Ethics statement

Not applicable as the bivalve species used in this study (*Crassostrea gigas*) does not require ethical permits. No permits were required for the collection of oysters from the hatchery.

Not applicable as oysters do not require ethical permits. No permits were required for the collection of oysters from the hatchery.

Other statements to use as examples:

The study plan and animal care procedures were approved by the NOAA Fisheries Northwest Fisheries Science Center. NOAA Fisheries does not require or provide for review by an Institutional Animal Care and Use Committee for research projects focused solely on fishes. The permit used to sample sablefish was a Scientific Research Permit issued through the NOAA-West coast Region Sustainable Fishery Division. In the present study, fish were euthanized in accordance with guidelines for the euthanasia of finfish of the American Veterinary Medical Association [27].

All the experiments were conducted according to the ethics committee of Ifremer. Oyster genitors were sampling in field and sites for sampling were accessible on foot or by boat. No permit was required to harvest oysters. A very small fraction of the oysters (< 0.1% in every site) was collected without disturbance. Except for the production of biparental families of oysters, the study did not used samples collected from the field.

Another one just says NA

Another one say: Not applicable as oysters do not require ethical permits. No permits were required for the collection of oysters from the hatchery.

The study was approved by the ethics committee of the French institute of research and exploitation of the sea (Ifremer). The breeding of *P. margaritifera* followed institutional and national guidelines.

PLEASE DO NOT MAKE ADDITIONAL CHANGES, WORKING IN WORD VERSION AS OF 8/8/25

Target Journal: BMC Genomics

Primordial Germ Cell Specification and Early Developmental Cell States in Pacific Oyster

Authors: Mackenzie R. Gavery*, Lauren Vandepas*, Lauren M. Saunders, Brent Vadopalas, J. Adam Luckenbach, Cole Trapnell, Steven Roberts

*contributed equally

Affiliations

Adam and Mac: Environmental and Fisheries Sciences Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd E, Seattle, WA 98112, USA

Adam: Center for Reproductive Biology, Washington State University, Pullman, WA 99164, USA

Lauren V. and Steven and Brent: School of Aquatic & Fishery Sciences, University of Washington, Seattle WA, USA

Lauren S.: Centre for Organismal Studies, Heidelberg University, Heidelberg, Germany

Cole: Department of Genome Sciences, University of Washington, Seattle WA, USA

ABSTRACT

Background: Primordial germ cells (PGCs) are the precursor cells of gametes and pivotal in understanding reproductive and developmental biology. Importantly, having a thorough understanding of PGC specification is leading to critical advances in sterility induction in aquaculture species. In shellfish, however, the ability to develop these approaches is hampered

by the lack of information available regarding germ cell specification. The goal of this study was to identify genes uniquely expressed in these earliest germ cells of the economically and ecologically important bivalve mollusc, the Pacific oyster (*Crassostrea* (*Magallana*) *gigas*).

Results: To capture specification of the PGCs - which represent a rare cell type - during embryonic development, we measured single-cell transcriptomes during cleavage, blastula, and gastrulation stages of *C. gigas* development. Using expression of germ cell marker gene, *vasa*, we identified cells in both cleavage-stage embryos and blastulae that likely represent the developing germ cell lineage, but had yet to fully differentiate and segregate from somatic cell types. However, at the gastrula stage, *vasa* expression was limited primarily to a single cluster of cells. Other genes uniquely expressed in these *vasa*-positive cells include those with functions in transcriptional repression, chromatin architecture, and DNA repair. Interestingly, some genes with no known homologies are also uniquely expressed in this cluster, perhaps representing novel PGC-associated genes in bivalves.

Conclusions: We identified a suite of candidate genes that can be explored for their role in oyster PGC specification and advance efforts to develop methods to achieve reproductive sterility via germ cell disruption in cultured shellfish. In addition, this effort produced a transcriptional atlas of early developmental cell states in bivalve embryos, providing a wealth of information on genes contributing to other important developmental processes, such as growth and shell production. These are the earliest developmental stages examined via single-cell RNA sequencing in a lophotrochozoan.

BACKGROUND

Primordial germ cells (PGCs) are specialized cells that give rise to self-renewing germinal stem cells (GSCs; germline), which subsequently differentiate into gametes in the gonad. Because disruption of germline development can result in sterility, perturbation of PGC development is being explored as a means to produce sterile animals for use in aquaculture (Wargelius et al. 2016, Kleppe et al. 2022, Xu et al. 2023a, Xu et al. 2023b). Sterile or non-reproductive shellfish are both a market-driven need and an ecologically sustainable approach to increasing food production via aquaculture. Sterility has clear advantages in shellfish aquaculture including the ability to improve rates of growth and meat quality, prevent accidental establishment of non-native species into the environment, and preclude genetic contamination of wild, native bivalve populations by farmed conspecifics (Nell 2002, Piferrer et al. 2009).

The Pacific oyster *Crassostrea* (*Magallana*) *gigas* (Thunberg 1793; but see Salvi et al. 2014, Bayne et al. 2017) is one of the world's most economically important bivalves, approaching a million metric tons farmed annually (FAO 2020). *Crassostrea* sp. are also among the best characterized molluscan model systems for evolution and development (evo-devo), genetics, and cell biology (reviewed in Robledo et al. 2019) and provide important ecosystem services (Grabowski & Peterson 2007). Efforts to generate sterile oysters for aquaculture to date have involved creation of triploid (3n) individuals (Downing & Allen 1987, Guo et al. 2009, Nell 2002). The shellfish aquaculture industry produces triploid oysters due to their reduced gonad development, year-round marketability, and improved performance traits when farmed (Allen &

Downing, 1986, 1991; Dégremont et al., 2012; Normand et al., 2008; Shpigel et al., 1992). Although use of triploids in aquaculture has been widely adopted, increasing reports of triploid mortality under environmentally stressful conditions (Houssin et al. 2019, Wadsworth 2019, Guevelou et al. 2019) highlights the critical need for alternative methods to induce sterility in farmed oysters.

A potential alternative to ploidy manipulation in bivalves is the induction of sterility via perturbation of genes essential for PGC formation. The power of this biotechnological approach has been realized recently in finfish species, where suppression of the germ-cell specific gene, *dead end* (*dnd*), produced zebrafish and Atlantic salmon with no detectable germ cells or gonadal development (Wong and Zohar, 2015, Wargelius et al., 2016). A thorough understanding of genes involved in germ cell fate is a critical step toward controlling reproduction through molecular approaches such as these. However, beyond descriptions of expression of the germline marker gene *vasa* (Fabioux et al., 2004a,b; Fabioux et al. 2009; Xu et al. 2025), PGCs are poorly characterized in bivalves. Further, known marker genes specific to vertebrate PGCs and germline such as *dnd* (Weidinger et al. 2003) are not present in molluscs, and identification of genes involved in oyster PGC specification has been elusive due to a lack of targeted sequencing efforts with sufficient cell-level resolution, as germ cell precursors make up only a tiny fraction of the cells in developing oyster embryos.

Single-cell RNA-seq (scRNA-seq) has emerged as a powerful technique to identify and molecularly characterize rare cell populations in developing organisms, including other marine invertebrates (Foster et al. 2020; Massri et al. 2021; Salamanca-Diaz et al. 2022; Piovani et al. 2023). Further, single-cell sequencing of early embryogenesis in non-vertebrates can reveal transcriptional signatures associated with early specification and differentiation of cell types, including of PGCs (Packer et al. 2019, Foster et al. 2020, Sladitschek et al. 2020). Here, we apply scRNA-seq to profile *C. gigas* early cleavage stages, blastulae, and gastrulae to comprehensively identify and characterize single-cell gene expression in whole developing animals with a primary focus on PGC specification. We show that PGC sequestration occurs prior to gastrulation and identify novel PGC-specific genes that will be used as candidates for future gene knockdown/silencing experiments to develop sterile shellfish for aquaculture efforts. We present the earliest single cell developmental time series atlas for a lophotrochozoan, data which provide novel insights into formation of a variety of larval tissues in the model bivalve *C. gigas*.

METHODS

Experimental design and *C. gigas* sampling

Gravid adult *C. gigas* were generously provided by Taylor Shellfish Hatchery (Quilcene, WA, USA). All steps were conducted at room temperature (approximately 20°C). Male or female gonad tissues were submerged in filtered seawater (FSW) (0.2 µm filter) and gametes gently removed by agitating the tissue using a lab spatula. Isolated oocytes were passed through a 90 µm filter, then collected with a 20 µm filter to remove debris. Oocytes were then allowed to hydrate for 1 h in FSW. Oocytes were visually assessed with a microscope to ensure complete hydration prior to fertilization. Spermatozoa were separated from the gonad of a male oyster using a lab spatula in approximately 20 mL of FSW. Sperm were observed with a microscope to

qualitatively confirm motility. A small volume of sperm suspension was mixed with eggs to initiate fertilization (at 1:200 sperm solution to eggs). Mixed gametes were incubated for 10 min at room temperature. Fertilized oocytes were again screened on a 20 µM filter and rinsed with FSW to remove excess spermatozoa, and subsequently maintained in FSW during development.

To measure genes uniquely expressed in *C. gigas* PGCs and/or PGC precursors, we collected embryos at the earliest reported developmental stages for the expression of *vasa* (Fabioux et al. 2004). We collected early embryos from 8 time-points representing a continuum of development from cleavage- to blastula-stage embryos (Figure S1). Development in oysters can be asynchronous and samples were visually assessed to ensure that the majority (> 50%) of embryos were at similar developmental stages. To collect cleavage-stage embryos, samples were taken after each cell division (~35 min) and placed on ice to arrest further development. Embryos were counted and 10,000 embryos for each of the first four cleavage samples, and 100,000 embryos for the remaining four cleavage samples (to account for the smaller proportion of cells that may represent the putative germ cells in the later cleavage-stage embryos) were pooled for cell dissociation and library preparation. To collect blastula-stage embryos, a sample was taken approximately 35 min after the last cleavage stage (approximately 6.5 hours-post fertilization (hpf)). Approximately 300,000 blastula-stage embryos were collected in FSW for cell dissociation and library preparation. To collect gastrulae, embryos were sampled 10 hpf (Strathmann 1987). Approximately 40,000 gastrulae were collected in FSW for cell dissociation and library preparation.

Cell dissociation and scRNA-seq library preparation

Gastrula-stage embryos were rinsed three times, while gradually decreasing the salinity: first in FSW, next in 0.75x calcium- and magnesium-free artificial seawater (ASW), then in 0.5x ASW, and then finally in 0.3x ASW. Embryos were dissociated to single-cell suspensions with 500 µL of 10 mg/mL protease (Native *Bacillus licheniformis* protease; Creative Enzymes, Catalog Number: NATE0633) in 0.3x ASW incubated at room temperature for 15 minutes with intermittent gentle pipetting with a P1000 pipette. Cells were visualized on a microscope to monitor dissociation efficiency. The dissociation reaction was halted by transferring cells to ice-cold Dulbecco's phosphate-buffered saline (dPBS, ThermoFisher; Catalog Number: 14190144, pH 7.0-7.3) containing 10% fetal bovine serum (FBS; Thermo Fisher; Catalog Number A4736401). The dissociated cells were filtered through a 30 µm filter into a 15 mL tube, centrifuged (600 rcf, 5 min) and resuspended in 5 mL 0.3x ASW. This rinse was repeated, followed by resuspension in 200 µL 0.3x ASW. Cells were counted and viability assessed using 0.2% Trypan blue (HyClone; Catalog Number: SV30084) on a Bright-Line Hemacytometer (Hausser Scientific, Catalog Number: 3110). To process cleavage-stage and blastula samples, cell dissociation was performed similarly with the exception that embryos were rinsed once in room temperature FSW and then resuspended directly in dPBS for subsequent dissociation, filtering, and rinsing steps.

Cells were prepared for sequencing using the 10X Genomics Chromium platform (Zheng et al. 2017). Single-cell mRNA libraries were prepared using the Chromium Next GEM Single Cell 3’ GEM, Library & Gel Bead Kit v3.1 (10x Genomics, Pleasanton, CA, USA). Approximately 24,000 cells (split equally across 4 reactions) were targeted for the gastrula sample,

approximately 18,000 cells (split equally across 3 reactions) were targeted for capture for the cleavage pool and 6,000 cells were targeted for the blastula sample. Quality control and quantification assays were performed using a Qubit fluorometer (Thermo Fisher, Catalog Number::Q33238) and a D1000 Screentape Assay (Agilent, Catalog Number:5067-5582). Libraries were sequenced on two Illumina NextSeq 500 75-cycle, high output kits (v2.5; Illumina, Catalog Number: 20024906).

Processing of scRNA-seq sequencing reads

Sequencing reads were processed using the Cell Ranger pipeline (v3.1.0, 10X Genomics). A custom *C. gigas* STAR genome index was built using gene annotations from the "cggigas_uk_roslin_v1" version of the genome (GenBank: GCA_902806645.1) and filtered for protein-coding genes using the *mkref* command. Cell barcodes and unique molecular identifiers (UMIs) were determined using the Cell Ranger *count* command. To ensure accurate calling of cells, a more stringent filter requiring 2000 UMIs per cell was applied to cell calling from Cell Ranger (Fig S2).GitHub repo with code: [this is coming]

In total, approximately 8 x 10^8 GB of sequencing data were generated across the cleavage-stage, blastulae and gastrulae libraries (NCBI SRA: PRJNA906172). Post-filtering, a total of 18,511 cells (9,314 median UMIs and a median 2,828 genes per cell) were analyzed for the gastrulae, 7,658 cells (4,056 median UMIs and 1,949 median genes per cell) for the cleavage-stage embryos, and 2,836 cells (5,232 median UMIs and 2,277 median genes per cell) for the blastulae embryos. Mapping was consistent across libraries with 80-85% of reads mapping to the *C. gigas* genome. See Table S1 for additional sequencing and mapping information. Data for cleavage-stage and blastula cells were combined for downstream analysis.

Dimensionality reduction, clustering and cell-type annotation

After cell calling and quality control, single cell data analysis was performed using Monocle3 (Cao et al., Nature, 2019) (v1.3.4), following a standard processing pipeline: (1) log normalization of gene counts and PCA analysis with retention of the 20 top principal components for (2) projection of data into two dimensions with Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018) (*umap.min_dist = 0.2, umap.n_neighbors = 15L*) and (3) Leiden clustering to group cells. UMAP visualization and clustering was performed separately for the gastrulae, and combined for the cleavage-stage and blastula embryos. The function *align_cds* (Haghverdi et al. 2018) was used to minimize batch effects of replicate libraries in low dimensional space. All genes were provided as input into Principal Component Analysis (PCA). For cell clustering, we manually adjusted the resolution parameter towards modest overclustering.

Genes that were highly expressed and specific to a particular cluster (referred to here as 'marker genes') were identified in order to facilitate the annotation of cell clusters to a particular cell type. Specifically, the function *top_markers (group_cells_by = "cluster"*, *reference_cells = 1000* (Monocle3 v.1.3.4 (Trapnell et al. 2014, Cao et al. 2019)) was used to identify the genes most specifically expressed in each cluster. The 25 most specifically expressed genes by "marker_score" (a value between 0-1 based on the fraction of cells expressing the gene scaled by a measure of how specific the gene's expression is to the cluster) are reported in Table 2.

Annotation of *C. gigas* genes to homologs in model invertebrates

In order to facilitate direct comparisons of *C. gigas* marker genes identified here to those in other model marine invertebrates, an iterative BLAST was performed to *Caenorhabditis elegans, Drosophila melanogaster* and *Strongylocentrotus purpuratus* databases (the top BLAST hit and corresponding e-value reported in Tables S2-S4). Specifically, the full set of *C. gigas* genes (GCF_902806645.1_cgigas_uk_roslin_v1_translated_cds.faa) was compared to SwissProt database (20210613_ncbi_sp_v5/swissprot (blastp)), *C. elegans* protein database (Caenorhabditis_elegans.WBcel235.pep (blastp)), *D. melanogaster* nucleotide database (dmel-all-CDS-r6.37 (tblastn)) and the *S. purpuratus* protein database (ProteinsSpur5.0 (blastp)) using NCBI BLAST (Altschul et al. 1990). Marker genes used to assign cell or tissue identities to UMAP clusters were manually screened using iterative reciprocal protein BLAST to assess predicted gene homology.

Semi-quantitative PCR

RNA was extracted from pooled embryos or larvae and whole tissue from juvenile *C. gigas* using Tri Reagent (ThermoFisher, Catalog Number: 15596018), following the manufacturer's protocol. The amount of cDNA was normalized so that 200 ng of cDNA were loaded into each PCR reaction. PCR amplifications were performed in a 25 mL reaction volume consisting of 9.5 µL nuclease-free water, 12.5 µL GoTaq master mix (Promega, Catalog Number: M7122), 1 µL 10 µM forward primer, 1 µL 10 µM reverse primer, 1 µL cDNA template. Amplification was performed on a T100 Thermal Cycler (Bio-Rad): initial denaturation, 2 m at 95°C, followed by 35 cycles of 30 sec at 95°C for denaturation, 30 sec at 60°C annealing, 60 s at 72°C for extension, and final extension at 72°C for 10 min. The PCR products were then assessed by gel electrophoresis at 100 V for 30 min in a 1.25% (w/v) agarose gel in 1 x TAE buffer using SYBR Safe (Thermo Fisher, Catalog Number: S33102).

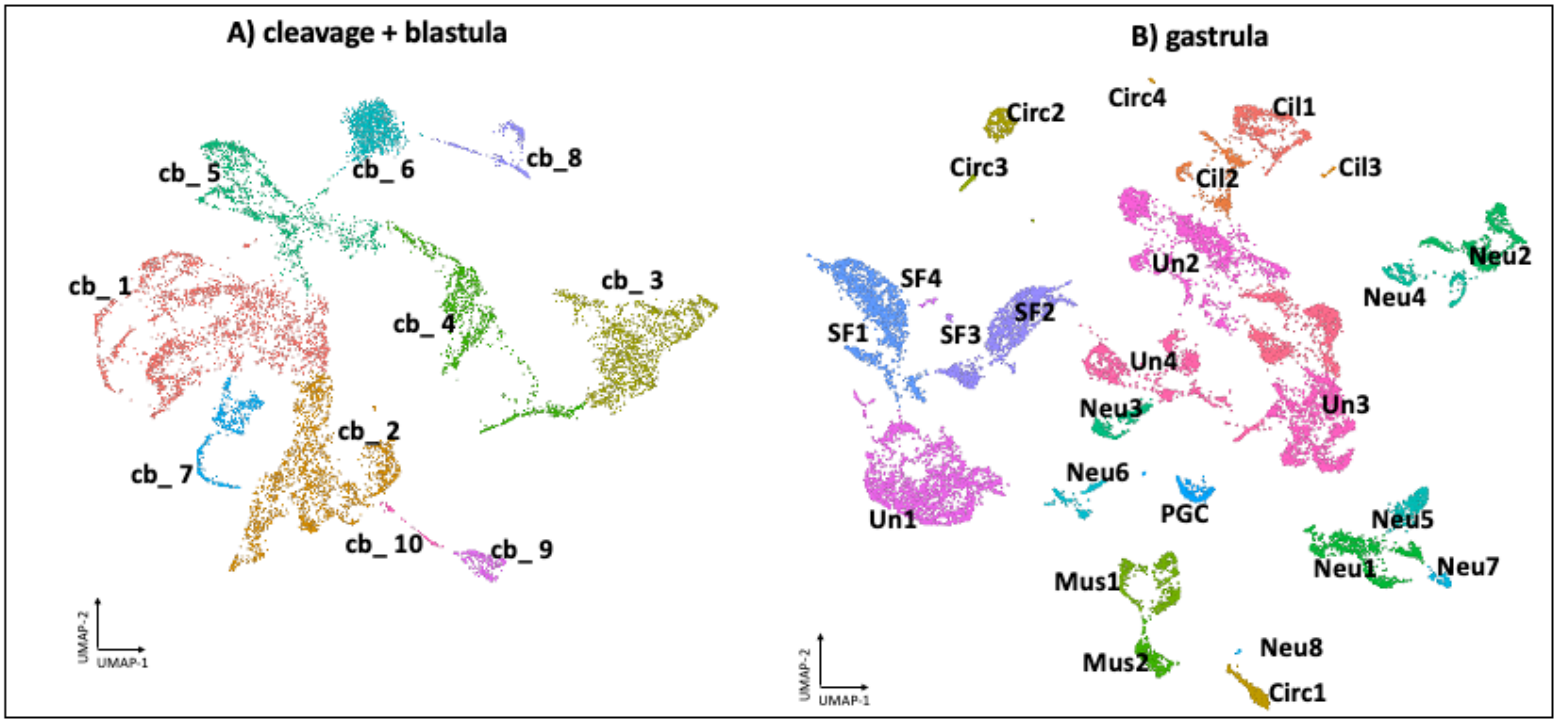
GAPDH was used as a housekeeping control gene and expression of PGC-related genes *nanos, vasa, sperm-specific protein PHI-2B/PHI-3* (*spPHI*)*, uncharacterized gene LOC105328839* were assessed using semi-quantitative PCR (Table S5). Two biological replicates (either pooled larvae from different parent crosses or individual juvenile oysters) and at least two technical replicates were performed for each timepoint for each gene. Gel band quantification was performed using FIJI software (Schindelin et al. 2012).

RESULTS

To identify potential progenitors of bivalve PGCs, we used single-cell transcriptomics to analyze gene expression profiles in *C. gigas* early developmental stages: cleavage-stages + blastulae and gastrulae (Fig. 1; Fig. S1). A total of 7,658 and 2,836 cells for the cleavage-stage and blastulae embryos respectively and 18,511 cells were analyzed for the gastrulae. Due to maternal mRNAs and undifferentiated cell states (Andeol 1994, Tadros and Lipschitz 2009), sequencing data from the earliest developmental stages (cleavage + blastula embryos) yielded few distinct cell clusters (Fig 1A). In contrast, gastrula stage embryos showed a strikingly high degree of heterogeneity (Fig 1B; Table S2). From this stage, we identified early genes associated with tissue progenitors such as PGCs, shell field, muscle, and neurons (Table 2).

Figure 1. Single-cell transcriptional heterogeneity of early *C. gigas* embryos increases with developmental time. (A) UMAP visualization of single-cell clustering of the cleavage (pre-blastula) +

blastula stage that revealed 10 putative clusters. (B) UMAP visualization of single-cell clustering of the gastrula stage. Twenty-seven putative clusters representing broad cell types were resolved using UMAP clustering analysis (Cil, ciliated cells; Circ, circulatory system-related; Mes, undefined mesodermal derivatives; Mus, muscle-related; Neu, neuronal; PGC, primordial germ cell; SF, shell field; Un, undefined).



1. Early developmental origins of PGCs in the oyster *C. gigas*

The germline is sequestered from somatic cells early in embryonic development in most animals (reviewed in Extavour 2007) and typically has low amounts of transcriptional activity to protect germ cells from introduction of mutations (Nakamura et al. 2010). However, PGCs do express a specific suite of conserved genes including *vasa*, a deadbox RNA helicase that is a PGC marker in many metazoans (Gustafson and Wessel 2010; Colonnetta et al. 2022; Miramón-Puértolas et al. 2024). Previous studies in *C. gigas* have demonstrated that *vasa* is highly expressed in PGCs and is necessary for normal PGC development (Fabioux et al. 2004; Fabioux et al. 2009). To identify PGCs in developing *C. gigas* embryos and determine co-expressed genes that may be important for PGC specification, we analyzed single cell gene expression profiles of gastrulae and blastulae and cleavage-stage embryos (Figure 2, Table1, Fig S3).

1a. *PGCs are transcriptionally distinct by the gastrula stage in the oyster* C. gigas

We first searched our datasets for potential PGCs using the previously characterized oyster germline marker gene *vasa* (Fabioux et al. 2004). In our *C. gigas* gastrulae dataset, one cell cluster showed high expression of *vasa* and also had the highest proportion of *vasa*-expressing cells within the cluster (Figure 2B). Although *vasa* expression was relatively specific to this cluster at this developmental stage (one of the top 25 marker genes for this

cluster), a number of other genes were expressed with even higher specificity than *vasa*. One of those genes is *nanos*, an evolutionarily conserved repressor of transcription shown to be expressed exclusively in the oyster germline (Forbes and Lehmann 1998; Xu et al. 2018), which provided further support that the cells in this cluster represent developing PGCs (Fig. 2B; Table 1). The high and specific expression of both *vasa* and *nanos* indicates that a single cluster of cells identified here represents the PGCs in oyster gastrulae (Fig 2).

This cluster (which we will refer to as PGCs) also specifically expressed genes involved in transcriptional silencing, stemness, germline and gamete development, and chromatin remodeling and maintenance (Fig. 2A; Table 1). For example, the gene with the highest overall expression in the PGC cluster is a protamine-like histone variant, annotated as 'sperm-specific protein PHI-2B/PHI-3' (*spPHI*) (Fig 2; Table 1). An orthologous gene has been described in the bivalve *Mytilus californianus*, where it is uniquely expressed in spermatozoa (Carlos et al. 1993). It is possible that the high expression of *spPHI* observed in the PGC cluster is associated with maintaining transcriptional quiescence, a feature associated with early germ cell formation (Lebedeva et al. 2018).

Many genes uniquely expressed in the oyster gastrula PGC cluster have previously been associated with germline development and maintenance in other animals, a pattern that has been described previously (Forbes and Lehmann 1998). The gene fem-1 homolog A-A (*fem-1*), a stem-loop binding protein involved in processing translation and degradation of replication dependent histone mRNAs, is also uniquely expressed in the PGC cluster (Fig 2A; Table 1). First identified in *Caenorhabditis elegans*, *fem-1* plays a vital role in sex determination in diverse taxa, including bivalves (Doniach & Hodgkin, 1984; Galindo-Torres et al. 2019; Teaniniuraitemoana et al. 2014). Also specifically expressed in the oyster gastrula PGC cluster is structural maintenance of chromatin protein 6 (*smc6*) (Fig 2A; Table 1), which has been shown to play a role in maintaining germline integrity in *C. elegans* (Bickel et al. 2010). Similarly, ATRX chromatin remodeler (*ATRX*) is expressed highly in female gonad tissue in the oyster *Crassostrea hongkongensis* (Tong et al. 2015). These data taken together support the identification of a specific cluster of cells in our single-cell sequencing analysis that likely represents PGCs in *C. gigas* gastrulae.

Interestingly, there are also genes uniquely expressed in the PGC cluster with unknown germline functions (Fig 2A; Table 1). These include an uncharacterized *C. gigas* gene "*LOC105328839*," (Fig 2A-B; Table 1) a mollusc-specific protein that exhibits some amino acid sequence homology to mucins. Marker gene analysis also identified *mucin 5-AC* as being specifically expressed in this cluster (Table 1). Mucins play a diverse role in bivalves including locomotion, digestion, wound healing, biomineralization and defense (Prezant 1990), but little is known about the role of mucins in bivalve embryonic development.

1b. *Some genes associated with PGCs are expressed prior to gastrulation*

Cells from the earliest cleavage stages (2- to 512-cell) through the blastulae were analyzed together to identify genes uniquely expressed in the earliest germ cells and to potentially track PGC formation in *C. gigas* (Fig 1A, Fig2B, Fig S1). Genes associated with PGC development were expressed at these early developmental stages. Early embryos display wide expression of *vasa*; however, some clusters show relatively elevated levels of *vasa* expression (Fig 2B). Many marker genes identified as being uniquely expressed in the gastrula PGC cluster

were either not detected or showed broad expression in the cleavage-blastula cells (Fig S3). Interestingly, *spPHI*, which is specific to PGCs in oyster gastrula (Table 1), was co-expressed in *vasa*-expressing cells in these early embryos (Fig 2A).

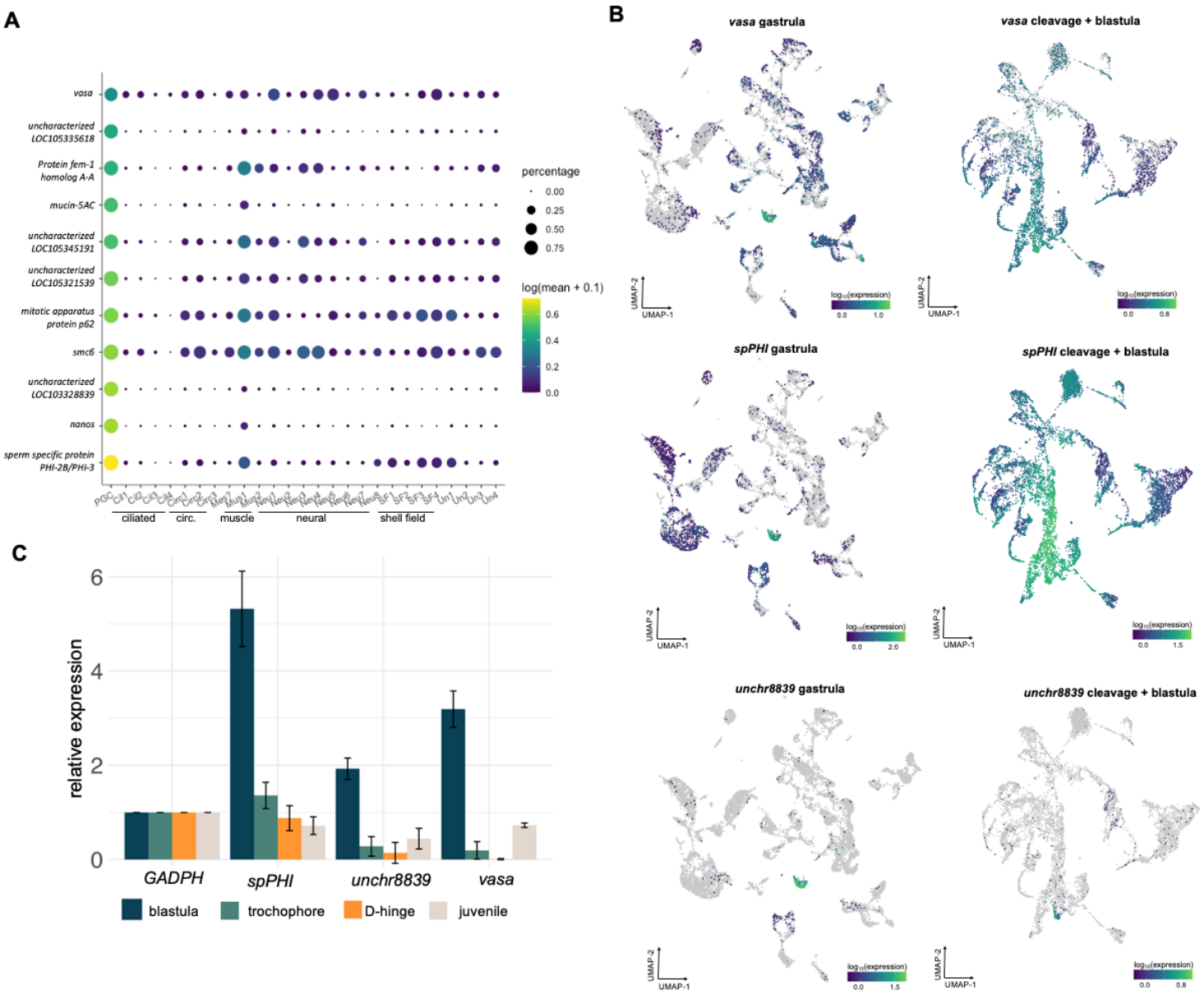
To temporally assess expression of some PGC marker genes identified through our sequencing analysis, we performed semi-quantitative PCR on pooled blastulae, trochophores, and early D-hinge larvae, as well as individual 8-month-old juveniles. Expression of *spPHI*, uncharacterized gene *LOC105328839*, and *vasa* is relatively high in blastula but decreases significantly as larval development progresses (Fig S4), with *spPHI* showing the most substantial drop in relative expression between blastula and trochophore stages. These data suggest that the genes *spPHI* and uncharacterized gene *LOC105328839* in particular are preferentially expressed early in development, or are restricted to a small number of cells later in development, supporting a potential role in initial PGC formation.

Table 1. Top 10 marker genes for the PGC cluster in *C. gigas* gastrulae

Gene ID	Gene Name	Marker Score	Swiss Prot ID	Conserved Domains	Protein IDs
LOC105328839	uncharacterized LOC105328839	0.73	N/A	N/A	XP_034312373.1
LOC105347430	protein nanos	0.70	P60321	zinc finger - Nanos superfamily (cl05351)	XP_011454832.3
LOC105327445	sperm-specific protein PHI-2B/PHI-3	0.63	Q3HNG7	linker histone 1/5 domain (cl00073)	XP_011426236.2
LOC105318512	mucin-5AC	0.63	N/A	N/A	XP_011413998.2
LOC105335618	uncharacterized LOC105335618	0.60	N/A	N/A	XP_034338648.1
LOC105321539	uncharacterized LOC105321539	0.51	N/A	N/A	XP_011418154.2
LOC105345191	uncharacterized LOC105345191	0.36	Q93075	metallo-dependent_hydrolases (cl00281)	XP_011451571.2
LOC105321610	structural maintenance of chromosomes protein 6	0.35	Q6P9I7	RecF/RecN/SMC N terminal domain (cl37666), Smc (cl34174)	XP_034314486.1
LOC105346676	mitotic apparatus protein p62	0.34	P91753	Nucleoplasmin-like domain (cl03870)	XP_034328868.1
LOC109619157	protein fem-1 homolog A-A	0.32	Q7T3P8	Ankyrin repeats (cl39094)	XP_034314173.1

Marker Score = The fraction of cells expressing the gene scaled by specificity.

Figure 2. Tracing the expression of *vasa* identifies a single putative PGC cluster in *C. gigas* gastrulae. (A) Dot plot heat map displaying the expression of marker genes identified in the PGC cluster in gastrulae across all clusters. Dot color represents expression level and dot size represents fractional representation of cells expressing a gene. (B) UMAP visualization of expression of select PGC marker genes in gastrulae (left) and cleavage + blastula embryos (right): *vasa* expression is primarily limited to a single cluster in the gastrula stage whereas expression is broader in the earlier cleavage+blastula stages (top); *spPHI* shows high specificity to the PGC cluster in gastrulae, while expression is widespread in cleavage + blastulae embryos (middle); an uncharacterized oyster-specific gene (*LOC105328839*) shows high levels of expression specifically in PGC cluster in gastrulae, but is only expressed in a small number of cells in cleavage+blastula embryos (bottom).



2. Single-cell transcriptomics highlights tissue-specific progenitors in *C. gigas* gastrulae

Morphogenetic events and cell signaling during gastrulation define germ layers - endoderm, ectoderm, mesoderm - and lead to the establishment of the gut (Hejnol and Martindale, 2008; Lyons et al. 2015). In spiralian lophotrochozoan species like oysters, mesoderm develops from both ectodermal and endodermal-associated bipotential precursor cells, with ectomesodermal cells located near the blastopore (Perry et al. 2015). While gene expression atlases are not as well defined for bivalves compared to more established developmental models like sea urchin (Foster et al. 2020; Massri et al. 2021) and *Drosophila* (Sakaguchi et al. 2023), we were able to broadly define oyster gastrula cell clusters by assessing expression of evolutionarily conserved metazoan marker genes, including

identification of cell populations corresponding to highly ciliated cells, muscle, shell field, epidermis, neuronal precursors and some undifferentiated cells (Fig 3; Table 2 and Table S2). To annotate these 27 clusters with putative cell-type information (Fig. 1B), we queried the dataset using a combination of highly conserved bilaterian marker genes and previously identified genes expressed in specific tissues or cells in other molluscs.

2.1 Evolutionarily conserved germ layer and ciliated cell markers are expressed in C. gigas *gastrulae*

Bivalve embryos and larvae are often highly ciliated, with prototroch precursors developing during the gastrula stage (Cragg 1991). We identified multiple clusters representing putative ciliated cells with high levels of expression of well-established ciliary genes (Cil1-4; Table 2; Fig 3). For example, *rootletin* (CROCC (Ciliary Rootlet Coiled-Coil)) is one of the most specifically expressed genes in the ciliated cell clusters, along with *trichohyalin*, *dynein*, *tektin*, *foxJ*, and tubulins, which are expressed in ciliated cell types in other bilaterian larvae (Arenas-Mena et al. 2007; Paganos et al. 2021; Salamanca-Diaz et al. 2021; Piovani et al. 2023). Interestingly, *caveolin*, shown to be expressed in the developing prototroch of *Dreissena* trochophore larvae (Salamanca-Diaz et al. 2021), is one of the most specifically expressed genes in cluster Un3, which also has ectomesodermal expression signatures (Table 2).

We also detected expression of conserved genes involved in germ layer specification. Brachyury (*bra*) is a conserved T-box transcription factor with mesodermal expression across bilaterian taxa, including in the gastropod *Crepidula* (Perry et al. 2015) and the mussel *Dreissena* (Schulreich et al. 2022), as well as in the developing in shell field of gastropod larvae (Lartillot et al. 2002; Jackson and Degnan 2016). In *C. gigas* gastrulae, mesodermal markers *bra, twist,* and *snail2* expression is high in portions of some undefined clusters (Un2-5), as well as some cells in shell field cluster SF2 and muscle cluster Mus1 (Fig 3B), indicating that these clusters likely represent mesodermal derivatives. Some typical mesodermal genes like *hes* and *mox* (Sachslehner et al. 2021) were not detected in any of our datasets, possibly due to the early developmental stages targeted.

Bone morphogenetic protein 3 (*bmp3*), which is expressed in endomesodermal tissues in larval pearl oyster (Fan et al. 2019), shows widespread expression in putative endomesodermal clusters in *C. gigas* gastrulae, including portions of putative muscle precursors (Fig 3B). Hemocyte marker transketolase-like protein 2 (*tktl2.3*; Piovani et al. 2023) is highly expressed in putative circulatory-associated cluster Circ2, with lower expression in shell field clusters (Fig 3B). Caudal *(cad/cdx)* is a homeobox-containing transcription factor important for gastrulation, anterior/posterior axis specification, and shown previously to be expressed in ectodermal and ectomesodermal cells in mollusc embryos (Le Gouar et al. 2003; Perry et al. 2015; Johnson and Lambert 2021). In *C. gigas* gastrulae, *cdx* is expressed in clusters Neu6, Circ1, Mus2, and portions of undefined clusters Un2-5 (Fig 3B, Table 2). In the snail *Patella*, homeobox transcription factor goosecoid (*gsc*) is expressed in all three germ layers in the anterior of the embryo during gastrulation (Lartillot et al. 2002). We detected high expression of *gsc* in undefined clusters 2,3,5, as well as shell field cluster SF1 (Fig 3B).

Fig 3. Single-cell transcriptomics highlights tissue-specific progenitors in *C. gigas* gastrulae defined by expression of evolutionarily conserved germ layer markers and novel genes identified via cluster analysis. (A) Dot plot illustrating representative marker gene expression associated with each

UMAP cluster. Dot color represents expression level and dot size notes fractional representation of cells expressing a gene. (B) Expression patterns of select genes associated with developmental markers in UMAP space.

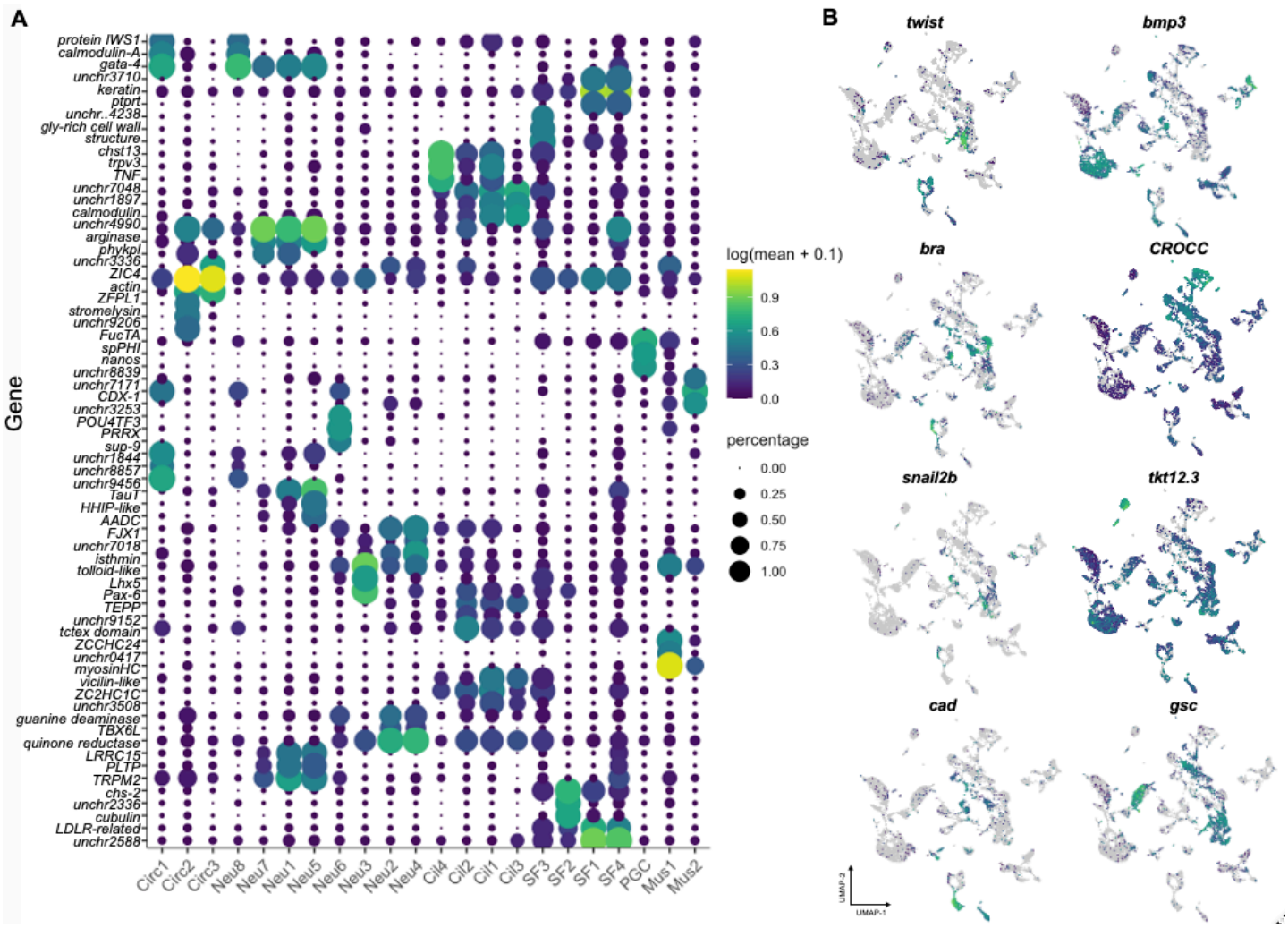


Table 2. Cluster identification.

UMAP Cluster	Cluster Name	Selected Informative Expressed Genes	Associated Germ Layer	Associated Tissue or Cell Type
1	Un1	HoxA7, HoxA1, HoxA5, HoxB7, eve, bmp1, bmp3, HTR	mesoderm	undefined
2	Un2	myosin ii HC, gsc, cdx, bmp2/4, gata2/3	endomesoderm	undefined
3	Un3	foxA, gsc, twist, chordin, dachshund, six3/6, bmp2/4, sox2	ectomesoderm	undefined
4	SF1	HoxA5, keratin, shematrin, prisilkin-39, bmp1	ectoderm	shell field
5	Un5	brachyury, cdx, foxA	mesoderm?	gut-associated?
6	SF2	bmp1, bmp2/4, bmp3, chs2, keratin, shematrin, prisilkin-39, gsc, sox2, gata2/3, HoxA5	ectoderm	shell field
7	Un4	brachyury, cdx, snail2b, gata2/3, myosin LC, HoxB7, bmp3, sox2	ectomesoderm	undefined
8	Neu1	six3/6, TRPM2, FMRFaR, ETNPPL, sox2, foxA, SERT	ectoderm	neural
9	Neu2	otp, gsc, HoxA1, bmp3, six3/6, prospero, sox2, FMRFaR, HTR, VMAT, four-jointed	ectoderm	neural
10	Cil1	zinc finger C2HC domain-containing protein 1C , CFAP44, CROCC, DNAH7, CFAP77, HTR1, four-jointed	ectoderm	ciliated cells
11	Mus1	bmp3, myosin LC, myosin HC, dachshund, brachyury, snail2a, snail2b, twist, otp	endomesoderm	muscle
12	Cil2	TUBA1A, PARK2, RSPH1, HTR1, four-jointed	ectoderm	ciliated cells
13	Neu3	pax6, LHX5, bmp1, PCDH9, myosin HC II-like	ectoderm	neural, sensory
14	Neu4	snail2b, six3/6, neurogenin-1, HTR, four-jointed, sox11	ectoderm	neural
15	Neu5	six3/6, DOPA decarboxylase, HHIPL1, gat3, dachshund, SERT	ectoderm	neural
16	Circ1	dachshund, foxA, caudal, cripto	ectomesoderm	circulatory
17	Neu6	PRRX2, sup-9, homeobox unc-42, caudal, VMAT	ectomesoderm	neural
18	Mus2	dachshund, myosins, bmp3, caudal, sox11	endomesoderm	muscle
19	PGC	nanos, vasa	PGCs	germ cells
20	Circ2	-1, nkx2.5	mesoderm?	circulatory
21	Circ3	myophilin, ZIC-4, rax, SBSB3, FMRFaR, nkx2.5	mesoderm?	circulatory
22	Neu7	six3/6, dachshund, foxQ2, FMRFaR, SERT	ectoderm	neural
23	Cil3	myosin HC II/trichohylin, foxJ	ectoderm	ciliated cells
24	Cil4	HTR1	ectoderm	ciliated cells
25	SF3	sodium-dependent glucose transporter 1A, mucin-17, HoxA5	ectoderm	shell field
26	SF4	shematrin, cytoskeletal keratins, HoxA5	ectoderm	shell field
27	Neu8	dachshund, DUXB, sox11, cdx, foxA	ectoderm	neural, sensory

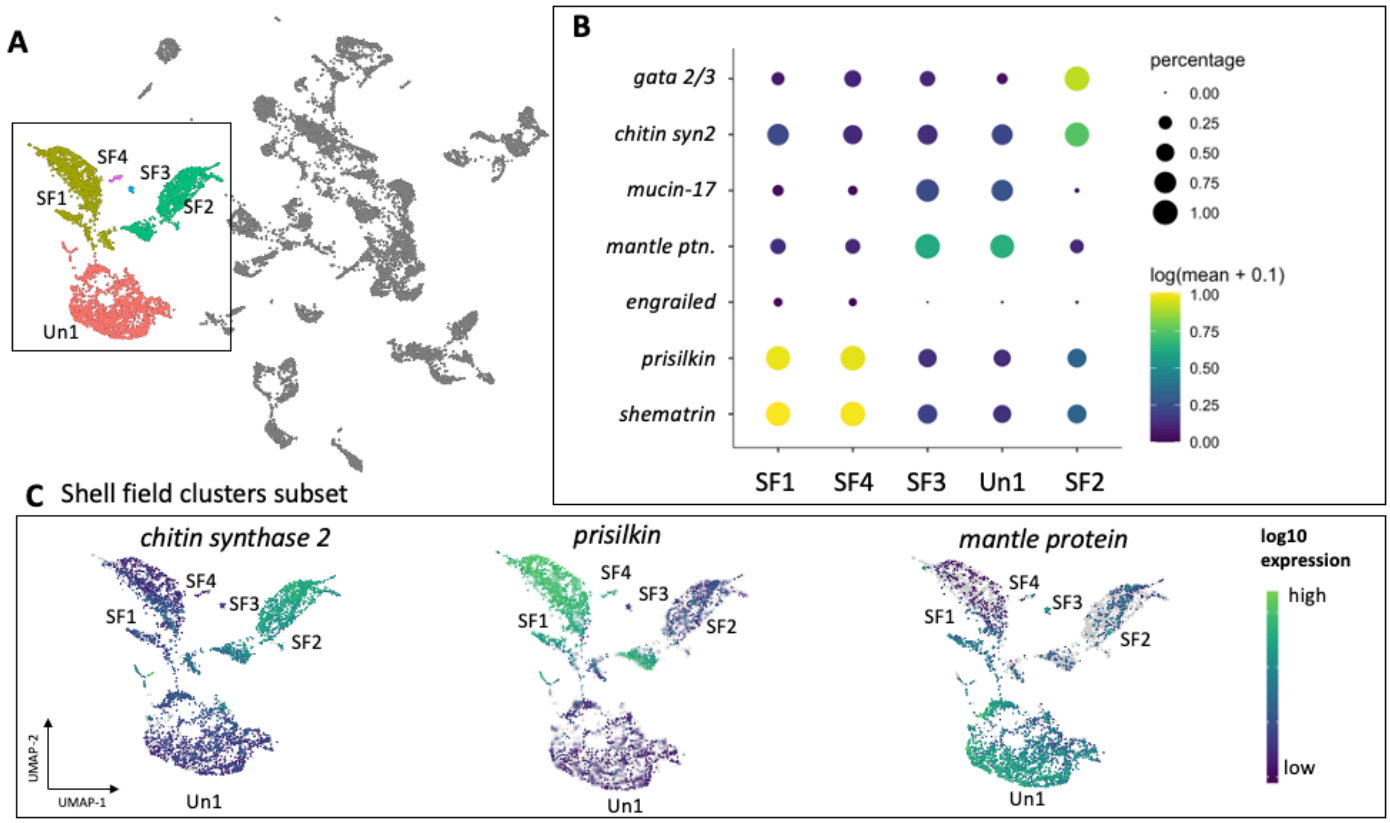
2.2 Biomineralization and ectodermal genes define shell field progenitors

Mollusc shells are composed primarily of calcium carbonate ($CaCO_3$), with smaller proportions of organic compounds like chitin and glycoproteins (reviewed in Suzuki and Nagasawa 2013). Larval molluscs have a structure called the shell field or shell gland that secretes shell material (Kniprath, 1981; Eyster and Morse, 1984). The shell field first appears as a visually identifiable structure during the gastrula stage as a thickened cap of ectoderm on the embryo's dorsal side (Moor 1983). Genes involved in synthesis of mollusc shell matrix have been identified in *C. gigas* (reviewed in Huang and Zhang 2022) and shown to be expressed in late trochophore larvae (Piovani et al. 2023). We found that *C. gigas* gastrulae have multiple cell

clusters expressing genes associated with developing shell field, with biomineralization pathway components being highly expressed in SF1, SF2 and SF4 (Fig 4; Table S2). Additionally, *gsc*, which is expressed in both the adult mantle and larval shell field in gastropods (Lartillot et al. 2002) is highly expressed in SF2 (Fig 3B). Gastrula shell field clusters also express *sox2*, *gata2/3*, and *engrailed*, other known genes involved in the developing shell field in *C. gigas* trochophores and veligers (Table 2; Huan et al., 2013; Liu et al., 2017; Min et al. 2023). *Hox1*, which is expressed in the shell field of freshwater mussel trochophores (Salamanca-Diaz et al. 2021) shows high levels of expression in clusters Un-1, as well as Neu4.

Interestingly, even at this early developmental stage, expression patterns of mantle or shell field-associated genes show differences between shell field clusters (Fig 4B,C). For example, genes involved in chitin synthesis are highly expressed in SF2, while *shematrin* and *prisilkin-39* are expressed in SF1 and SF4 (Fig 4), suggesting that these tissue-level shell field distinctions are becoming defined by the mid-gastrula stage. SF3 shows preferential expression of *mucin-17*; mucins have been shown to be involved in biomineralization processes in mussels (Marin et al. 2000). Interestingly, *mantle protein* (LOC105326721) is expressed at a higher level in Un1 and SF3 compared to the other putative shell field clusters, though these two clusters do not have substantial expression of biomineralization genes.

Figure 4. Multiple clusters expressing biomineralization and ectodermal genes suggest tissue-level shell field organization is becoming defined by the mid-gastrula stage. (A) UMAP color inset coded by clusters associated with the developing shell field in *C. gigas* gastrula embryos. (B) Dot plot illustrating expression of genes expressed in shell field subclusters. Dot color represents expression level and dot size represents fractional representation of cells expressing a gene. (C) Expression patterns of select genes associated with shell field clusters in UMAP space.



2.3 Muscle precursors

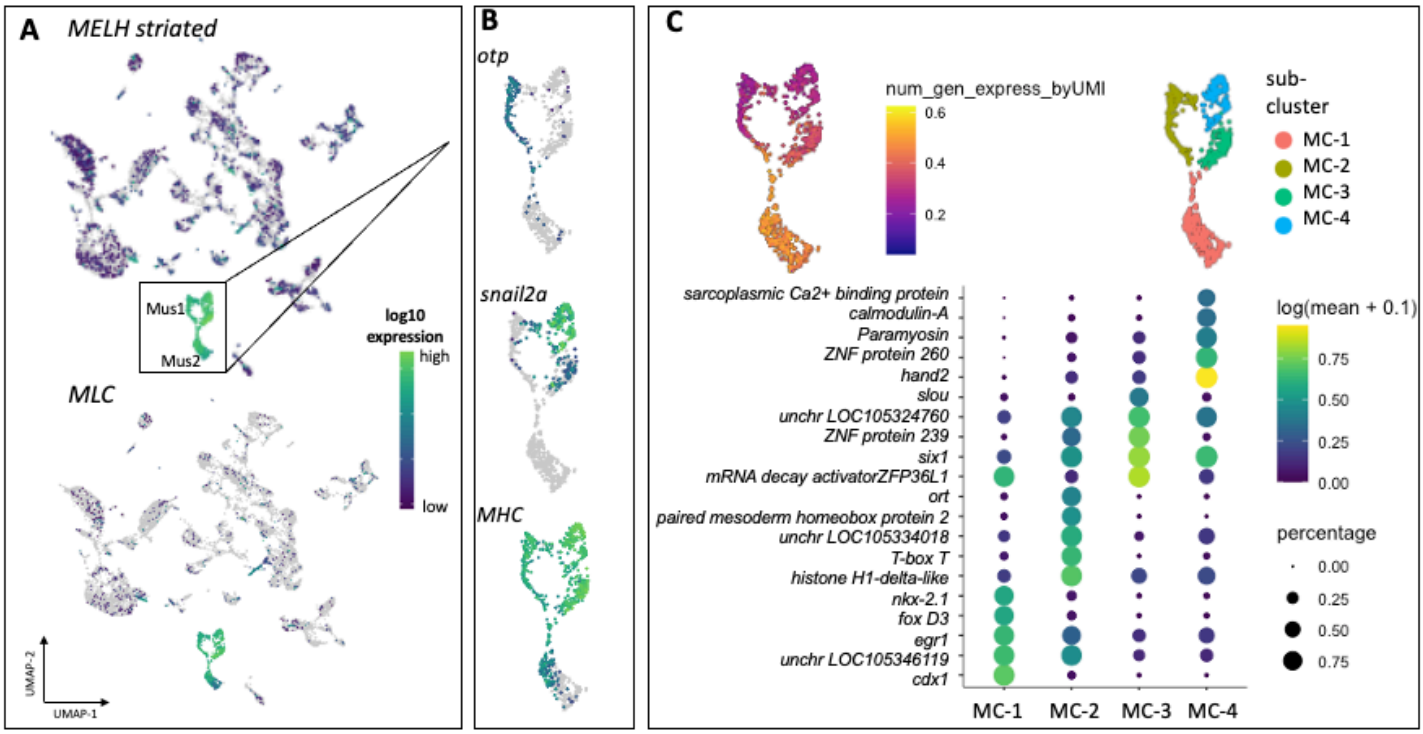
Two gastrula cell clusters show expression of known muscle marker genes including *myosins*, *troponins*, and *twist*, indicating that these clusters likely represent muscle precursors (Mus1, Mus2; Fig. 5; Table S2). Myosin essential light chain (*MELC*; striated adductor muscle; LOC105317061), which has been demonstrated by *in situ* hybridization to be expressed as early as blastula stage in *C. gigas* (Yu et al. 2017), is highly expressed in both clusters Mus1, Mus2 (Fig. 5A). An additional *MELC* associated with smooth adductor muscle (LOC105320997) also shows high levels of expression in the putative muscle precursor clusters. *Twist*, a basic helix-loop-helix (bHLH) transcription factor essential for development of the mesoderm, has been identified as a marker gene for muscle progenitor cells within the adductor muscle in adult scallops (Sun et al. 2021). In *C. gigas* gastrulae, *twist* is expressed in cluster Mus1, and its highest expression was observed in portions of putative ectomesodermal cluster Un3 (Fig 3B).

Within the clusters of muscle precursors, there appear to be some further separation of cells based on gene expression differences. For example, *otp* is expressed in only half of cluster Mus1 (Fig 5B). Interestingly, *bmp3*, which is expressed in developing endomesoderm in the pearl oyster (Fan et al. 2019), is expressed throughout Mus2 and in the 'left' portion of the Musc1 cell grouping, but is absent from the rest of that cluster (Fig. 3B). Myosin heavy chain (*mhc*, LOC105338907) is highly expressed in cluster Mus1, with some expression of adjacent cells in cluster Mus2 (Fig 5B).

To assess developing muscle cell progression through the differentiation process, we performed focused clustering and pseudotime analyses of clusters Mus1 and EnM-Mus2. These

two clusters subsequently yielded four sub-clusters (MC 1-4) in the pseudotime analysis and a potential developmental trajectory based on the number of genes expressed per cell, which can be used as a proxy for developmental progress (Gulati et al. 2020) (Fig 5C). Muscle cluster-1 is defined by high expression of transcription factors associated with endomesodermal and muscle specification (Fig 5C), consistent with an analysis showing a relatively higher number of genes expressed/cell suggestive of a "younger" developmental age (Gulati et al. 2020) (Fig 5C). Mus3 shows slightly fewer genes expressed per cell compared to Mus1, but cells in this cluster also express genes associated with muscle tissue differentiation. For example, a homolog of the *Drosophila* gene *slouch* (LOC105345169) - a homeodomain-containing protein involved in muscle specification (Knirr et al. 1999) - is one of the defining marker genes for Mus3 (Fig 5C).

Figure 5. Two clusters expressing genes associated with the development of larval muscle in *C. gigas* gastrulae. (A) Expression of myosin essential light chain - striated muscle (*MELH striated*) and myosin light chain (*MLC*) are highest in the two putative muscle-associated clusters of gastrulae in UMAP space. (B) Expression of muscle-related genes *otp*, *snail2a* and *MHC* show variation in expression within each putative muscle cluster. (C) Visualizing the number of genes expressed per cell suggests a potential developmental trajectory, with cells from cluster Mus1 representing the developmentally oldest cells. (D) Sub-cluster analysis of muscle-associated clusters identified four sub-clusters. (E) Dot plot illustrating expression of muscle-associated marker genes of each associated sub-cluster. Dot color represents expression level and dot size represents fractional representation of cells expressing a gene.



2.4 Putative neural and sensory cell precursors

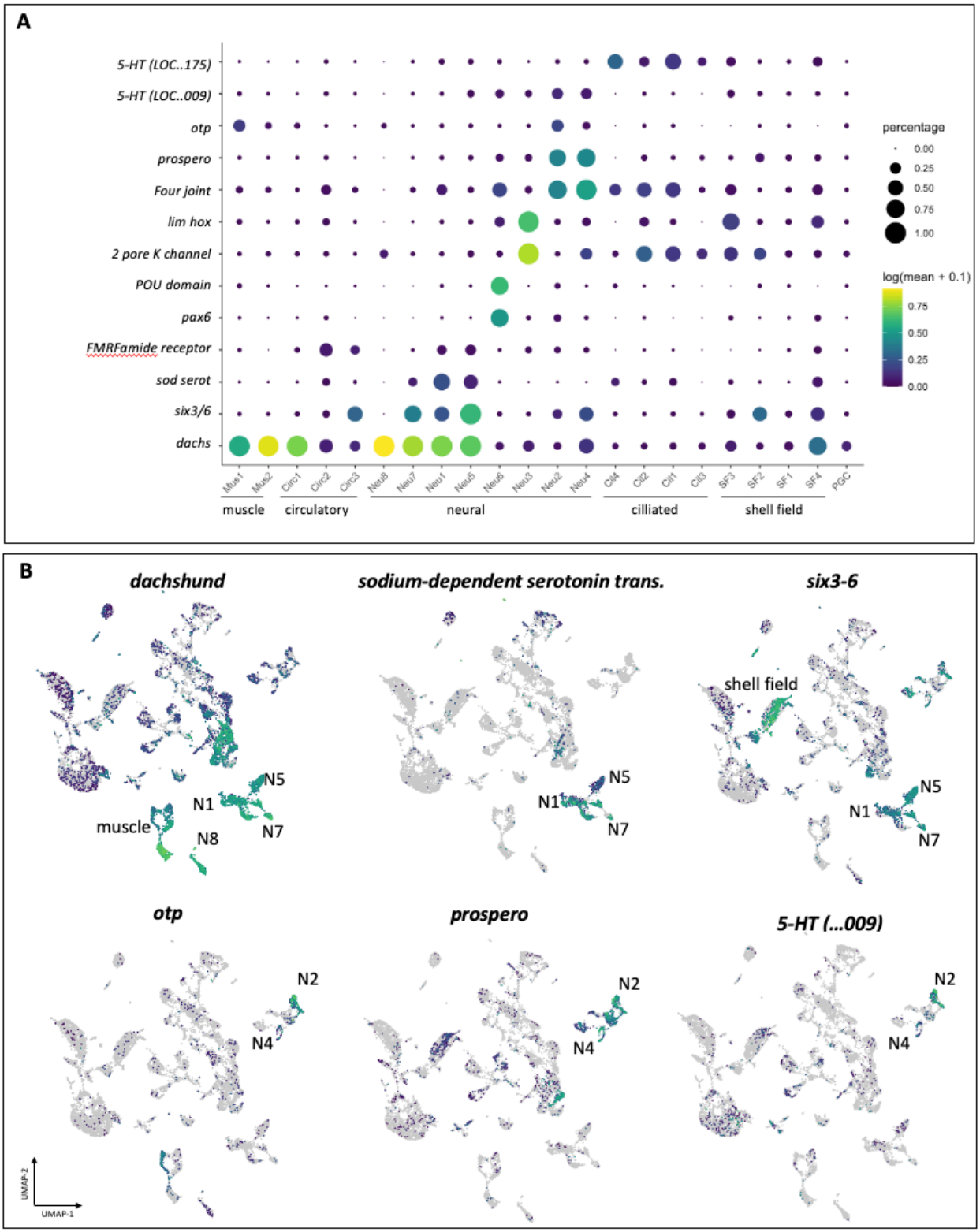
Immunostaining and histology of larval bivalve neuroanatomy shows that veliger larvae possess an apical sensory organ and several paired ganglia (visceral, cerebral, and pleural) connected via a lateral cord (Ellis and Kempf 2011; Yurchenko et al. 2019). Mollusc trochophore

larvae possess fewer defined neural structures compared to veligers, though trochophores have an apical sensory organ and some identifiable nerves (Voronezhskaya 2004; Yurchenko et al. 2019; Piovani et al. 2023). In our gastrula-stage samples, we identified multiple cell clusters expressing conserved neuronal or sensory-related genes (Neu1-8; Fig 6, Table 2, Table S2). Strikingly, we detected disparate clusters of putative neural or sensory cells based on distinct patterns of neural gene expression, even at this relatively early embryonic stage (Fig 6). These include clusters of cells expressing the neuronal transcription factor *prospero* (highest in clusters Un3, Neu2, Neu4; Doe et al. 1991), four-jointed box kinase (*four-jointed*; clusters Neu2, Neu4, Cil1, Cil2; Zeidler et al. 1999), *sodium-dependent serotonin transporter* (clusters Neu1, Neu5, Neu7), and the neural and eye-associated nuclear protein *dachshund* (clusters Neu5, Neu7; Popov et al. 2010).

Emergence of peripheral sensory neurons has also been reported during the early trochophore stage in *Crassostrea*, preceding the appearance of the apical organ (Yurchenko et al. 2018; Yurchenko et al. 2019; Piovani et al. 2023). The 5-hydroxytryptamine receptor (*5-HTR*; LOC105348009, LOC105345175), associated with the developing apical organ in oyster trochophores (Yurchenko et al. 2018), shows high levels of expression in clusters Neu1, Neu4, and Cil1 (Fig. 6A-B). Orthopedia (*otp*) is a homeodomain-containing transcription factor expressed during nervous system development in diverse clades including molluscs (Nederbragt et al. 2002; Perry et al. 2015). Interestingly, we detected *otp* expression in only one neural-associated cluster (Neu2). Nederbragt et al. described a small subset of *otp*-expressing cells descended from 1m111 micromeres that later associate with the apical tuft (Nederbragt et al 2002). The homeodomain-containing transcription factor *six3/6* is involved in the development of sensory structures in diverse metazoans, and has been shown to be expressed in the neural structures of mollusc embryos and larvae (Perry et al. 2015; Redl et al. 2018). We found that *six3/6* is expressed in most predicted neuronal or sensory cell clusters, as well as in some shell field clusters (Fig 6).

In *C. gigas* trochophore larvae, FMRF+ neurons have been detected in early ventral neurons and associated with/surrounding the apical organ through histological staining (Yurchenko et al. 2018). The *FMRFamide receptor* gene shows some expression in our gastrulae dataset, though at relatively low levels (Fig 6A; Fig S5). Some known neural genes such as *hex* (LOC105339244) are expressed in other mollusc larvae but were not detected in our cleavage+blastula (Fig S5) or gastrula datasets, possibly because these genes are not expressed until later in larval development.

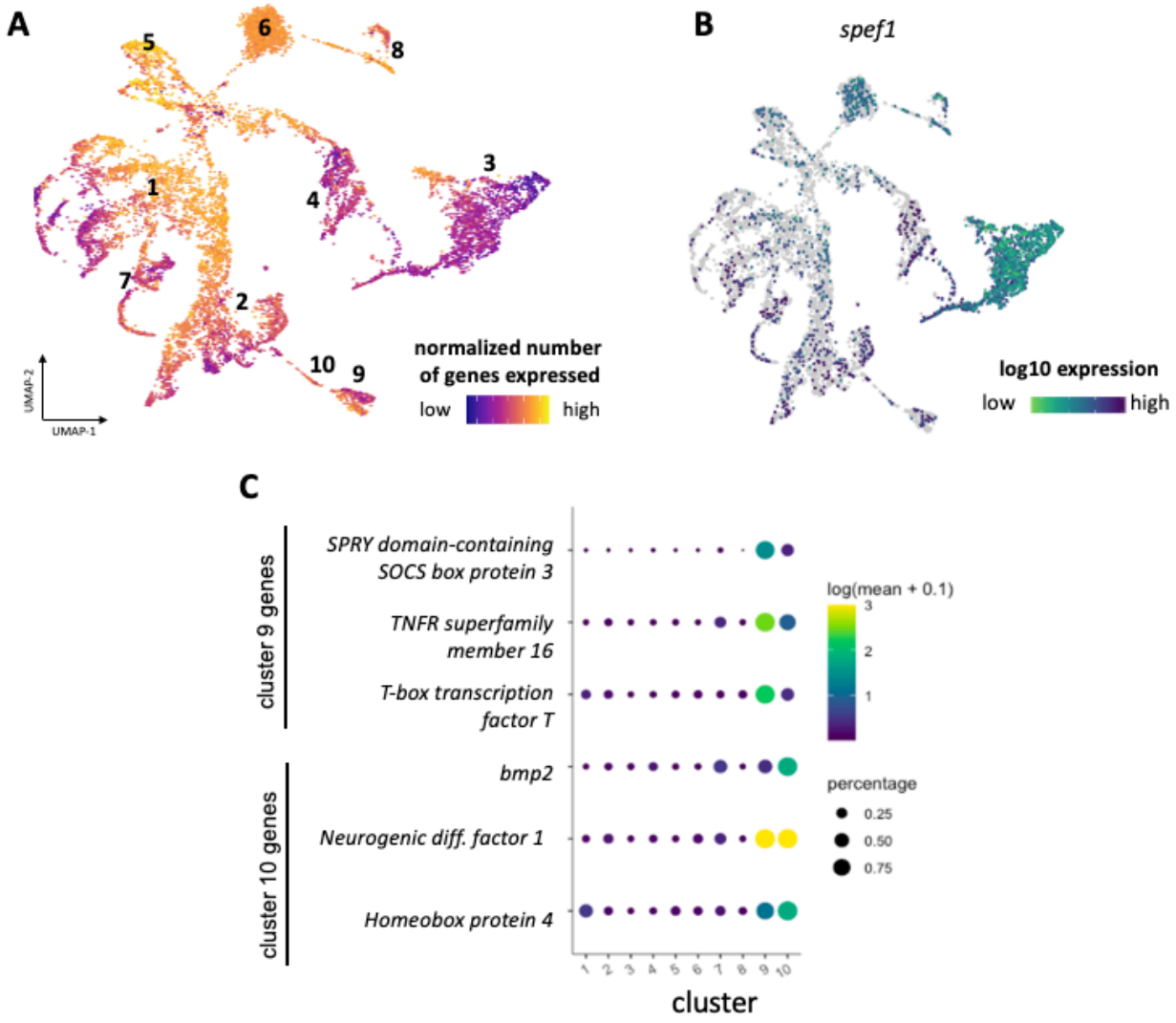
Figure 6. Multiple cell clusters expressing conserved neuronal or sensory-related genes are present in *C. gigas* gastrulae. Dot plot (A) and UMAP (B) representation of expression of select genes associated with neural clusters. For the dot plot and UMAPs, dot color represents expression level, and for the dot plot, dot size represents fractional representation of cells expressing a gene. For the UMAP, clusters with a high proportion of cells expressing a given gene are labeled for orientation.



3. Early signatures of transcriptional states of cleavage-stage and blastula embryos
We sequenced a transcriptional atlas of earlier developmental time points including cells from cleavage-stage and blastula embryos, which represent a continuum of the earliest stages of

embryonic development (Fig 7). Overall, there was much lower heterogeneity in the cleavage + blastula cells, as expected, based on the shared pool of maternal mRNAs and the cells' undifferentiated state (Andeol 1994, Tadros and Lipschitz 2009). Cluster analysis identified 10 major clusters in these cleavage+blastula cells (Fig 1). While cluster annotation is challenging at this early stage, we could see signals of ciliated cells of the rotating oyster blastula in cluster cb_3, which were expressing genes such as sperm flagellar protein 1 (*spef1*; LOC105339310) and EF-hand calcium binding domain-containing protein 6 (*efcab6*; LOC105337196) (Fig 7B, Table S3). Additionally, a marker gene analysis showed clusters cb_9 and cb_10 had the highest specificity of all the clusters (Fig 7C) and many of these specific genes were transcription factors or other regulatory proteins (Fig 7D). Developmental ordering, approached two ways, (1) visualization of cells expressing the fewest genes (which are predicted to be the most differentiated cells relative to progenitors (Gulati et al. 2020)) and (2) using visualization of the blastula cells in UMAP space and (Fig 7A and Fig S6 respectively) confirmed that cluster cb_9 and portions of cluster cb_3 are likely to be some of the developmentally oldest cells. Blastula samples were also analyzed separately. Clustering and marker gene analysis of blastula cells only shows enrichment for diverse transcription factor families (Fig S7, Table S4), an indication of embryonic genome activation (Latham and Schultz 2001; Graf et al. 2014; Colonnetta et al. 2023).

Figure 7. Early signatures of transcriptional states of cleavage-stage and blastula embryos, representing a continuum of the earliest stages of embryonic development in *C. gigas*. (A) UMAP of cleavage + blastula color-coded by the normalized number of genes expressed per cell (i.e. number of genes expressed/number of UMI per cell). (B) UMAP representation of expression of select gene associated with the ciliated cluster of cleavage + blastula cells. (C) Dot plot representation of expression of top 3 marker genes associated with cluster 9 and 10 of the cleavage + blastula cells. Dot color represents expression level and for the dot plot dot size represents fractional representation of cells expressing a gene.



DISCUSSION

Sterile shellfish are both a market-driven need and an ecologically sustainable approach to increase food production via aquaculture. Technological advances in sterility induction via inactivation of genes essential for germ cell formation and development in finfish aquaculture (Wong and Zohar, 2015, Wargelius et al., 2016, Xu et al. 2023), have poised the shellfish industry to incorporate this approach, but the lack of knowledge of the genes essential for PGC specification remains a significant barrier. To overcome this limitation, we have applied scRNA-seq to early stages of Pacific oyster embryos to identify genes involved in PGC specification. We present a robust suite of early developmental gene expression data using scRNA-seq in the Pacific oyster *C. gigas*. In gastrulae, though no discrete tissues are formed and the trochophore body plan is not yet defined, we identified cell clusters expressing evolutionarily conserved marker genes associated with bivalve larval tissues or cell types

including muscle, shell field, neuronal cells, ciliated cells, and PGCs. We show that the molecular identity of oyster PGCs is defined by gastrula-stage, and identified a number of novel genes specifically expressed in these putative PGCs. These data also provide a rich transcriptional atlas that can be mined to address both basic and applied research questions in Pacific oyster, an emerging molluscan model species with significant commercial and ecological importance.

Characterization of genes associated with PGC formation

We were particularly interested in the unique transcriptomic signatures associated with formation of PGCs in *C. gigas*. These data are necessary to identify potential gene targets that could be disrupted in the pursuit of generating reproductively sterile shellfish for use in aquaculture. The early developmental stages sequenced in this study were selected to focus on genes potentially involved in PGC specification, a biological process that occurs very early in embryonic development. Using the conserved germline marker gene *vasa* to initially select cells in oyster gastrula that likely represent PGCs, we also detected transcriptomic signatures in these *vasa*-positive cells that include both novel and previously described germline associated genes in other invertebrates.

A gene uniquely expressed in the PGC cluster is a histone variant, sperm-specific protein PHI-2B/PHI-3 (*spPHI*), which has not been functionally characterized in *C. gigas* but exhibits sequence similarity to a protamine-like gene uniquely expressed in spermatozoa of *Mytilus californianus* (Carlos et al. 1993). It is interesting to consider the role for a protamine-like or H1 linker gene in germ cell specification in *C. gigas* as many animals have embryo and germline-specific H1 linker proteins (Perez-Montero et al. 2015). In *C. gigas*, *spPHI* is broadly expressed in the earliest embryos, but expression is restricted to the PGC cluster by gastrulation. This pattern is similar to expression patterns in *Drosophila*, in which the embryonic H1 (dBigH1) gene is expressed only in the germline, where it acts to maintain transcriptional quiescence (Perez-Montero et al. 2013). A thorough spatial, temporal, and functional characterization of the *spPHI* transcripts in *C. gigas* is necessary to determine the role of *spPHI* in embryogenesis and germline formation in oysters.

In addition to known candidate genes involved in germline identity in other animals, this study sought to identify mollusc- or oyster-specific PGC genes that may not have been detected using a candidate approach. Indeed, an uncharacterized gene *LOC105328839* has the highest expression specificity to the PGC cluster in *C. gigas* gastrulae. This gene appears to be bivalve-specific and has few conserved protein domains that may suggest a specific molecular characterization and role in the embryo. Importantly, for both s*pPH*I and *LOC105328839*, we show that expression peaks early in development, prior to the larval trochophore stage and remains relatively low through the juvenile stage, suggesting their embryonic functions may be exclusively associated with the developing germline.

Interestingly, with the exception of *nanos* and *vasa*, our study did not identify most of the previously identified gonad-specific genes reported in *C. gigas* adult gonadal tissue (Wang et al. 2024). It is possible that there may be few shared genes involved in both embryonic PGC specification and adult germinal stem cell populations in *C. gigas*, although differences in methodologies and detection between studies (i.e. sequencing depth) may be a confounding factor (Ziegenhain et al. 2017). Nevertheless, our identification of genes associated with PGCs

in *C. gigas* may help with future characterization of GSCs and processes related to gametogenesis in oysters.

In addition to providing information about potential genes involved in PGC specification in bivalves, the experimental design employed here allows for some insight into mechanisms of specification. There is a general lack of consensus in the literature on whether molluscan PGCs are specified via preformation or epigenesis (Extavour and Akam 2003; Obata and Komaru 2012). At the molecular level, previous work in *C.gigas* showed *vasa* expression to be asymmetrically localized in cleavage-stage embryos, suggesting that the oyster germline is specified via preformation (Fabioux et al. 2004). Our results are generally consistent with that supposition, as *vasa* expression is widespread in cleavage and blastula-stage embryos; however, cells with the highest abundance of *vasa* transcripts are concentrated within a smaller region of cluster 2 (Fig 2). These *vasa* transcripts potentially represent maternal transcripts undergoing sequestration to the presumptive germline. More recently, however, a study evaluating embryonic expression of *nanos* in Pacific oysters led to speculation that the process of formation of PGCs in *C. gigas* may involve elements of both epigenesis and preformation, perhaps in an evolutionarily transitional state (Xu et al. 2018).

In our study, *nanos* expression was highly specific to the PGC cluster in gastrulae but not highly expressed in any cluster of the cleavage-blastula stages. Other evidence, such as the expression of PGC cluster-specific genes, suggests PGC formation in oysters is occurring via induction of signaling and transcriptomic activity from the embryo. Interestingly, even in well-characterized model systems such as *Drosophila*, the traditional paradigm of exclusivity with regards to whether preformation or epigenesis yields formation of PGCs is being questioned as new data emerge suggesting both maternal and inductive signals are important (Colonetta 2022). While the results here represent a transcriptional snapshot in developmental time, and we cannot definitively resolve the mechanism of PGC formation in *C. gigas*, they imply the importance of both maternal and zygotically-derived cell differentiation factors.

Early signatures of larval tissue formation in C. gigas

In invertebrate gastrulae, single-cell sequencing has been an effective technique to identify cell-type differentiation trajectories, discover novel genes associated with developing tissues, and better understand gene regulatory networks (Foster et al. 2021; Massri et al. 2021; Satoh et al. 2022; Sakaguchi et al. 2023). While there are single-cell transcriptomic resources for oyster trochophore larvae (Piovani et al. 2023) and a variety of adult tissues (de La Forest Divonne et al. 2024; Wang et al. 2024), earlier developmental time points had not previously been reported in oysters or other lophotrochozoan taxa. We posit that the cleavage-through-blastula and gastrula single-cell RNA-seq datasets presented here are robust resources for exploring transcriptional activities, cell states, and cell-type differentiation early in oyster embryonic development.

Gastrula sequencing data revealed several distinct cell clusters that express genes known to be involved in mollusc shell formation (Lopez-Anido et al. 2024). Studies to date exploring molecular mechanisms of pre-larval shell formation have typically focused on co-expression of individual genes. Our findings are consistent with the results of Liu et al. (2020a), who identified multiple defined cell populations from the shell gland of *C. gigas* gastrulae using in-situ hybridization for known mollusc mantle genes. Even at this early

developmental stage, we identified four clusters associated with the developing shell field as well as a fifth cluster that highly expresses *mantle protein* (Fig 4). Embryonic insights into the sequestration of distinct regions of developing mantle tissues can inform evolutionary mechanisms driving the observed array of mollusc shell morphologies (Marie 2012; Jackson 2021; Lopez-Anido et al. 2024). Additionally, these data can inform how larval shell formation may be perturbed by environmental stressors, such as ocean acidification. Ocean acidification can impact many physiological systems in marine invertebrates, but larval shell formation has been shown to be particularly sensitive (reviewed in Parker et al. 2013). Transcriptomic studies of whole bivalve larvae have hinted at mechanisms underlying developmental delays observed upon OA exposure (de Witt et al. 2018, Liu et al. 2020b). The data provided here, at single-cell resolution, could provide additional information and identify potential transcriptomic targets to gain further insights into this process.

Many planktonic marine larvae, including bivalve trochophores and veligers, are highly ciliated, allowing them to move through the water column. In both the cleavage-blastula and gastrula datasets, we identified cell clusters with gene expression signatures associated with larval ciliated cells (Piovani et al. 2021). Genes associated with microtubules, cilia, and flagellar proteins were highly expressed in a subset of gastrula clusters, indicating the presence of several ciliated cell types. In bivalve larvae, sensory cells such as those associated with the apical tuft in trochophores, can also be ciliated (Nikishchenko et al. 2024). We identified multiple clusters in *C. gigas* gastrulae that are expressing genes known to be present in early peripheral sensory cells and other neural progenitors, including *5-hydroxytryptamine receptor* which is expressed in the apical organ in oyster trochophores (Yurchenko et al. 2018). Planktonic larvae, particularly in species with feeding (planktotrophic) larvae such as Pacific oyster, must be able to sense their environment, control swimming direction through phototaxis and gravity sensing, and find prey. It is perhaps not unexpected that even at the gastrula stage, gene expression for developing sensory cells and neurons was detected in our data.

Our data represent the earliest developmental time-points sequenced at the single-cell level in a molluscan species. By identifying genes uniquely expressed within each cluster of putative differentiating cell or tissue types including both evolutionarily conserved marker genes and novel uncharacterized mollusc-specific genes, these data can be a resource for pairing uncharacterized genes with those known to be involved in cell-type or tissue information. The gastrula single-cell atlas presented here should also inform future studies exploring cell-type identity and developmental trajectories in larval bivalves. It is important to note that although the gastrula cells represent a single sampling time-point, oyster development is asynchronous (Allen et al. 1989), and therefore we expected the embryos to represent a range of developmental stages centered around the gastrulae. We observed this most clearly in the putative muscle clusters, where there appeared to be regions of different levels of transcriptional activity based on numbers of distinct genes being expressed per cell.

We combined the earliest embryos into a pool representing all cleavage stages, and found that by visualizing the number of genes expressed per cell in 2d space as a proxy for developmental age (Gulati et al. 2020) and anchoring the 2d map with a distinct blastula library, we could identify a developmental progression of cells that is consistent with their transcriptional continuity in low dimensional space. This strategy was a crucial aspect of being able to survey

this entire early developmental progression at single-cell resolution at a reasonable cost and still be able to evaluate developmental timing of germline-related transcripts.

Notes:

Tree of life of early embryos

How do these datasets contribute to broader knowledge of evolution

Identification of conserved GRNs for major developmental milestones, tissue types, germ layers

Having full transcriptomes of developing neuron, to find other important genes involved in cell type specification or function

Striking how with a few marker genes from what was available, we were able to give identities to most cell clusters in our gastrula samples. We weren't sure how much we would be able to interpret or analyze or conclude in this very early developmental stage, but were surprised to find a lot of organization in these embryos based on gene expression profiles. Even though distinct morphological structures are not yet visible in gastrula, gene expression signatures associated with specific tissues or cell types were already active. This makes sense, because the organization of the trochophore larva is being set up throughout gastrulation. This transcriptional knowledge in oyster gastrula can be applied to broader embryonic tree of life efforts that are seeking to understand the foundational drivers of developmental cell type differentiation.

CONCLUSIONS

These sequencing data and cell type atlas provide a rich dataset exploring transcriptional activities and cell-type differentiation during early oyster development. The gastrula single-cell atlas presented here should also inform future studies exploring cell-type identity and developmental trajectories in larval bivalves. Notably, we have identified a suite of candidate genes that can be explored for their role in oyster PGC specification and to advance efforts to develop methods to achieve reproductive sterility via germ cell elimination in cultured Pacific oysters and potentially other shellfish species. Taken together, this dataset shows the utility of scRNA-seq to describe cell type differentiation and PGC specification during mollusc development, a group that has been understudied from this perspective.

ACKNOWLEDGEMENTS

Internal Competitive Aquaculture Funds (ICAF) from the NOAA Office of Aquaculture to Mackenzie Gavery and Adam Luckenbach. This work publication was funded in part by a grant from Washington Sea Grant, University of Washington, pursuant to National Oceanic and Atmospheric Administration Award No. NA18OAR4170095 to Steven Roberts. The views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA or any of its sub-agencies. We are also grateful to Olivia Fitzgerald for administrative support.

AUTHOR CONTRIBUTIONS

This project was conceived and coordinated by MG, BV, JAL and SR. The experiments were carried out by MG and LS. Data analyses were performed by MG, LV, LS, SR. The paper was written and revised by MG, LV, LS, BV, JAL, CT and SR.

COMPETING INTERESTS

C.T. is a scientific advisory board member, consultant and/or co-founder of Algen Biotechnologies, Altius Therapeutics and Scale Bioscience.

REFERENCES

Agriculture Organization of the United Nations. Fisheries Department. The state of world fisheries and aquaculture. Food and Agriculture Organization of the United Nations; 2018.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of molecular biology. 1990 Oct 5;215(3):403-10.

Andéol Y. Early transcription in different animal species: implication for transition from maternal to zygotic control in development. Roux's archives of developmental biology. 1994 Oct;204:3-10.

Arenas-Mena C, Wong KS, Arandi-Forosani N. Ciliary band gene expression patterns in the embryo and trochophore larva of an indirectly developing polychaete. Gene Expression Patterns. 2007 Apr 1;7(5):544-9.

Bayne BL, Ahrens M, Allen SK, D'auriac MA, Backeljau T, Beninger P, Bohn R, Boudry P, Davis J, Green T, Guo X. The proposed dropping of the genus Crassostrea for all Pacific cupped oysters and its replacement by a new genus Magallana: a dissenting view. Journal of Shellfish Research. 2017 Dec;36(3):545-7.

Bickel JS, Chen L, Hayward J, Yeap SL, Alkers AE, Chan RC. Structural maintenance of chromosomes (SMC) proteins promote homolog-independent recombination repair in meiosis crucial for germ cell genomic stability. PLoS genetics. 2010 Jul 22;6(7):e1001028.

Carlos S, Jutglar L, Borrell I, Hunt DF, Ausio J. Sequence and characterization of a sperm-specific histone H1-like protein of Mytilus californianus. Journal of Biological Chemistry. 1993 Jan 5;268(1):185-94.

Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, Adey A. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science. 2017 Aug 18;357(6352):661-7.

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C. The single-cell transcriptional landscape of mammalian organogenesis. Nature. 2019 Feb 28;566(7745):496-502.

Colonnetta MM, Goyal Y, Johnson HE, Syal S, Schedl P, Deshpande G. Preformation and epigenesis converge to specify primordial germ cell fate in the early Drosophila embryo. PLoS genetics. 2022 Jan 5;18(1):e1010002.

Colonnetta MM, Schedl P, Deshpande G. Germline/soma distinction in Drosophila embryos requires regulators of zygotic genome activation. Elife. 2023 Jan 4;12:e78188.

Cragg SM. The biology of scallop larvae. Scallop: Biology, Ecology and Aquaculture. 1991:75-132.

de La Forest Divonne S, Pouzadoux J, Romatif O, Montagnani C, Mitta G, Destoumieux-Garzon D, Gourbal B, Charrière GM, Vignal E. Diversity and functional specialization of oyster immune cells uncovered by integrative single cell level investigations. bioRxiv. 2024 Jul 23:2024-07.

De Wit P, Durland E, Ventura A, Langdon CJ. Gene expression correlated with delay in shell formation in larval Pacific oysters (Crassostrea gigas) exposed to experimental ocean acidification provides insights into shell formation mechanisms. BMC genomics. 2018 Dec;19:1-5.

Doe CQ, Chu-LaGraff Q, Wright DM, Scott MP. The prospero gene specifies cell fates in the Drosophila central nervous system. Cell. 1991 May 3;65(3):451-64.

Doniach T, Hodgkin J. A sex-determining gene, fem-1, required for both male and hermaphrodite development in Caenorhabditis elegans. Developmental biology. 1984 Nov 1;106(1):223-35.

Downing SL, Allen Jr SK. Induced triploidy in the Pacific oyster, Crassostrea gigas: optimal treatments with cytochalasin B depend on temperature. Aquaculture. 1987 Mar 1;61(1):1-5.

Ellis I, Kempf SC. Characterization of the central nervous system and various peripheral innervations during larval development of the oyster Crassostrea virginica. Invertebrate Biology. 2011 Sep;130(3):236-50.

Extavour CG, Akam M. Mechanisms of germ cell specification across the metazoans: epigenesis and preformation.

Extavour CG. Evolution of the bilaterian germ line: lineage origin and modulation of specification mechanisms. Integrative and comparative biology. 2007 Nov 1;47(5):770-85.

Eyster LS, Morse MP. Early shell formation during molluscan embryogenesis, with new studies on the surf clam, Spisula solidissima. American Zoologist. 1984 Nov 1;24(4):871-82.

Fabioux C, Pouvreau S, Le Roux F, Huvet A. The oyster vasa-like gene: a specific marker of the germline in Crassostrea gigas. Biochemical and biophysical research communications. 2004 Mar 19;315(4):897-904.

Fabioux C, Huvet A, Lelong C, Robert R, Pouvreau S, Daniel JY, Minguant C, Le Pennec M. Oyster vasa-like gene as a marker of the germline cell development in Crassostrea gigas. Biochemical and biophysical research communications. 2004 Jul 23;320(2):592-8.

Fabioux C, Corporeau C, Quillien V, Favrel P, Huvet A. In vivo RNA interference in oyster–vasa silencing inhibits germ cell development. The FEBS Journal. 2009 May;276(9):2566-73.

Fan S, Zhou D, Xu Y, Yu D. Cloning and functional analysis of BMP3 in the pearl oyster (Pinctada fucata). Journal of Applied Animal Research. 2019 Jan 1;47(1):250-61.

Forbes A, Lehmann R. Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. Development. 1998 Feb 15;125(4):679-90.

Foster S, Oulhen N, Wessel G. A single cell RNA sequencing resource for early sea urchin development. Development. 2020 Sep 1;147(17):dev191528.

Galindo-Torres P, Ventura-López C, Llera-Herrera R, Ibarra AM. A natural antisense transcript of the fem-1 gene was found expressed in female gonads during the characterization, expression profile, and cellular localization of the fem-1 gene in Pacific white shrimp Penaeus vannamei. Gene. 2019 Jul 20;706:19-31.

Guévélou E, Carnegie RB, Small JM, Hudson K, Reece KS, Rybovich MM. Tracking triploid mortalities of eastern oysters Crassostrea virginica in the Virginia portion of the Chesapeake Bay. Journal of Shellfish Research. 2019 Apr;38(1):101-13.

Graf A, Krebs S, Zakhartchenko V, Schwalb B, Blum H, Wolf E. Fine mapping of genome activation in bovine embryos by RNA sequencing. Proceedings of the National Academy of Sciences. 2014 Mar 18;111(11):4139-44.

Grabowski JH, Peterson CH. Restoring oyster reefs to recover ecosystem services. Theoretical ecology series. 2007 Jan 1;4(10.1016):80017-7.

Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, Ilagan F, Kuo AH, Hsieh RW, Cai S, Zabala M. Single-cell transcriptional diversity is a hallmark of developmental potential. Science. 2020 Jan 24;367(6476):405-11.

Guo X, Wang Y, Xu Z, Yang H. Chromosome set manipulation in shellfish. InNew technologies in aquaculture 2009 Jan 1 (pp. 165-194). Woodhead Publishing.

Gustafson EA, Wessel GM. Vasa genes: emerging roles in the germ line and in multipotent cells. Bioessays. 2010 Jul;32(7):626-37.

Hejnol A, Martindale MQ. Acoel development indicates the independent evolution of the bilaterian mouth and anus. Nature. 2008 Nov 20;456(7220):382-6.

Houssin M, Trancart S, Denechere L, Oden E, Adeline B, Lepoitevin M, Pitel PH. Abnormal mortality of triploid adult Pacific oysters: Is there a correlation with high gametogenesis in Normandy, France?. Aquaculture. 2019 Apr 30;505:63-71.

Huan P, Liu G, Wang H, Liu B. Identification of a tyrosinase gene potentially involved in early larval shell biogenesis of the Pacific oyster Crassostrea gigas. Development genes and evolution. 2013 Nov;223:389-94.

Huang J, Zhang R. The mineralization of molluscan shells: Some unsolved problems and special considerations. Frontiers in Marine Science. 2022 Jun 27;9:874534.

Jackson DJ, Degnan BM. The importance of evo-devo to an integrated understanding of molluscan biomineralisation. Journal of structural biology. 2016 Nov 1;196(2):67-74.

Jackson HJ, Larsson J, Davison A. Quantitative measures and 3D shell models reveal interactions between bands and their position on growing snail shells. Ecology and Evolution. 2021 Jun;11(11):6634-48.

Johnson AB, Fogel NS, Lambert JD. Growth and morphogenesis of the gastropod shell. Proceedings of the national academy of sciences. 2019 Apr 2;116(14):6878-83.

Johnson AB, Lambert JD. The Caudal ParaHox gene is required for hindgut development in the mollusc Tritia (aka Ilyanassa). Developmental Biology. 2021 Feb 1;470:1-9.

Kleppe L, Fjelldal PG, Andersson E, Hansen T, Sanden M, Bruvik A, Skaftnesmo KO, Furmanek T, Kjærner-Semb E, Crespo D, Flavell S. Full production cycle performance of gene-edited, sterile Atlantic salmon-growth, smoltification, welfare indicators and fillet composition. Aquaculture. 2022 Nov 15;560:738456.

Kniprath E. Ontogeny of the molluscan shell field: a review. Zoologica Scripta. 1981 Jan;10(1):61-79.

Knirr S, Azpiazu N, Frasch M. The role of the NK-homeobox gene slouch (S59) in somatic muscle patterning. Development. 1999 Oct 15;126(20):4525-35.

Lartillot N, Lespinet O, Vervoort M, Adoutte A. Expression pattern of Brachyury in the mollusc Patella vulgata suggests a conserved role in the establishment of the AP axis in Bilateria.

Latham KE, Schultz RM. Embryonic genome activation. Front Biosci. 2001 Jun 1;6(D748-D759):10-2741.

Lebedeva LA, Yakovlev KV, Kozlov EN, Schedl P, Deshpande G, Shidlovskii YV. Transcriptional quiescence in primordial germ cells. Critical reviews in biochemistry and molecular biology. 2018 Nov 2;53(6):579-95.

Le Gouar M, Lartillot N, Adoutte A, Vervoort M. The expression of a caudal homologue in a mollusc, Patella vulgata. Gene expression patterns. 2003 Mar 1;3(1):35-7.

Liu G, Huan P, Liu B. A SoxC gene related to larval shell development and co-expression analysis of different shell formation genes in early larvae of oyster. Development Genes and Evolution. 2017 Jun;227:181-8.

Liu G, Huan P, Liu B. Identification of three cell populations from the shell gland of a bivalve mollusc. Development genes and evolution. 2020 Jan;230(1):39-45.

Liu Z, Zhang Y, Zhou Z, Zong Y, Zheng Y, Liu C, Kong N, Gao Q, Wang L, Song L. Metabolomic and transcriptomic profiling reveals the alteration of energy metabolism in oyster larvae during initial shell formation and under experimental ocean acidification. Scientific Reports. 2020 Apr 9;10(1):6111.

Lopez-Anido RN, Batzel GO, Ramirez G, Wang Y, Neal S, Lesoway MP, Goodheart JA, Lyons DC. The adult shell matrix protein repertoire of the marine snail Crepidula is dominated by conserved genes that are also expressed in larvae. BMC Ecology and Evolution. 2024 Sep 14;24(1):120.

Lyons DC, Perry KJ, Henry JQ. Spiralian gastrulation: germ layer formation, morphogenesis, and fate of the blastopore in the slipper snail Crepidula fornicata. EvoDevo. 2015 Dec;6:1-34.

Marie B, Joubert C, Tayalé A, Zanella-Cléon I, Belliard C, Piquemal D, Cochennec-Laureau N, Marin F, Gueguen Y, Montagnani C. Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. Proceedings of the National Academy of Sciences. 2012 Dec 18;109(51):20986-91.

Marin F, Corstjens P, de Gaulejac B, de Vrind-De Jong E, Westbroek P. Mucins and molluscan calcification: molecular characterization of mucoperlin, a novel mucin-like protein from the nacreous shell layer of the fan mussel Pinna nobilis (Bivalvia, Pteriomorphia). Journal of Biological Chemistry. 2000 Jul 7;275(27):20667-75.

Massri AJ, Greenstreet L, Afanassiev A, Berrio A, Wray GA, Schiebinger G, McClay DR. Developmental single-cell transcriptomics in the Lytechinus variegatus sea urchin embryo. Development. 2021 Oct 1;148(19):dev198614.

Min Y, Li Q, Yu H. Characterization of larval shell formation and CgPOU2F1, CgSox5, and CgPax6 gene expression during shell morphogenesis in Crassostrea gigas. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 2023 Jan 1;263:110783.

Miramón-Puértolas P, Pascual-Carreras E, Steinmetz PR. A population of Vasa2 and Piwi1 expressing cells generates germ cells and neurons in a sea anemone. Nature Communications. 2024 Oct 10;15(1):8765.

Moor B. The Mollusca: Development, Vol. 3. Verdonk NH, Van Den Biggelaar JAM, Tompa AS, editors. New York: Academic Press; 1983.

Nederbragt AJ, te Welscher P, van den Driesche S, van Loon AE, Dictus WJ. Novel and conserved roles for orthodenticle/otx and orthopedia/otp orthologs in the gastropod mollusc Patella vulgata. Development genes and evolution. 2002 Aug;212:330-7.

Nell JA. Farming triploid oysters. Aquaculture. 2002 Jul 31;210(1-4):69-88.

Nakamura A, Shirae-Kurabayashi M, Hanyu-Nakamura K. Repression of early zygotic transcription in the germline. Current opinion in cell biology. 2010 Dec 1;22(6):709-14.

Nikishchenko VE, Dyachuk VA. Comparison of neurogenesis in bivalves with different types of development. Scientific Reports. 2024 Aug 22;14(1):19495.

Obata M, Komaru A. The mechanisms of primordial germ cell determination during embryogenesis in molluscan species. Invertebrate Survival Journal. 2012 Dec 20;9(2):223-9.

Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, Waterston RH. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science. 2019 Sep 20;365(6459):eaax1971.

Paganos P, Voronov D, Musser JM, Arendt D, Arnone MI. Single-cell RNA sequencing of the Strongylocentrotus purpuratus larva reveals the blueprint of major cell types and nervous system of a non-chordate deuterostome. Elife. 2021 Nov 25;10:e70416.

Parker LM, Ross PM, O'Connor WA, Pörtner HO, Scanes E, Wright JM. Predicting the response of molluscs to the impact of ocean acidification. Biology. 2013 Apr 2;2(2):651-92.

Pérez-Montero S, Carbonell A, Azorín F. Germline-specific H1 variants: the "sexy" linker histones. Chromosoma. 2016 Mar;125(1):1-3.

Perry KJ, Lyons DC, Truchado‐Garcia M, Fischer AH, Helfrich LW, Johansson KB, Diamond JC, Grande C, Henry JQ. Deployment of regulatory genes during gastrulation and germ layer

specification in a model spiralian mollusc Crepidula. Developmental Dynamics. 2015 Oct;244(10):1215-48.

Piferrer, F., Beaumont, A., Falguière, J.C., Flajšhans, M., Haffray, P. and Colombo, L., 2009. Polyploid fish and shellfish: production, biology and applications to aquaculture for performance improvement and genetic containment. *Aquaculture*, *293*(3-4), pp.125-156.

Piovani L, Leite DJ, Yañez Guerra LA, Simpson F, Musser JM, Salvador-Martínez I, Marlétaz F, Jékely G, Telford MJ. Single-cell atlases of two lophotrochozoan larvae highlight their complex evolutionary histories. Science Advances. 2023 Aug 2;9(31):eadg6034.

Popov VM, Wu K, Zhou J, Powell MJ, Mardon G, Wang C, Pestell RG. The Dachshund gene in development and hormone-responsive tumorigenesis. Trends in Endocrinology & Metabolism. 2010 Jan 1;21(1):41-9.

Prezant RS. Form, function and phylogeny of bivalve mucins. Hong Kong: Hong Kong University Press.

Redl E, Scherholz M, Wollesen T, Todt C, Wanninger A. Expression of six3 and otx in Solenogastres (Mollusca) supports an ancestral role in bilaterian anterior‐posterior axis patterning. Evolution & Development. 2018 Jan;20(1):17-28.

Robledo JA, Yadavalli R, Allam B, Espinosa EP, Gerdol M, Greco S, Stevick RJ, Gómez-Chiarri M, Zhang Y, Heil CA, Tracy AN. From the raw bar to the bench: Bivalves as models for human health. Developmental & Comparative Immunology. 2019 Mar 1;92:260-82.

Sachslehner A, Zieger E, Calcino A, Wanninger A. HES and Mox genes are expressed during early mesoderm formation in a mollusk with putative ancestral features. Scientific Reports. 2021 Sep 9;11(1):18030.

Sakaguchi S, Mizuno S, Okochi Y, Tanegashima C, Nishimura O, Uemura T, Kadota M, Naoki H, Kondo T. Single-cell transcriptome atlas of Drosophila gastrula 2.0. Cell Reports. 2023 Jul 25;42(7).

Salamanca-Díaz DA, Schulreich SM, Cole AG, Wanninger A. Single-cell RNA sequencing atlas from a bivalve larva enhances classical cell lineage studies. Frontiers in Ecology and Evolution. 2022 Jan 26;9:783984.

Salvi D, Macali A, Mariottini P. Molecular phylogenetics and systematics of the bivalve family Ostreidae based on rRNA sequence-structure models and multilocus species tree. PloS one. 2014 Sep 24;9(9):e108696.

Satoh N, Hisata K, Foster S, Morita S, Nishitsuji K, Oulhen N, Tominaga H, Wessel GM. A single-cell RNA-seq analysis of Brachyury-expressing cell clusters suggests a morphogenesis-associated signal center of oral ectoderm in sea urchin embryos. Developmental Biology. 2022 Mar 1;483:128-42.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY. Fiji: an open-source platform for biological-image analysis. Nature methods. 2012 Jul;9(7):676-82.

Schulreich SM, Salamanca-Díaz DA, Zieger E, Calcino AD, Wanninger A. A mosaic of conserved and novel modes of gene expression and morphogenesis in mesoderm and muscle formation of a larval bivalve. Organisms Diversity & Evolution. 2022 Dec;22(4):893-913.

Sladitschek HL, Fiuza UM, Pavlinic D, Benes V, Hufnagel L, Neveu PA. MorphoSeq: full single-cell transcriptome dynamics up to gastrulation in a chordate. Cell. 2020 May 14;181(4):922-35.

Strathmann MF. Reproduction and development of marine invertebrates of the northern Pacific coast: data and methods for the study of eggs, embryos, and larvae. University of Washington Press; 1987.

Sun X, Li L, Wu B, Ge J, Zheng Y, Yu T, Zhou L, Zhang T, Yang A, Liu Z. Cell type diversity in scallop adductor muscles revealed by single-cell RNA-Seq. Genomics. 2021 Nov 1;113(6):3582-98.

Suzuki M, Nagasawa H. Mollusk shell structures and their formation mechanism. Canadian Journal of Zoology. 2013;91(6):349-66.

Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. Development. 2009 Sep 15;136(18):3033-42.

Teaniniuraitemoana V, Huvet A, Levy P, Klopp C, Lhuillier E, Gaertner-Mazouni N, Gueguen Y, Le Moullac G. Gonad transcriptome analysis of pearl oyster Pinctada margaritifera: identification of potential sex differentiation and sex determining genes. BMC genomics. 2014 Dec;15:1-20.

Thunberg CP. Teckning och beskrifning pa en stor Ostronsort ifran Japan. K. Svenska Vetenskapsakademien, Handlinger. (No Title). 1793;14:140-2.

Tong Y, Zhang Y, Huang J, Xiao S, Zhang Y, Li J, Chen J, Yu Z. Transcriptomics analysis of Crassostrea hongkongensis for the discovery of reproduction-related genes. PloS one. 2015 Aug 10;10(8):e0134280.

Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature biotechnology. 2014 Apr;32(4):381-6.

Wadsworth P, Casas S, La Peyre J, Walton W. Elevated mortalities of triploid eastern oysters cultured off-bottom in northern Gulf of Mexico. Aquaculture. 2019 Apr 30;505:363-73.

Wargelius A, Leininger S, Skaftnesmo KO, Kleppe L, Andersson E, Taranger GL, Schulz RW, Edvardsen RB. Dnd knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. Scientific reports. 2016 Feb 18;6(1):21284.

Wang H, Yu H, Li Q. Integrative analysis of single-nucleus RNA-seq and bulk RNA-seq reveals germline cells development dynamics and niches in the Pacific oyster gonad. Iscience. 2024 Apr 19;27(4).

Weidinger G, Stebler J, Slanchev K, Dumstrei K, Wise C, Lovell-Badge R, Thisse C, Thisse B, Raz E. Dead end, a novel vertebrate germ plasm component, is required for zebrafish primordial germ cell migration and survival. Current Biology. 2003 Aug 19;13(16):1429-34.

Wong TT, Zohar Y. Production of reproductively sterile fish by a non-transgenic gene silencing technology. Scientific Reports. 2015 Oct 29;5(1):15822.

Voronezhskaya EE, Khabarova MY, Nezlin LP. Apical sensory neurones mediate developmental retardation induced by conspecific environmental stimuli in freshwater pulmonate snails.

Xu R, Li Q, Yu H, Kong L. Oocyte maturation and origin of the germline as revealed by the expression of Nanos-like in the Pacific oyster Crassostrea gigas. Gene. 2018 Jul 15;663:41-50.

Xu L, Zhao M, Zohar Y, Wong TT. Induction of Reproductive Sterility in Coho Salmon (Oncorhynchus kisutch) by an Immersion-Based Gene Silencing Technology. Journal of Marine Science and Engineering. 2023 Nov 21;11(12):2208.

Xu L, Zhao M, Ryu JH, Hayman ES, Fairgrieve WT, Zohar Y, Luckenbach JA, Wong TT. Reproductive sterility in aquaculture: A review of induction methods and an emerging approach with application to Pacific Northwest finfish species. Reviews in Aquaculture. 2023 Jan;15(1):220-41.

Xu L, Small JM, Hood SM, Zhao M, Plough LV, Wong TT. Morpholino oligomer delivery via bath immersion for use in reverse genetic studies on the early development of eastern oysters (Crassostrea virginica). Aquaculture. 2025 Feb 5:742261.

Yu H, Li H, Li Q. Molecular characterization and expression profiles of myosin essential light chain gene in the Pacific oyster Crassostrea gigas. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 2017 Nov 1;213:1-7.

Yurchenko OV, Savelieva AV, Kolotuchina NK, Voronezhskaya EE, Dyachuk VA. Peripheral sensory neurons govern development of the nervous system in bivalve larvae. EvoDevo. 2019 Dec;10:1-9.

Zeidler MP, Perrimon N, Strutt DI. The four-jointed gene is required in the Drosophila eye for ommatidial polarity specification. Current Biology. 1999 Dec 2;9(23):1363-72.

Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT. Massively parallel digital transcriptional profiling of single cells. Nature communications. 2017 Jan 16;8(1):14049.

Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, Leonhardt H, Heyn H, Hellmann I, Enard W. Comparative analysis of single-cell RNA sequencing methods. Molecular cell. 2017 Feb 16;65(4):631-43.