of each cluster. **(A2–D2)** Molecular functions in GO functional annotation of the differential genes of each cluster.

### SUPPLEMENTARY FIGURE 6

Expression profiles of cell differentiation-related genes in each hemocytes cluster. **(A)** Heatmap showing the expression of genes related to the NOTCH signaling pathway among various clusters. **(B–D)** Differential expression of genes related to cell differentiation among cell clusters.

### SUPPLEMENTARY FIGURE 7

Statistics on the number and percentage of different sub-clusters in each hemocyte cluster. **(A)** Cluster 0/TGase<sup>+</sup> cluster, **(B)** Cluster 1/CTL<sup>+</sup> cluster, **(C)** Cluster 2/Crustin<sup>+</sup> cluster

### SUPPLEMENTARY FIGURE 8

Identification of specific cell sub-clusters. **(A)** TGase<sup>+</sup> cluster. **(B)** CTL<sup>+</sup> cluster. **(C)** Crustin<sup>+</sup> cluster.

### SUPPLEMENTARY FIGURE 9

Schematic showing the functional characteristics of hemocytes morphological subpopulations, and their potential lineage flow based on hemocyte cluster profiles. **(A)** Schematic plot shows the functional characteristics of hemocytes morphological subpopulations. **(B)** The potential lineage flow of hemocytes morphological subpopulations. **(C)** The potential lineage flow of hemocytes morphological subpopulations based on hemocyte sub-cluster profiles.

### SUPPLEMENTARY FILE 1

Top 20 marker genes for each cell type cluster. Shown are lists of the top 20 genes that show marker gene-like characteristics for each cell type cluster.

### SUPPLEMENTARY FILE 2

GO functional annotation and KEGG pathway enrichment analysis of the differential genes among the hemocyte clusters.

### SUPPLEMENTARY FILE 3

Differential expression profiles of key functional genes in various clusters.

### SUPPLEMENTARY FILE 4

Expression profiles of cell differentiation-related genes in each hemocytes cluster.

### SUPPLEMENTARY FILE 5

Expression profiles of genes for identifying the specific cell sub-clusters in each hemocytes cluster.

### SUPPLEMENTARY FILE 6

Source data for CT values of each gene in different hemocyte subpopulations.

## References

1. Stankus A. State of world aquaculture 2020 and regional reviews: FAO webinar series. *FAO Aquacult Newsletter* (2021) 63:17–8.
2. Sánchez-Muros MJ, Rentería P, Vizcaino A, Barroso FG. Innovative protein sources in shrimp (*Litopenaeus vannamei*) feeding. *Rev Aquacult* (2020) 12:186–203. doi: 10.1111/raq.12312
3. *Global aquaculture production 1950–2016* (2018). Available at: <http://www.fao.org/fishery/statistics/global-aquaculture-production/query/en>.
4. Stentiford GD, Neil DM, Peeler EJ, Shields JD, Small HJ, Flegel TW, et al. Disease will limit future food supply from the global crustacean fishery and aquaculture sectors. *J Invertebr Pathol* (2012) 110:141–57. doi: 10.1016/j.jip.2012.03.013
5. Aweya JJ, Zheng Z, Zheng X, Yao D, Zhang Y. The expanding repertoire of immune-related molecules with antimicrobial activity in penaeid shrimps: a review. *Rev Aquacult* (2021) 13:1907–37. doi: 10.1111/raq.12551
6. Ghosh J, Lun CM, Majeske AJ, Sacchi S, Schrankel CS, Smith LC. Invertebrate immune diversity. *Dev Comp Immunol* (2011) 35:959–74. doi: 10.1016/j.dci.2010.12.009
7. Cui C, Liang QR, Tang XQ, Xing J, Sheng XZ, Zhan WB. Differential apoptotic responses of hemocyte subpopulations to white spot syndrome virus

- infection in *fenneropenaeus chinensis*. *Front Immunol* (2020) 11:3174. doi: 10.3389/fimmu.2020.594390
8. Li FH, Xiang JH. Recent advances in researches on the innate immunity of shrimp in China. *Dev Comp Immunol* (2013) 39:11–26. doi: 10.1016/j.dci.2012.03.016
  9. Söderhäll I. Crustacean hematopoiesis. *Dev Comp Immunol* (2016) 58:129–41. doi: 10.1016/j.dci.2015.12.009
  10. Zanjani NT, Miranda-Saksena M, Cunningham AL, Dehghani F. Antimicrobial peptides of marine crustaceans: The potential and challenges of developing therapeutic agents. *Curr Med Chem* (2018) 25:2245–59. doi: 10.2174/0929867324666171106155936
  11. Johansson MW, Keyser P, Sritunyalucksana K, Söderhäll K. Crustacean haemocytes and haematopoiesis. *Aquaculture* (2000) 191:45–52. doi: 10.1016/S0044-8486(00)00418-X
  12. Vargas-Albores F, Gollas-Galván T, Hernández-López J. Functional characterization of farfantepenaeus californiensis, litopenaeus vannamei and *I. stylirostris* haemocyte separated using density gradient centrifugation. *Aquac Res* (2005) 36:352–60. doi: 10.1111/j.1365-2109.2004.01207.x
  13. Wu F, Xie Z, Yan M, Li Q, Song J, Hu M, et al. Classification and characterization of hemocytes from two Asian horseshoe crab species *tachyleus tridentatus* and *carcinuscorpius rotundicauda*. *Sci Rep* (2019) 9:1–10. doi: 10.1038/s41598-019-43630-8
  14. Söderhäll K, Smith VJ. Separation of the haemocyte populations of *carcinusmaenas* and other marine decapods, and prophenoloxidase distribution. *Dev Comp Immunol* (1983) 7:229–39. doi: 10.1016/0145-305X(83)90004-6
  15. Van de Braak CBT, Botterblom MHA, Liu W, Taverne N, van der Knaap WPW, Rombout JHWM. The role of the haematopoietic tissue in haemocyte production and maturation in the black tiger shrimp (*Penaeus monodon*). *Fish Shellfish Immunol* (2002) 12:253–72. doi: 10.1006/fsim.2001.0369
  16. Dantas-Lima JJ, Tuan VV, Corteel M, Grauwet K, An NTT, Sorgeloos P, et al. Separation of *penaeus vannamei* haemocyte subpopulations by iodixanol density gradient centrifugation. *Aquaculture* (2013) 408:128–35. doi: 10.1016/j.aquaculture.2013.04.031
  17. Estrada N, Velázquez E, Rodríguez-Jaramillo C, Ascencio F. Carbohydrate moieties and cytoenzymatic characterization of hemocytes in whiteleg shrimp *litopenaeus vannamei*. *Int J Cell Biol* (2016) 2016:1–9. doi: 10.1155/2016/9032181
  18. Koiwai K, Kondo H, Hirono I. Isolation and molecular characterization of hemocyte sub-populations in kuruma shrimp *marsupenaeus japonicus*. *Fish Sci* (2019) 85:521–32. doi: 10.1007/s12562-019-01311-5
  19. Sung H, Sun R. Use of monoclonal antibodies to classify hemocyte subpopulations of tiger shrimp (*Penaeus monodon*). *J Crustac Biol* (2002) 22:337–44. doi: 10.1163/20021975-99990240
  20. Winotaphan P, Sithigorngul P, Muenpol O, Longyant S, Rukpratanporn S, Chaivisuthangkura P, et al. Monoclonal antibodies specific to haemocytes of black tiger prawn *penaeus monodon*. *Fish Shellfish Immunol* (2005) 18:189–98. doi: 10.1016/j.fsi.2004.05.007
  21. Liu MJ, Liu S, Liu HP. Recent insights into hematopoiesis in crustaceans. *Fish Shellfish Immunol Rep* (2021) 2:100040. doi: 10.1016/j.fsirep.2021.100040
  22. Li F, Zheng Z, Li H, Fu R, Xu L, Yang F. Crayfish hemocytes develop along the granular cell lineage. *Sci Rep* (2021) 11:1–16. doi: 10.1038/s41598-021-92473-9
  23. Kurucz É, Váczi B, Márkus R, Laurinyecz B, Vilmos P, Zsámboki J, et al. Definition of *drosophila* hemocyte subsets by cell-type specific antigens. *Acta Biol Hung* (2017) 58:95–111. doi: 10.1556/ABiol.58.2007.Suppl.8
  24. Xing J, Chang YH, Tang XQ, Sheng XZ, Zhan WB. Separation of haemocyte subpopulations in shrimp *fenneropenaeus chinensis* by immunomagnetic bead using monoclonal antibody against granulocytes. *Fish Shellfish Immunol* (2017) 60:114–8. doi: 10.1016/j.fsi.2016.11.034
  25. Macosko EZ, Basu A, Satija R, Nemes J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* (2015) 161:1202–14. doi: 10.1016/j.cell.2015.05.002
  26. Satija R, Shalek AK. Heterogeneity in immune responses: from populations to single cells. *Trends Immunol* (2014) 35:219–29. doi: 10.1016/j.it.2014.03.004
  27. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* (2019) 566:496–502. doi: 10.1038/s41586-019-0969-x
  28. Trapnell C. Defining cell types and states with single-cell genomics. *Genome Res* (2015) 25:1491–8. doi: 10.1101/gr.190595.115
  29. Wagner A, Regev A, Yosef N. Revealing the vectors of cellular identity with single-cell genomics. *Nat Biotechnol* (2016) 34:1145–60. doi: 10.1038/nbt.3711
  30. Drokhllyansky E, Smillie CS, Van Wittenbergh E, Ericsson M, Griffin GK, Eraslan G, et al. The human and mouse enteric nervous system at single-cell resolution. *Cell* (2020) 182:1606–22. doi: 10.1016/j.cell.2020.08.003
  31. Fu Y, Huang X, Zhang P, van de Leemput J, Han Z. Single-cell RNA sequencing identifies novel cell types in *drosophila* blood. *J Genet Genomics* (2020) 47:175–86. doi: 10.1016/j.jgg.2020.02.004
  32. Zhang TQ, Xu ZG, Shang GD, Wang JW. A single-cell RNA sequencing profiles the developmental landscape of *arabidopsis* root. *Mol Plant* (2019) 12:648–60. doi: 10.1016/j.molp.2019.04.004
  33. Raddi G, Barletta ABF, Efremova M, Ramirez JL, Cantera R, Teichmann SA, et al. Mosquito cellular immunity at single-cell resolution. *Science* (2020) 369:1128–32. doi: 10.1126/science.abc0322
  34. Tattikota SG, Cho B, Liu Y, Hu Y, Barrera V, Steinbaugh MJ, et al. A single-cell survey of *drosophila* blood. *eLife* (2020) 9:e54818. doi: 10.7554/eLife.54818
  35. Meng J, Wang WX. Highly sensitive and specific responses of oyster hemocytes to copper exposure: Single-cell transcriptomic analysis of different cell populations. *Environ Sci Technol* (2022) 56:2497–510. doi: 10.1021/acs.est.1c07510
  36. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* (2013) 29:15–21. doi: 10.1093/bioinformatics/bts635
  37. Zhang X, Yuan J, Sun Y, Li S, Gao YI, Yu Y, et al. *Penaeid* shrimp genome provides insights into benthic adaptation and frequent molting. *Nat Commun* (2019) 10:1–14. doi: 10.1038/s41467-018-08197-4
  38. Chung NC, Storey JD. Statistical significance of variables driving systematic variation in high-dimensional data. *Bioinformatics* (2015) 31:545–54. doi: 10.1093/bioinformatics/btu674
  39. McDavid A, Finak G, Chattopadhyay PK, Dominguez M, Lamoreaux L, Ma SS, et al. Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* (2013) 29:461–7. doi: 10.1093/bioinformatics/bts714
  40. Perrin D, Zuccon G. Recursive module extraction using louvain and PageRank. *F1000Res* (2018) 7:1286. doi: 10.12688/f1000research.15845.1
  41. Fonville JM, Carter CL, Pizarro L, Steven RT, Palmer AD, Griffiths RL, et al. Hyperspectral visualization of mass spectrometry imaging data. *Anal Chem* (2013) 85:1415–23. doi: 10.1021/ac302330a
  42. Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. *Bioinformatics* (2007) 23:257–8. doi: 10.1093/bioinformatics/btl567
  43. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* (2000) 28:27–30. doi: 10.1093/nar/28.1.27
  44. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IW, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* (2019) 37:38–44. doi: 10.1038/nbt.4314
  45. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WMIII, et al. Comprehensive integration of single-cell data. *Cell* (2019) 177:1888–902. doi: 10.1016/j.cell.2019.05.031
  46. Niu J, Huang Y, Liu X, Zhang Z, Tang J, Wang B, et al. Single-cell RNA-seq reveals different subsets of non-specific cytotoxic cells in teleost. *Genomics* (2020) 112:5170–9. doi: 10.1016/j.ygeno.2020.09.031
  47. Grenard P, Bates MK, Aeschlimann D. Evolution of transglutaminase genes: identification of a transglutaminase gene cluster on human chromosome 15q15: structure of the gene encoding transglutaminase X and a novel gene family member, transglutaminase z. *J Biol Chem* (2001) 276:33066–78. doi: 10.1074/jbc.M102553200
  48. Lorand L, Graham RM. Transglutaminases: crosslinking enzymes with pleiotropic functions. *Nat Rev Mol Cell Biol* (2003) 4:140–56. doi: 10.1038/nrm1014
  49. Chen YN, Chen WC, Cheng W. The second type of transglutaminase regulates immune and stress responses in white shrimp, *litopenaeus vannamei*. *Fish Shellfish Immunol* (2014) 37:30–7. doi: 10.1016/j.fsi.2014.01.010
  50. Maningas MBB, Kondo H, Hirono I. Molecular mechanisms of the shrimp clotting system. *Fish Shellfish Immunol* (2013) 34:968–72. doi: 10.1016/j.fsi.2012.09.018
  51. Yeh MS, Liu CH, Hung CW, Cheng W. cDNA cloning, identification, tissue localisation, and transcription profile of a transglutaminase from white shrimp, *litopenaeus vannamei*, after infection by *vibrio alginolyticus*. *Fish Shellfish Immunol* (2009) 27:748–56. doi: 10.1016/j.fsi.2009.09.006
  52. Lin X, Söderhäll K, Söderhäll I. Transglutaminase activity in the hematopoietic tissue of a crustacean, *pacifastacus leniusculus*, importance in hemocyte homeostasis. *BMC Immunol* (2008) 9:1–11. doi: 10.1186/1471-2172-9-58
  53. Noonin C, Lin X, Jiravanichpaisal P, Söderhäll K, Söderhäll I. Invertebrate hematopoiesis: an anterior proliferation center as a link between the hematopoietic tissue and the brain. *Stem Cells Dev* (2012) 21:3173–86. doi: 10.1089/scd.2012.0077
  54. Jin LH, Choi JK, Kim B, Cho HS, Kim J, Kim-Ha J, et al. Requirement of split ends for epigenetic regulation of notch signal-dependent genes during infection-induced hemocyte differentiation. *Mol Cell Biol* (2009) 29:1515. doi: 10.1016/S0965-1748(02)00092-9

55. Mukherjee T, Kim WS, Mandal L, Banerjee U. Interaction between notch and hif- $\alpha$  in development and survival of drosophila blood cells. *Science* (2011) 332:1210–3. doi: 10.1126/science.1199643
56. Shang Y, Smith S, Hu X. Role of notch signaling in regulating innate immunity and inflammation in health and disease. *Protein Cell* (2016) 7:159–74. doi: 10.1007/s13238-016-0250-0
57. Ferguson GB, Martinez-Agosto JA. Yorkie and scalloped signaling regulates notch-dependent lineage specification during drosophila hematopoiesis. *Curr Biol* (2014) 24:2665–72. doi: 10.1016/j.cub.2014.09.081
58. Lavine MD, Strand MR. Insect hemocytes and their role in immunity. *Insect Biochem Mol Biol* (2002) 32:1295–309. doi: 10.1016/S0965-1748(02)00092-9
59. Small C, Ramroop J, Otazo M, Huang LH, Saleque S, Govind S. An unexpected link between notch signaling and ROS in restricting the differentiation of hematopoietic progenitors in drosophila. *Genetics* (2014) 197:471–83. doi: 10.1534/genetics.113.159210
60. Janeway CA Jr., Medzhitov R. Innate immune recognition. *Annu Rev Immunol* (2002) 20:197–216. doi: 10.1146/annurev.immunol.20.083001.084359
61. Zhao ZY, Yin ZX, Xu XP, Weng SP, Rao XY, Dai ZX, et al. A novel c-type lectin from the shrimp *litopenaeus vannamei* possesses anti-white spot syndrome virus activity. *J Virol* (2009) 83:347. doi: 10.1128/JVI.00707-08
62. Liu Y, Liu J, Pang X, Liu T, Ning Z, Cheng G. The roles of direct recognition by animal lectins in antiviral immunity and viral pathogenesis. *Molecules* (2015) 20:2272–95. doi: 10.3390/molecules20022272
63. Kondo M, Matsuyama H, Yano T. The opsonic effect of lectin on phagocytosis by hemocytes of kuruma prawn, *penaeus japonicus*. *Fish Pathol* (1992) 27:217–22. doi: 10.3147/jsfp.27.217
64. Ma THT, Tiu SHK, He JG, Chan SM. Molecular cloning of a c-type lectin (LvLT) from the shrimp *litopenaeus vannamei*: early gene down-regulation after WSSV infection. *Fish Shellfish Immunol* (2007) 23:430–7. doi: 10.1016/j.fsi.2006.12.005
65. Yang H, Luo T, Li F, Li S, Xu X. Purification and characterisation of a calcium-independent lectin (PjLec) from the haemolymph of the shrimp *penaeus japonicus*. *Fish Shellfish Immunol* (2007) 22:88–97. doi: 10.1016/j.fsi.2006.03.015
66. Chen YH, Jia XT, Zhao L, Li CZ, Zhang S, Chen YG, et al. Identification and functional characterization of Dicer2 and five single VWC domain proteins of *litopenaeus vannamei*. *Dev Comp Immunol* (2011) 35:661–71. doi: 10.1016/j.dci.2011.01.010
67. Sheldon TJ, Miguel-Aliaga I, Gould AP, Taylor WR, Conklin D. A novel family of single VWC-domain proteins in invertebrates. *FEBS Lett* (2007) 581:5268–74. doi: 10.1016/j.febslet.2007.10.016
68. Gross PS, Bartlett TC, Browdy CL, Chapman RW, Warr GW. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the pacific white shrimp, *litopenaeus vannamei*, and the Atlantic white shrimp, *L. setiferus*. *Dev Comp Immunol* (2001) 25:565–77. doi: 10.1016/S0145-305X(01)00018-0
69. Stoss TD, Nickell MD, Hardin D, Derby CD, McClintock TS. Inducible transcript expressed by reactive epithelial cells at sites of olfactory sensory neuron proliferation. *J Neurobiol* (2004) 58:355–68. doi: 10.1002/neu.10294
70. Vargas-Albores F, Yepiz-Plascencia G, Jiménez-Vega F, Ávila-Villa A. Structural and functional differences of *litopenaeus vannamei* crustins. *Comp Biochem Physiol B-Biochem Mol Biol* (2004) 138:415–22. doi: 10.1016/j.cbpc.2004.05.007
71. Li M, Ma C, Zhu P, Yang Y, Lei A, Chen X. A new crustin is involved in the innate immune response of shrimp *litopenaeus vannamei*. *Fish Shellfish Immunol* (2019) 94:398–406. doi: 10.1016/j.fsi.2019.09.028
72. Zhang J, Li F, Wang Z, Xiang JH. Cloning and recombinant expression of a crustin-like gene from Chinese shrimp, *fenneropenaeus chinensis*. *J Biotechnol* (2007) 127:605–14. doi: 10.1016/j.jbiotec.2006.08.013
73. Koiwai K, Koyama T, Tsuda S, Toyoda A, Kikuchi K, Suzuki H, et al. Single-cell RNA-seq analysis reveals *penaeid* shrimp hemocyte subpopulations and cell differentiation process. *eLife* (2021) 10:e66954. doi: 10.7554/eLife.66954.sa2
74. Ahmed S, Stanley D, Kim Y. An insect prostaglandin E2 synthase acts in immunity and reproduction. *Front Physiol* (2018) 9:1231. doi: 10.3389/fphys.2018.01231
75. Kedia K, Wendler JP, Baker ES, Burnum-Johnson KE, Jarsberg LG, Stratton KG, et al. Application of multiplexed ion mobility spectrometry towards the identification of host protein signatures of treatment effect in pulmonary tuberculosis. *Tuberculosis* (2018) 112:52–61. doi: 10.1016/j.tube.2018.07.005
76. Leung LLK, Morser J. Carboxypeptidase B2 and carboxypeptidase n in the cross-talk between coagulation, thrombosis, inflammation, and innate immunity. *J Thromb Haemost* (2018) 16:1474–86. doi: 10.1111/jth.14199
77. Li D, Scherfer C, Korayem AM, Zhao Z, Schmidt O, Theopold U. Insect hemolymph clotting: evidence for interaction between the coagulation system and the prophenoloxidase activating cascade. *Insect Biochem Mol Biol* (2002) 32:919–28. doi: 10.1016/S0965-1748(02)00030-9
78. Ferguson SW, Nguyen J. Exosomes as therapeutics: the implications of molecular composition and exosomal heterogeneity. *J Control Release* (2016) 228:179–90. doi: 10.1016/j.jconrel.2016.02.037
79. Meldolesi J. Exosomes and ectosomes in intercellular communication. *Curr Biol* (2018) 28:R435–44. doi: 10.1016/j.cub.2018.01.059
80. Shibata T, Hadano J, Kawasaki D, Dong X, Kawabata SI. Drosophila TG-a transglutaminase is secreted via an unconventional golgi-independent mechanism involving exosomes and two types of fatty acylations. *J Biol Chem* (2017) 292:10723–1073. doi: 10.1074/jbc.M117.779710
81. Zeringer E, Barta T, Li M, Vlassov AV. Strategies for isolation of exosomes. *Cold Spring Harb Protoc* (2015) 4:319–23. doi: 10.1101/pdb.top074476
82. Jiravanichpaisal P, Lee BL, Söderhäll K. Cell-mediated immunity in arthropods: hematopoiesis, coagulation, melanization and opsonization. *Immunobiology* (2006) 211:213–36. doi: 10.1016/j.imbio.2005.10.015
83. Davey PA, Power AM, Santos R, Bertemes P, Ladurner P, Palmowski P, et al. Omics-based molecular analyses of adhesion by aquatic invertebrates. *Biol Rev* (2021) 96:1051–75. doi: 10.1111/brv.12691
84. Davey PA, Power AM, Santos R, Bertemes P, Ladurner P, Palmowski P, et al. Coagulation in arthropods: defence, wound closure and healing. *Trends Immunol* (2004) 25:289–94. doi: 10.1016/j.it.2004.03.004
85. Daimon E, Shibukawa Y, Wada Y. Calponin 3 regulates stress fiber formation in dermal fibroblasts during wound healing. *Arch Dermatol Res* (2013) 305:571–84. doi: 10.1007/s00403-013-1343-8
86. Kaestner L, Bogdanova A, Egee S. Calcium channels and calcium-regulated channels in human red blood cells. *Adv Exp Med Biol* (2020) 1131:625–48. doi: 10.1007/978-3-030-12457-1\_25
87. Wang XW, Wang JX. Diversity and multiple functions of lectins in shrimp immunity. *Dev Comp Immunol* (2013) 39:27–38. doi: 10.1016/j.dci.2012.04.009
88. Wang XW, Wang JX. Pattern recognition receptors acting in innate immune system of shrimp against pathogen infections. *Fish Shellfish Immunol* (2013) 34:981–9. doi: 10.1016/j.fsi.2012.08.008
89. Apel K, Hirt H. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* (2004) 55:373–99. doi: 10.1146/annurev.arplant.55.031903.141701
90. Yang Y, Bazhin AV, Werner J, Karakhanova S. Reactive oxygen species in the immune system. *Int Rev Immunol* (2013) 32:249–70. doi: 10.3109/08830185.2012.755176
91. Bazil JN, Beard DA, Vinnakota KC. Catalytic coupling of oxidative phosphorylation, ATP demand, and reactive oxygen species generation. *Biophys J* (2016) 110:962–71. doi: 10.1016/j.bpj.2015.09.036
92. Rosa RD, Barracco MA. Antimicrobial peptides in crustaceans. *ISJ-Invertebr. Surviv J* (2010) 7:262–84.
93. Cerenius L, Söderhäll K. The prophenoloxidase-activating system in invertebrates. *Immunol Rev* (2004) 198:116–26. doi: 10.1111/j.0105-2896.2004.00116.x
94. Kan H, Kim CH, Kwon HM, Park JW, Roh KB, Lee H, et al. Molecular control of phenoloxidase-induced melanin synthesis in an insect. *J Biol Chem* (2008) 283:25316–23. doi: 10.1074/jbc.M804364200
95. Tauszig S, Jouanguy E, Hoffmann JA, Imler JL. Toll-related receptors and the control of antimicrobial peptide expression in drosophila. *Proc Natl Acad Sci* (2000) 97:10520–5. doi: 10.1073/pnas.180130797
96. Lin X, Söderhäll I. Crustacean hematopoiesis and the astakine cytokines. *Blood* (2011) 117:6417–24. doi: 10.1182/blood-2010-11-320614
97. Baron M. An overview of the notch signalling pathway. *Semin Cell Dev Biol* (2003) 14:113–9. doi: 10.1016/S1084-9521(02)00179-9
98. Zhao W, Zheng Z, Aweya JJ, Wang F, Li S, Tuan TN, et al. *litopenaeus vannamei* notch interacts with COP9 signalosome complex subunit 1 (CNS1) to negatively regulate the NF- $\kappa$ B pathway. *J Proteomics* (2021) 232:104074. doi: 10.1016/j.jprot.2020.104074
99. Badia R, Pujantell M, Riveira-Muñoz E, Puig T, Torres-Torronteras J, Martí R, et al. The G1/S specific cyclin D2 is a regulator of HIV-1 restriction in non-proliferating cells. *PLoS Pathog* (2016) 12:e1005829. doi: 10.1371/journal.ppat.1005829
100. Li M, Dong Q, Zhu B. Aurora kinase b phosphorylates histone H3. 3 at serine 31 during mitosis. *Mamm Cells J Mol Biol* (2007) 429:2042–5. doi: 10.1016/j.jmb.2017.01.016
101. Murakami R, Okumura T, Uchiyama H. GATA factors as key regulatory molecules in the development of drosophila endoderm. *Dev Growth Differ* (2005) 47:581–9. doi: 10.1111/j.1440-169X.2005.00836.x
102. Lin X, Novotny M, Söderhäll K, Söderhäll I. Ancient cytokines, the role of astakines as hematopoietic growth factors. *J Biol Chem* (2010) 285:28577–86. doi: 10.1074/jbc.M110.138560

103. Heino TI, Kärpänen T, Wahlström G, Pulkkinen M, Eriksson U, Alitalo K, et al. The drosophila VEGF receptor homolog is expressed in hemocytes. *Mech Dev* (2001) 109:69–77. doi: 10.1016/S0925-4773(01)00510-X
104. Takemura M, Nakato H. Drosophila Sulfl is required for the termination of intestinal stem cell division during regeneration. *J Cell Sci* (2017) 130:332–43. doi: 10.1242/dev.149278
105. Moser M, Binder O, Wu Y, Aitseaomo J, Ren R, Bode C, et al. BMPER, a novel endothelial cell precursor-derived protein, antagonizes bone morphogenetic protein signaling and endothelial cell differentiation. *Mol Cell Biol* (2003) 23:5664. doi: 10.1128/MCB.23.16.5664-5679.2003
106. Tuddenham L, Wheeler G, Ntounia-Fousara S, Waters J, Hajihosseini MK, Clark I, et al. The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. *FEBS Lett* (2006) 580:4214–7. doi: 10.1016/j.febslet.2006.06.080
107. Lin X, Söderhäll K, Söderhäll I. Invertebrate hematopoiesis: an astakine-dependent novel hematopoietic factor. *J Immunol* (2011) 186:2073–9. doi: 10.4049/jimmunol.1001229
108. Wong LH, McGhie JD, Sim M, Anderson MA, Ahn S, Hannan RD, et al. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res* (2010) 20:351–60. doi: 10.1101/gr.101477.109
109. Rice SJ, Lai SC, Wood LW, Helsley KR, Runkle EA, Winslow MM, et al. MicroRNA-33a mediates the regulation of high mobility group at-hook 2 gene (*hmga2*) by thyroid transcription factor 1 (*ttf-1/nkx2-1*). *J Biol Chem* (2013) 288:16348–60. doi: 10.1074/jbc.M113.474643
110. Sallmyr A, Rashid I, Bhandari SK, Naila T, Tomkinson AE. Human DNA ligases in replication and repair. *DNA repair* (2020) 93:102908. doi: 10.1016/j.dnarep.2020.102908
111. Lele KM, Wolgemuth DJ. Distinct regions of the mouse cyclin A1 gene, *Ccna1*, confer male germ-cell specific expression and enhancer function. *Biol Reprod* (2004) 71:1340–7. doi: 10.1095/biolreprod.104.030387
112. Li J, Deng M, Wei Q, Liu T, Tong X, Ye X. Phosphorylation of MCM3 protein by cyclin *e*/cyclin-dependent kinase 2 (*Cdk2*) regulates its function in cell cycle. *J Biol Chem* (2011) 286:39776–85. doi: 10.1074/jbc.M111.226464
113. Uren AG, Wong L, Pakusch M, Fowler KJ, Burrows FJ, Vaux DL, et al. Survivin and the inner centromere protein INCENP show similar cell-cycle localization and gene knockout phenotype. *Curr Biol* (2000) 10:1319–28. doi: 10.1016/S0960-9822(00)00769-7
114. Sun J, Wang A, Zhang T. Flow cytometric analysis of defense functions of hemocytes from the penaeid shrimp, *penaeus vannamei*. *J World Aquacult Soc* (2010) 41:92–105. doi: 10.1111/j.1749-7345.2009.00316.x
115. Martin GG, Graves BL. Fine structure and classification of shrimp hemocytes. *J Morphol* (1985) 185:339–48. doi: 10.1002/jmor.1051850306
116. Cheng TC. *Invertebrate blood: Cells and serum factors* Vol. 6. Berlin: Springer Science & Business Media (2013).
117. Gargioni R, Barracco MA. Hemocytes of the palaemonids macrobrachium rosenbergii and *m. acanthurus*, and of the penaeid *penaeus paulensis*. *J Morphol* (1998) 236:209–21. doi: 10.1002/(SICI)1097-4687(199806)236:3<209::AID-JMOR4>3.0.CO;2-Y
118. Tsing A, Arcier JM, Brehélin M. Hemocytes of penaeid and palaemonid shrimps: morphology, cytochemistry, and hemograms. *J Invertebr Pathol* (1989) 53:64–77. doi: 10.1016/0022-2011(89)90075-X
119. Hose JE, Martin GG, Gerard AS. A decapod hemocyte classification scheme integrating morphology, cytochemistry, and function. *Biol Bull* (1990) 178:33–45. doi: 10.2307/1541535
120. Giulianini PG, Bierti M, Lorenzon S, Battistella S, Ferrero EA. Ultrastructural and functional characterization of circulating hemocytes from the freshwater crayfish *astacus leptodactylus*: cell types and their role after *in vivo* artificial non-self challenge. *Micron* (2007) 38:49–57. doi: 10.1016/j.micron.2006.03.019
121. Söderhäll I, Junkunlo K. A comparative global proteomic analysis of the hematopoietic lineages in the crustacean *pacifastacus leniusculus*. *Dev Comp Immunol* (2019) 92:170–8. doi: 10.1016/j.dci.2018.11.016
122. Sun M, Li S, Zhang X, Xiang J, Li F. Isolation and transcriptome analysis of three subpopulations of shrimp hemocytes reveals the underlying mechanism of their immune functions. *Dev Comp Immunol* (2020) 108:103689. doi: 10.1016/j.dci.2020.103689
123. Cerenius L, Söderhäll K. Crayfish immunity—recent findings. *Dev Comp Immunol* (2018) 80:94–8. doi: 10.1016/j.dci.2017.05.010
124. Li YD, Zhou FL, Yang QB, Jiang S, Huang JH, Yang LS, et al. Single-cell sequencing reveals types of hepatopancreatic cells and haemocytes in black tiger shrimp (*Penaeus monodon*) and their molecular responses to ammonia stress. *Front Immunol* (2022) 13:a. doi: 10.3389/fimmu.2022.883043
125. Yang P, Chen YH, Huang ZQ, Xia HD, Cheng L, Wu H, et al. Myeloid cell evolution uncovered by shrimp immune cell analysis at single-cell resolution. *bioRxiv* (2022), 492277. doi: 10.1101/2022.05.17.492277
126. Cho B, Yoon SH, Lee D, Koranteng F, Tattikota SG, Cha N, et al. Single-cell transcriptome maps of myeloid blood cell lineages in drosophila. *Nat Commun* (2020) 11(1):1–18. doi: 10.1038/s41467-020-18135-y
127. Kwon H, Mohammed M, Franzén O, Ankarkev J, Smith RC. Single-cell analysis of mosquito hemocytes identifies signatures of immune cell subtypes and cell differentiation. *Elife* (2021) 10:e66192. doi: 10.7554/eLife.66192.sa2
128. Li F, Chang X, Xu L, Yang F. Different roles of crayfish hemocytes in the uptake of foreign particles. *Fish Shellfish Immunol* (2018) 77:112–9. doi: 10.1016/j.fsi.2018.03.029
129. Persson M, Vey A, Söderhäll K. Encapsulation of foreign particles *in vitro* by separated blood cells from crayfish, *astacus leptodactylus*. *Cell Tissue Res* (1987) 247:409–15. doi: 10.1007/BF00218322
130. Duvic B, Hoffmann JA, Meister M, Royet J. Notch signaling controls lineage specification during drosophila larval hematopoiesis. *Curr Biol* (2002) 12:1923–7. doi: 10.1016/S0960-9822(02)01297-6
131. Ferguson GB, Martinez-Agosto JA. Kicking it up a notch for the best in show: Scalloped leads yorkie into the haematopoietic arena. *Fly* (2014) 8:206–17. doi: 10.1080/19336934.2015.1055427
132. Jariyapong P, Pudgerd A, Cheloh N, Hirono I, Kondo H, Vanichviriyakit R, et al. Hematopoietic tissue of macrobrachium rosenbergii plays dual roles as a source of hemocyte hematopoiesis and as a defensive mechanism against macrobrachium rosenbergii nodavirus infection. *Fish Shellfish Immunol* (2009) 86:756–63. doi: 10.1016/j.fsi.2018.12.021
133. Hammond JA, Smith VJ. Lipopolysaccharide induces DNA-synthesis in a sub-population of hemocytes from the swimming crab, *liocarcinus depurator*. *Dev Comp Immunol* (2002) 26:227–36. doi: 10.1016/S0145-305X(01)00069-6
134. Buttitta LA, Edgar BA. Mechanisms controlling cell cycle exit upon terminal differentiation. *Curr Opin Cell Biol* (2007) 19:697–704. doi: 10.1016/j.ceb.2007.10.004
135. Soufi A, Dalton S. Cycling through developmental decisions: how cell cycle dynamics control pluripotency, differentiation and reprogramming. *Development* (2016) 143:4301–11. doi: 10.1242/dev.142075
136. Gisselbrecht S, Skeath JB, Doe CQ, Michelson AM. Heartless encodes a fibroblast growth factor receptor (DFR1/DFGF-R2) involved in the directional migration of early mesodermal cells in the drosophila embryo. *Genes Dev* (1996) 10:3003–17. doi: 10.1101/gad.10.23.3003