

RESEARCH ARTICLE

SPECIAL ISSUE
UNCOVERING DEVELOPMENTAL DIVERSITY

Contrasting the development of larval and adult body plans during the evolution of biphasic lifecycles in sea urchins

Brennan D. McDonald^{1,*}, Abdull J. Massri¹, Alejandro Berrio¹, Maria Byrne², David R. McClay¹ and Gregory A. Wray^{1,‡}

ABSTRACT

Biphasic lifecycles are widespread among animals, but little is known about how the developmental transition between larvae and adults is regulated. Sea urchins are a unique system for studying this phenomenon because of the stark differences between their bilateral larval and pentaradial adult body plans. Here, we use single-cell RNA sequencing to analyze the development of *Heliocidaris erythrogramma* (*He*), a sea urchin species with an accelerated, non-feeding mode of larval development. The sequencing time course extends from embryogenesis to roughly a day before the onset of metamorphosis in *He* larvae, which is a period that has not been covered by previous datasets. We find that the non-feeding developmental strategy of *He* is associated with several changes in the specification of larval cell types compared to sea urchins with feeding larvae, such as the loss of a larva-specific skeletal cell population. Furthermore, the development of the larval and adult body plans in sea urchins may utilize largely different sets of regulatory genes. These findings lay the groundwork for extending existing developmental gene regulatory networks to cover additional stages of biphasic lifecycles.

KEY WORDS: Sea urchins, Biphasic lifecycles, Body plans, Single-cell RNA sequencing, Gene regulatory networks

INTRODUCTION

Most metazoan phyla contain species with biphasic lifecycles, also known as indirect development (Formery and Lowe, 2023; Moran, 1994; Nielsen, 2009; Strathmann, 1985; Thorson, 1950). Species that use this life history strategy first develop from embryos into a larval body plan, and then undergo metamorphosis to transition into a distinct adult body plan. Little is known about how the transition between the larval and adult body plans is regulated at the molecular level. Larval development of phyla such as the hemichordates and annelids is modeled to be controlled by the bilaterian anterior

patterning network, and the delayed activation of the trunk patterning network is proposed to initiate adult development (Formery and Lowe, 2023; Gonzalez et al., 2017; Lacalli, 2005; Martin-Zamora et al., 2023).

This model does not fit with sea urchin development. These animals have highly derived body plans owing to the pentaradial symmetry of their adults (Smith, 2008). They evolved from a bilaterally symmetrical ancestor, and their larvae retain this symmetry form (Peterson et al., 2000; Sumrall and Wray, 2007). Thus, sea urchins undergo a developmental transition between larvae with bilateral symmetry and juveniles with pentaradial symmetry when they metamorphose. A recent study found that the adults of sea stars, a close relative of sea urchins, do not employ the bilaterian trunk regulatory network to pattern the ectoderm (Formery et al., 2023), unlike hemichordates and annelids (Gonzalez et al., 2017; Martin-Zamora et al., 2023). This suggests that echinoderms may use mechanisms different from other bilaterians to control the transition between the larval and adult body plans. To obtain a better understanding of the relationship between larval and adult pattern regulation, a needed starting point is to see whether the well-characterized gene regulatory network (GRN) that controls embryonic and larval development in sea urchins (Davidson et al., 2002; Peter and Davidson, 2011; Rafiq et al., 2012; Sethi et al., 2009; Su et al., 2009) also plays a role in patterning the adult.

Late larval development and adult morphogenesis in most sea urchin species is difficult to study because development of the adult rudiment takes place over several weeks during the larval phase. This handicap is much reduced with selection of a model species that rapidly transits from the larval to adult body plan. The Australian sea urchin *Heliocidaris erythrogramma* (*He*) employs a lecithotrophic developmental strategy and enters metamorphosis about 4 days post-fertilization (Raff, 1992; Williams and Anderson, 1975). This life history strategy is enabled, not by feeding on plankton, but rather by providing sufficient energy content in its eggs to support the juvenile until it can feed on its own. In contrast, sea urchins with planktotrophic larvae (planktotrophs) require feeding on phytoplankton for an extended time to acquire the energy supply necessary to build the rudiment before metamorphosis (McEdward and Janies, 1997; Strathmann, 1985). Although sometimes termed a ‘direct’ developer, *He* has retained many characteristics of the sea urchin biphasic lifecycle and constructs a partial, bilaterally symmetrical larval body (Emlet, 1995). Thus, *He* can be used as a model for studying the transition between the larval and adult bodies within a manageable experimental timeframe.

He is also a useful model for understanding how biphasic lifecycles can evolve over short periods of time. Evolutionary transitions between different modes of larval development are common across marine invertebrate taxa, despite the extensive

¹Department of Biology, Duke University, Durham, NC 27708, USA. ²School of Life and Environmental Sciences, A11, University of Sydney, Sydney, NSW, 2006, Australia.

*Present address: Department of Biology, Stanford University, Stanford, CA 94305, USA.

‡Author for correspondence (gwray@duke.edu)

 B.D.M., 0009-0007-2321-0149; A.J.M., 0000-0003-1314-6650; A.B., 0000-0003-1259-2181; M.B., 0000-0002-8902-9808; D.R.M., 0000-0001-8824-2183; G.A.W., 0000-0001-5634-5081

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Handling Editor: Karen Sears

Received 30 April 2024; Accepted 16 September 2024

changes that often need to occur in embryonic development and larval morphology (Strathmann, 1978, 1985; Wray, 1996; Wray and Raff, 1991). Identifying the regulatory mechanisms that control the shift between developmental modes has been difficult, because species with different modes are often distantly related. The distinct advantage of studying *He* is that it is closely related to *H. tuberculata* (*Ht*), a sea urchin with the planktotrophic mode of development (Wray, 2022). Phylogenetic evidence suggests that *He* is separated by only about 5 million years of evolutionary divergence from *Ht* (Hart et al., 2011; Smith et al., 1990; Zigler et al., 2003). This, along with strong evidence that planktotrophic development is ancestral for sea urchins (McEdward and Miner, 2001; Wray, 1996), implies that *He*'s lecithotrophic developmental process is a derived evolutionary feature (Raff, 1992; Smith et al., 1990). At the same time, *He* and *Ht* adults have highly similar morphologies (Byrne and O'Hara, 2017), making them a good model for determining how different larval developmental processes can lead to the same adult body plan.

Here, we use *He* as a model for studying the evolution and development of biphasic lifecycles from two perspectives. First, we advance our understanding of how larval development changed during the evolution of lecithotrophy in *He*. Second, we focus on the regulation of larval and adult body plan development in *He* and compare this to the same processes in species with planktotrophic development. To address these objectives, we performed single-cell RNA sequencing (scRNA-seq) for *He* embryos and larvae at 12 time points from 6 to 60 h post-fertilization (hpf). The final atlas combines scRNA-seq data from five new time points of larval stages, along with seven embryonic time points that are the focus of a complementary paper (Massri et al., 2024 preprint). Temporal scRNA-seq atlases have been generated for several sea urchin species with planktotrophic development (Foster et al., 2020; Massri et al., 2021; Paganos et al., 2021), but these atlases stopped well before metamorphosis. By contrast, the final time point in the *He* atlas is close to the onset of metamorphosis, which is approximately 96 hpf in *He* larvae raised at 23°C. At 60 hpf, *He* larvae possess a highly developed adult rudiment, including partially formed primary podia and the early water vascular system (Fig. 1) (Minsuk and Raff, 2002). We describe several cell types in *He* with unique transcriptional trajectories compared to those in species with planktotrophic development. In particular, we focus on the loss of a larva-specific skeletogenic cell population, the emergence of an undifferentiated cell population, and the accelerated diversification of neural subpopulations. Furthermore, we compare gene expression patterns between early and late stages of the time course and find that genes from the sea urchin embryonic GRN may play a reduced role in controlling adult morphogenesis. We then leverage the presence of cells from adult tissues in our dataset to generate a list of candidate transcription factors (TFs) that may pattern the adult rudiment. Altogether, the extensive temporal coverage of our dataset makes it a valuable resource for addressing questions about larval and adult morphogenesis in sea urchins and perhaps other species with biphasic lifecycles.

RESULTS

A temporal scRNA-seq atlas of development in *He* reveals the diversity of cell types in larval sea urchins

Here, we expand upon a temporal analysis of embryonic development in *He* (Massri et al., 2024 preprint) by adding scRNA-seq data for five additional time points of larval stages, resulting in insights into formation of the adult rudiment. The final atlas contains 12 time points from 6 to 60 hpf and covers

development in *He* from cleavage to the mid-larva stage, allowing us to infer cell lineage trajectories between cell types in the embryo and larva (Fig. 1). Analysis of the dataset began by merging cell-level gene expression data from each of the time points into the same object using the scRNA-seq analysis tool Seurat (Hao et al., 2021). After filtering and normalization, there were 49,896 cells across all time points in the dataset. Cells in each time point had an average of 500-1500 genes per cell and 1000-3000 unique molecular identifiers per cell (Fig. S1). We then used the uniform manifold approximation projection (UMAP) algorithm to project the cells into two-dimensional space for visualization (Fig. 2A) and performed Louvain-based clustering to group cells with similar transcriptional identities. This returned 57 distinct cell clusters, to which we assigned 18 cell-type identities (Fig. 2B) based on known marker genes from studies in other sea urchin species (Fig. S2; Table S1). We also performed hybridization chain reaction *in situ* (HCR) to verify the spatial expression patterns of prominent marker genes from the scRNA-seq atlas. Using this information, we were able to recover in the *He* dataset most of the cell types that are known to be present in the larvae of sea urchins with planktotrophic development (Massri et al., 2021). For example, we annotated *pks1*-expressing cell clusters as pigment cells (Calestani et al., 2003), which aligned with the HCR expression of *pks1* in cells scattered throughout the ectoderm in 56 hpf *He* larvae (Fig. 2C). The HCR expression patterns of *onecut* and *foxq2* (Fig. 2D,E) in the ciliary band and animal pole domain of 56 hpf larvae, respectively, also aligned with known expression patterns in species with planktotrophic development (Poustka et al., 2004; Yaguchi et al., 2008).

Not all clusters aligned with the profiles identified in previous scRNA-seq studies. Most of these contained cells from later time points, suggesting that the previous scRNA-seq atlases failed to capture these late larval cell types. We used HCRs to map these clusters onto discrete tissues in 56 hpf larvae. For example, one transcript of *egfi*, a gene that encodes an extracellular matrix protein (Bisgrove et al., 1991), had strong expression in the rudiment ectoderm (Fig. 2F). We thus annotated the scRNA-seq cluster with strong *egfi* expression as 'Rudiment Ectoderm (Verified)'. To our knowledge, this is the first time that scRNA-seq has been used to capture the transcriptional profile of cells in sea urchin rudiment structures. *dynlc2*, a gene that encodes a protein component of the dynein light chain (Telmer et al., 2024), was expressed in the tips of the primary podia (Fig. 2G), which may be associated with the intracellular vesicle trafficking or sensory cilia that are known to be common in this tissue (Burke, 1980; Byrne et al., 2018). We grouped this with the other rudiment ectoderm clusters. As a final example, *hox7*, a member of the sea urchin Hox gene complex (Popodi et al., 1996), was expressed in the posterior ectoderm as well as part of the adult rudiment (Fig. 2H) that is transcriptionally distinct from other ectodermal domains (annotated as 'Posterior Ectoderm'). Overall, our scRNA-seq atlas of *He* development is a valuable resource for understanding the transcriptional profiles of tissues in late sea urchin larvae.

Larval cell types in *He* follow novel developmental trajectories compared to those in species with planktotrophic development

Several *He* larval cell types have different developmental trajectories compared to what was observed in single-cell atlases of planktotrophs (Foster et al., 2020; Massri et al., 2021; Paganos et al., 2021). A particularly noticeable difference was with the skeletogenic cell (SKC) cluster (Fig. 3A). This cluster contained cells that expressed the TF *alk1* (Ettensohn et al., 2003) and the

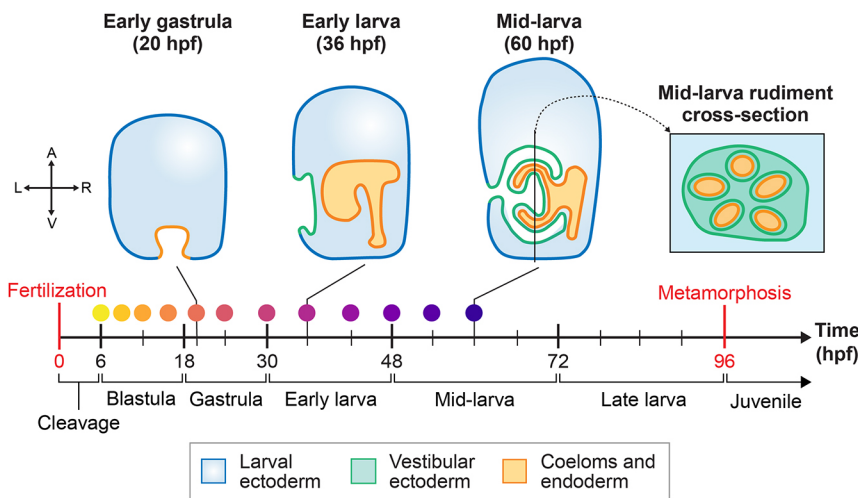


Fig. 1. Overview of larval development in *Heliocidaris erythrogramma*. Diagrams of several larval stages are shown at the top, which are vertical cross-sections through the left-right plane of the larva, except for the right-most diagram, which is a cross-section of the adult rudiment through the dorsal-ventral plane of the larva. Note that not all tissue types are shown. The timeline below the diagrams highlights the time intervals (in hours post-fertilization, hpf) corresponding to key stages in *He* development. The time intervals for each stage are based on raising *He* embryos and larvae at 23°C. Colored dots above the timeline correspond to the scRNA-seq samples collected for this study (see Fig. 2A). The coordinate plane on the left represents the positional axes of the embryo depicted in this diagram: A, animal; L, left; R, right; V, vegetal.

downstream differentiation genes *msp130*, *sm37* and *sm50*, which are commonly used markers of SKCs in other scRNA-seq studies (Fig. S3; Table S1) (Massri et al., 2021; Paganos et al., 2021). In the development of planktotrophs, there are two populations of SKCs, one for building the larval skeleton that is derived from the primary mesenchyme cells and a later population for building the adult skeleton that differentiates from the secondary mesenchyme cells (SMCs) (Tate et al., 2024; Yajima, 2007). A previous study in *He* found that the SKCs are specified much later in larval development (Davidson et al., 2022). Given that larval development in *He* is much shorter than in species with planktotrophic development, we wondered whether the late emergence of SKCs may be an indication that the two populations of SKCs are specified differently. To assess this, we first examined the expression of larval skeletogenic genes. Gao and Davidson (2008) established that *tbr*, *tel*, *foxo* and *foxb* are expressed in larval, but not adult, SKCs. In the *He* scRNA-seq atlas, few cells in the skeletogenic lineage showed co-expression between *alx1* and *tbr* (2/106=1.8% of *alx1*⁺ cells in the skeletogenic lineage), *alx1* and *foxb* (0/106=0% of *alx1*⁺ cells), and *alx1* and *foxo* (7/106=6.6% of *alx1*⁺ cells) and only showed moderate co-expression between *alx1* and *tel* (24/106=22.6% of *alx1*⁺ cells) (Fig. 3C-F). We then compared these expression patterns with those in *Lytechinus variegatus* (*Lv*), a sea urchin with planktotrophic development similar to *Ht* and for which there is already a published scRNA-seq atlas (Massri et al., 2021). Co-expression of these four gene pairs was more widespread in *Lv* than in *He* (Fig. 3B,C'-F'). Nearly all SKCs in *Lv* co-expressed *alx1* and *tbr* (1961/2211=88.9% of *alx1*⁺ cells), and significant fractions of SKCs co-expressed *alx1* and *foxb* (636/2211=28.8% of *alx1*⁺ cells), *alx1* and *foxo* (517/2211=23.4% of *alx1*⁺ cells), and *alx1* and *tel* (913/2211=41.3% of *alx1*⁺ cells). These results expand on the finding from Massri et al. (2024) that SKCs in *He* no longer co-express *alx1* and *foxb* and suggest that a subset of the interactions in the larval skeletogenic gene regulatory network has been lost in *He*.

We then examined whether the larval SKCs in *He* may have acquired characteristics of the adult SKC population in species with planktotrophic development. We focused on the expression of *scl*, *ese* and *gatac*, which are all markers for oral SMCs in planktotrophs (Materna et al., 2013). SKCs in the *He* scRNA-seq dataset showed significant co-expression between *alx1* and *scl* (93/106=87.7% of *alx1*⁺ cells) and moderate co-expression between *alx1* and *gatac* (37/106=34.9% of *alx1*⁺ cells) and *alx1* and *ese* (19/106=17.9% of

alx1⁺ cells) (Fig. 4A-C). Double HCRs for *alx1* and *scl* in early (36 hpf) *He* larvae provided additional verification that this co-expression event was taking place (Fig. 4D). In *Lv*, co-expression between *alx1* and oral SMC markers was more limited (Fig. 4A'-C'). The difference was particularly stark for *alx1* and *scl*, as only 14.6% (323/2211) of *alx1*⁺ cells showed co-expression (Fig. 4A'). In addition, whereas co-expression between *alx1* and *scl* began as soon as *alx1* started being expressed in *He* (Fig. 4E), *scl* expression began later in *alx1*-expressing cells in *Lv* (Fig. 4F). There were also lower co-expression levels of *alx1* and *ese* and *alx1* and *gatac* in *Lv* (Fig. 4B',C'). Thus, it appears that SKCs in *He* have gained expression of a subset of genes known to pattern a different mesenchymal population in sea urchins with planktotrophic development.

Another newly identified larval cell population in the *He* scRNA-seq atlas was a persistent cluster of undifferentiated cells (Fig. 5A). Although many cells in this cluster were from the 6 hpf time point, the cluster contained cells from all time points except the last, 60 hpf (Fig. 5B). This cell population persisted throughout larval development, suggesting it is not an artifact of the sample preparation for a single time point. To get a better sense of the gene expression profiles of cells in this cluster, we identified the genes that were significantly enriched in this cluster compared to all others ($n=342$) and performed gene ontology (GO) over-representation analysis for genes in this list to identify the biological processes that they likely facilitate. None of the enriched genes in this cluster was a known marker gene for differentiated cell types in species with planktotrophic development (Table S1). Similarly, most of the enriched GO terms were for categories related to regulation of the cell cycle and the cytoskeleton (Fig. 5D), which are not unique to any differentiated cell type. The undifferentiated nature of cells in this cluster was further demonstrated by how it grouped closest to the early ectoderm cluster, which contained multipotent cells from the 6 hpf time point, in a phylogenetic tree of cell types (Fig. S4). We then implemented the Waddington-OT (optimal transport) trajectory inference algorithm using a previously published pipeline (Massri et al., 2021; Schiebinger et al., 2019) to track cell fate transitions in the *He* dataset. It appeared that cells beginning in this cluster transition into several other cell types over the 6-60 hpf period covered by the scRNA-seq atlas, especially the ciliary band and larval ectoderm (Fig. 5C). Taken together, these data indicate that the cells in this population start out with undifferentiated transcriptional profiles, but eventually differentiate

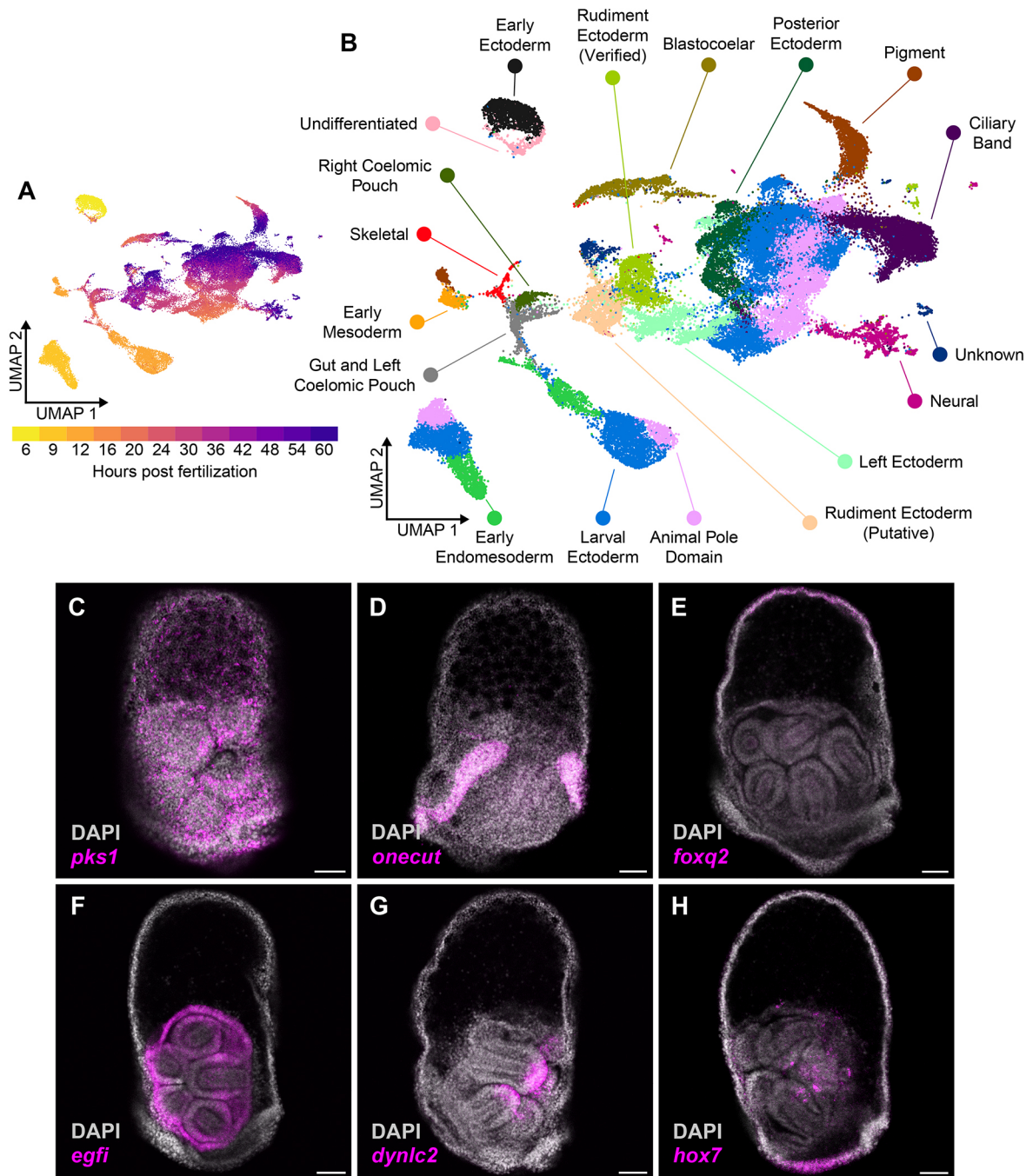


Fig. 2. Cell-type identification and trajectories in a temporal scRNA-seq atlas of development in *He*. (A) UMAP plot that projects the gene expression profiles of each cell into two-dimensional space, colored by sample time point. (B) UMAP that colors each cell by its cell type identity based on the expression of known cell-type marker genes. (C-H) HCR micrographs in 56 hpf *He* larvae for six marker genes that were used to identify cell types in the scRNA-seq atlas: *pks1*, a marker for pigment cells (C); *onecut*, a marker for the ciliary band (D); *foxq2*, a marker for the animal pole domain (E); *egfi*, expressed in the rudiment ectoderm (F); *dynlc2*, expressed at the tips of the primary podia (G); and *hox7*, expressed in the posterior ectoderm and inside the rudiment (H). Scale bars: 50 μ m.

into other cell fates during development. It is important to note that this cluster is unlikely to be the germline owing to the absence of co-expression between *nanos2*, *vasa* and *seawi*, which are known markers of this cell type in planktotrophs (Fig. 5E) (Juliano et al., 2006). The developmental identity of the germline remains unknown in *He*.

Finally, the *He* scRNA-seq atlas is a rich resource for understanding the larval nervous system of sea urchins. A

previous scRNA-seq study (Paganos et al., 2021) catalogued neural cell types in the early planktotrophic larvae of *Strongylocentrotus purpuratus*. The slower larval developmental process of this species meant that the nervous system was still in the early stages of development at the assayed time point (72 hpf). In contrast, the rapid larval development of *He* meant that the 60 hpf time point of the scRNA-seq atlas was likely to capture a wide array of differentiated neural cell types. To explore this further, we

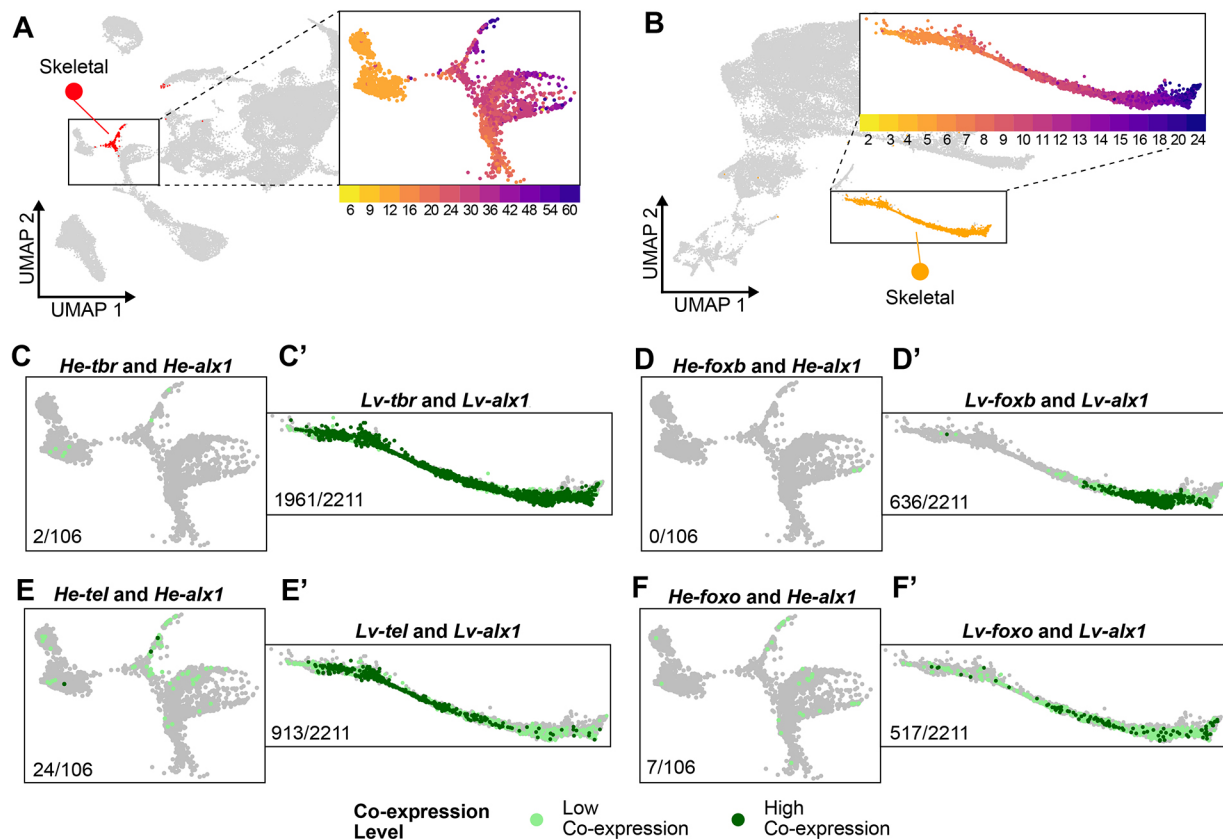


Fig. 3. *He* skeletal cells lack expression of genes known to be expressed in the larval SKCs of planktotrophs. (A) UMAP of *He* development highlighting the SKC cluster (red), with an inset of the SKC region of the UMAP colored by time point. (B) UMAP of *Lv* development (retrieved from Massri et al., 2021) highlighting the SKC cluster (orange), with an inset of the SKC region of the UMAP colored by time point. (C-F) Cells in the *He* scRNA-seq dataset (zoomed into the SKC region of the UMAP) showing co-expression of *alx1* and *tbr* (C), *alx1* and *foxb* (D), *alx1* and *tel* (E), and *alx1* and *foxo* (F). (C'-F') Cells in the *Lv* scRNA-seq dataset (zoomed into the SKC region of the UMAP) showing co-expression of *alx1* and *tbr* (C'), *alx1* and *foxb* (D'), *alx1* and *tel* (E'), and *alx1* and *foxo* (F'). Light green dots represent cells with low co-expression levels, and dark green dots represent cells with high co-expression levels. Numbers at the bottom left of each diagram indicate the number of both the low and high co-expressing cells for each gene pair out of the total number of *alx1*-expressing cells in the SKC lineage for each species.

isolated the scRNA-seq clusters expressing pan-neuronal genes (including *syn4* and *secretag*) and performed additional clustering to identify a finer range of neural cell types (Burke et al., 2006b). We then annotated each neural cell cluster based on the expression of marker genes identified in previous studies (Fig. S5; Table S2). In total, we identified 15 unique neural cell clusters, including both progenitor and terminally differentiated populations (Fig. 6A,B). Several of these clusters aligned with previously studied larval neural cell types in sea urchins. Neural progenitor cells from the earliest time points (12-20 hpf) showed expression of early anterior neuroectoderm (ANE) genes, such as *hbn* and *soxc*. Differentiated larval neural cell types appeared at later stages, including serotonergic neurons expressing *tph* and postoral and ciliary band neurons expressing *th*. However, several clusters did not match the expression profiles of known larval neural cell types. The most likely explanation for this was that these were neurons from the developing adult nervous system, given that many adult structures are already present in the rudiments of 60 hpf *He* larvae. We were only able to provide rough annotations for these clusters based on the limited research previously conducted on adult sea urchin nervous systems. For example, the expression of *isl*, *pax6* and *opsin2* in a late-stage cluster matched previous studies that profiled the light receptor neurons found in the podia of adult sea urchins (annotated as 'Podia Neurons') (Byrne et al., 2018; Ullrich-Lüter et al., 2011).

We next turned our attention to the regulation of the neural cell type diversification process in *He*. We compiled a list of 68 TFs known to be expressed in sea urchin neural cell types (Burke et al., 2006a) and selected the 30 TFs with the highest expression levels across all cells in the neural scRNA-seq cluster. We then used k-means clustering to group these genes based on similar expression levels in each of the 15 different neural cell clusters (Fig. 6C). The ten groups of TFs can be viewed as potential gene regulatory suites controlling the differentiation of neural cell types. Groups 1 and 2 contain TFs with high expression in both ANE and adult neural progenitor cells, such as *myt1*, *soxc* and *smadip*. Group 6, with *brn1/2/4*, *otp*, *onecut* and *z166*, had high expression in ciliary band and postoral neurons, aligning with studies that found a shared embryonic origin for these neural cell types (reviewed by McClay, 2022). This analysis also identified potential genes that may regulate the differentiation of the unidentified adult neural cell clusters, with the TFs in groups 9 and 10 showing high expression levels in these cells. We noticed that two of these genes, *hmx* and *lim1*, also had non-neural expression domains in adult rudiment clusters in the full scRNA-seq atlas. This finding was confirmed using HCR in 56 hpf *He* larvae. Both genes were expressed in ectodermal cells with similar morphology to neurons (Fig. 6D, top row) while simultaneously showing expression in rudiment tissues (Fig. 6D, bottom row). It appears that there is overlap between the patterning

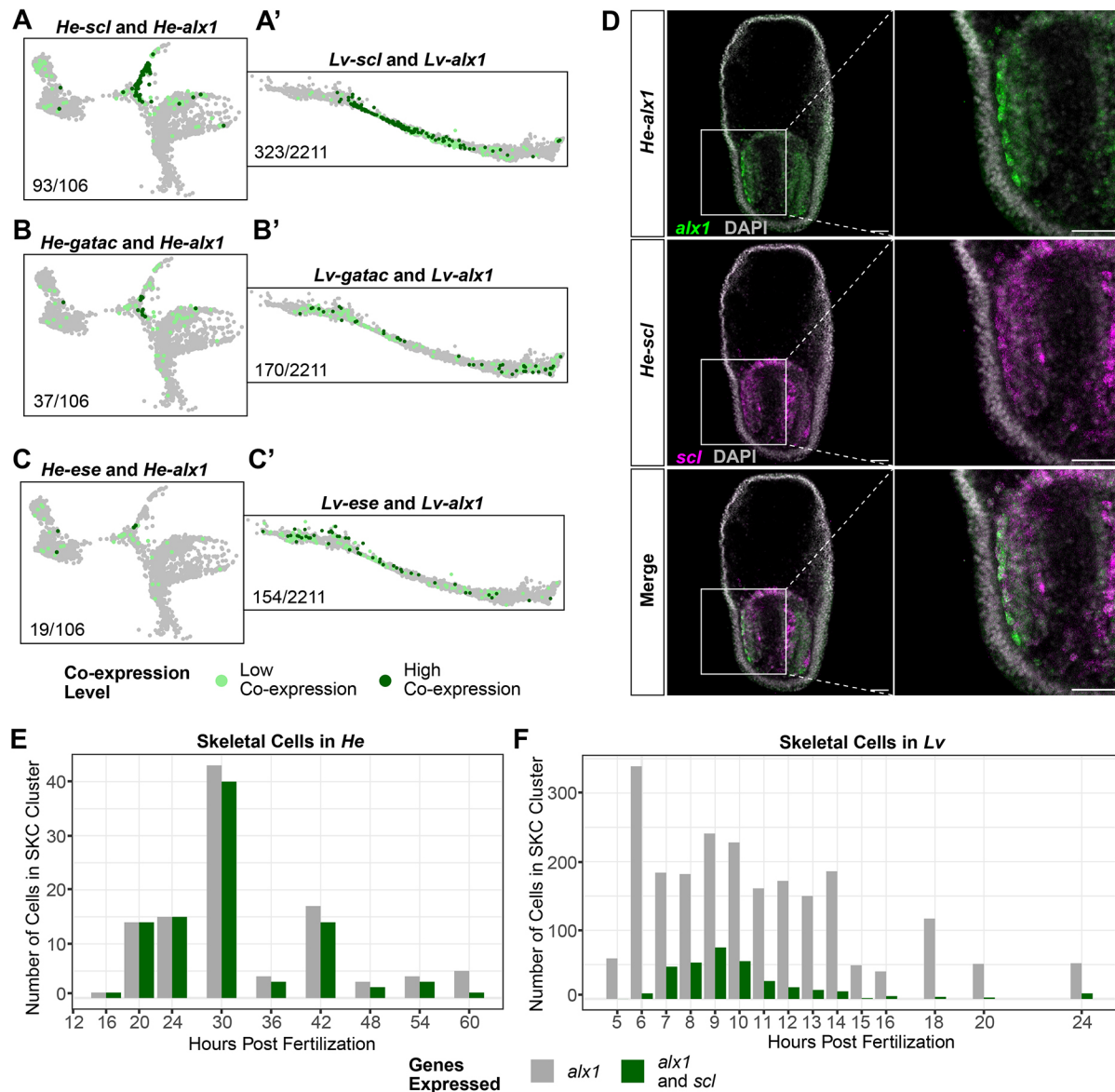


Fig. 4. *He* skeletal cells gained co-expression of genes known to be expressed in the secondary mesenchyme cells of planktotrophic sea urchins. (A-C) Cells in the *He* scRNA-seq dataset (zoomed into the SKC region of the UMAP) showing co-expression of *alx1* and *scl* (A), *alx1* and *gatac* (B), and *alx1* and *ese* (C). (A'-C') Cells in the *Lv* scRNA-seq dataset (zoomed into the SKC region of the UMAP) showing co-expression of *alx1* and *scl* (A'), *alx1* and *gatac* (B'), and *alx1* and *ese* (C'). Light green dots represent cells with low co-expression levels, and dark green dots represent cells with high co-expression levels. Numbers at the bottom left of each diagram indicate the number of both the low and high co-expressing cells for each gene pair out of the total number of *alx1*-expressing cells in the SKC lineage for each species. (D) HCR micrographs of *alx1* (green) and *scl* (magenta) in 36 hpf *He* larvae. Co-expression occurs in a large number of mesenchymal cells surrounding the vestigial gut of early larvae. (E,F) Charts showing the number of SKCs in the *He* (E) and *Lv* (F) scRNA-seq time courses showing expression of *alx1* (gray) and *alx1* and *scl* (green) at each sample time point. Scale bars: 50 μ m.

of the nervous system and the adult rudiment in *He* larvae, with similar genes potentially controlling neurogenesis in both larvae and adults.

Different gene regulatory suites appear to control the development of the larval and adult body plans

In organisms with biphasic lifecycles, the switch between the larval and adult body plans does not entirely occur during the brief timeframe of metamorphosis (Formery and Lowe, 2023). Rather, many adult tissues are constructed during larval development, as is the case with the sea urchin rudiment. We thus turned to the *He* scRNA-seq atlas to compare gene expression patterns between early, embryonic stages and later, larval stages when the rudiment is

emerging. We began by performing a pseudobulk analysis of the scRNA-seq dataset. We performed fuzzy c-means clustering to group genes with similar temporal expression profiles into nine clusters (Fig. 7A). Genes that belong to each cluster will have similar shapes to their gene expression profiles over time. Clusters 3, 5 and 8 formed a group that showed highest expression during the 6 and 9 hpf time points, which we labeled as 'High Early'. Clusters 1, 6 and 7 showed highest expression in the middle of the time course from 12 to 30 hpf, which we labeled as 'High Middle'. Finally, clusters 2, 4 and 9 peaked in expression at 36 hpf or later, which we labeled as 'High Late'. We then performed GO over-representation analysis to see whether certain biological processes were enriched for genes in some clusters compared to others (Fig. 7B). 'High

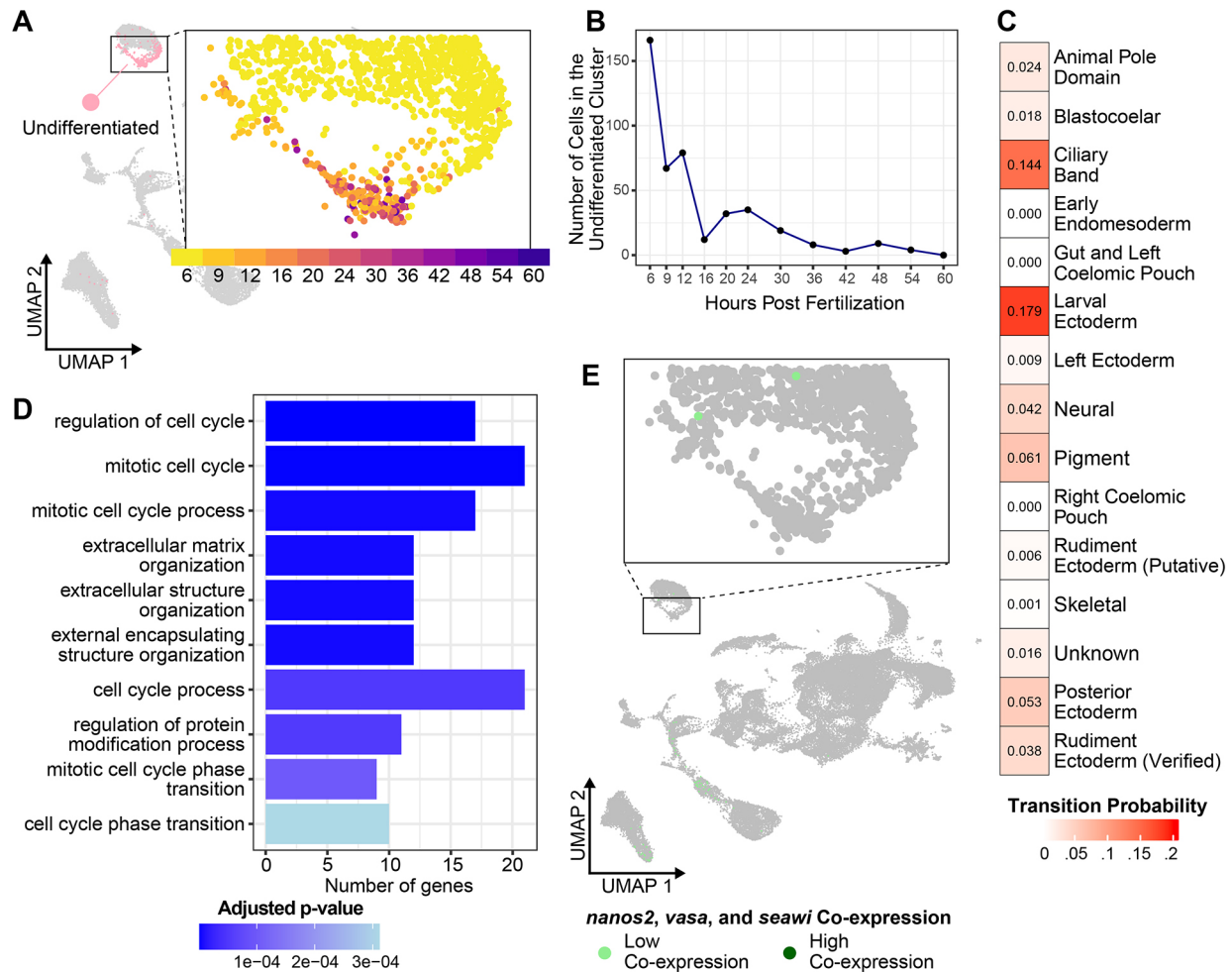


Fig. 5. A population of undifferentiated cells persists through the accelerated larval period of *He*. (A) UMAP of *He* development highlighting the undifferentiated cell cluster (pink), with an inset of the undifferentiated cell region of the UMAP colored by time point. (B) Plot showing the number of cells belonging to the undifferentiated cell cluster at each time point in the scRNA-seq time course. (C) Transition table showing the Waddington-OT trajectory inference results for the undifferentiated cell cluster. Transition probability values indicate the likelihood that cells beginning in the undifferentiated cell cluster transition into a different cell type identity. (D) Chart showing the top enriched GO terms in the undifferentiated cluster, colored by adjusted *P*-value. (E) UMAP of *He* development showing co-expression between *nanos2*, *seawi* and *vasa*, three of the main markers of the germline in planktotrophic sea urchins.

Early clusters were enriched for functions related to metabolism, cell cycle regulation, and organelle remodeling, which may be due to the rapid cell divisions that occur early in embryogenesis. Many of the enriched terms in the 'High Middle' clusters were for cell division-related processes, as well as transcription and translation. This may reflect widespread larval cell differentiation during late embryogenesis. Clusters in the 'High Late' category showed enrichment for nervous system functions, aligning with the rapid diversification of neural cells during later larval stages (Fig. 6). There was also enrichment for terms related to cell migration and movement in the 'High Late' category. Cell movement may be common in later developmental stages as larval cells are rearranged during the morphogenesis of adult tissues.

The gene function analysis suggested that a unique set of processes may be taking place during later stages of larval development in *He*. This prompted us to investigate whether the well-defined GRN for sea urchin embryonic development (Davidson et al., 2002; Peter and Davidson, 2011; Rafiq et al., 2012; Sethi et al., 2009; Su et al., 2009) plays a similarly important role in late larval and adult development. Many of the genes in the embryonic GRN are TFs and signaling molecules that are

commonly used in many developmental processes, so it would not be unexpected to see similar genes being used in both time frames. To assess this, we retrieved a curated list of embryonic GRN genes from the *He* genome used in previous analyses (Davidson et al., 2022) and classified whether the temporal expression profile of each gene fell into the 'High Early', 'High Middle' and 'High Late' categories obtained from the fuzzy c-means analysis. As a comparison, we also generated a list of all the potential DNA-binding TFs that are found in the *He* genome based on GO terms and a previously curated list (see Materials and Methods). The number of embryonic GRN genes and TFs that fell into each temporal category is shown in Fig. 8A. Note that many, but not all, embryonic GRN genes were TFs, so the numbers of GRN genes and TFs are not directly comparable. However, it was striking to see that the vast majority of GRN genes had expression profiles that peaked early on or in the middle of *He* development (100 with early or middle peaks versus 38 with late peaks). In comparison, there were roughly equal numbers of TFs that had expression peaks in each temporal category. It thus appears that genes in the embryonic GRN may play a reduced role in patterning the late larvae and adult rudiment of *He*. Of the few embryonic GRN genes that had late expression peaks,

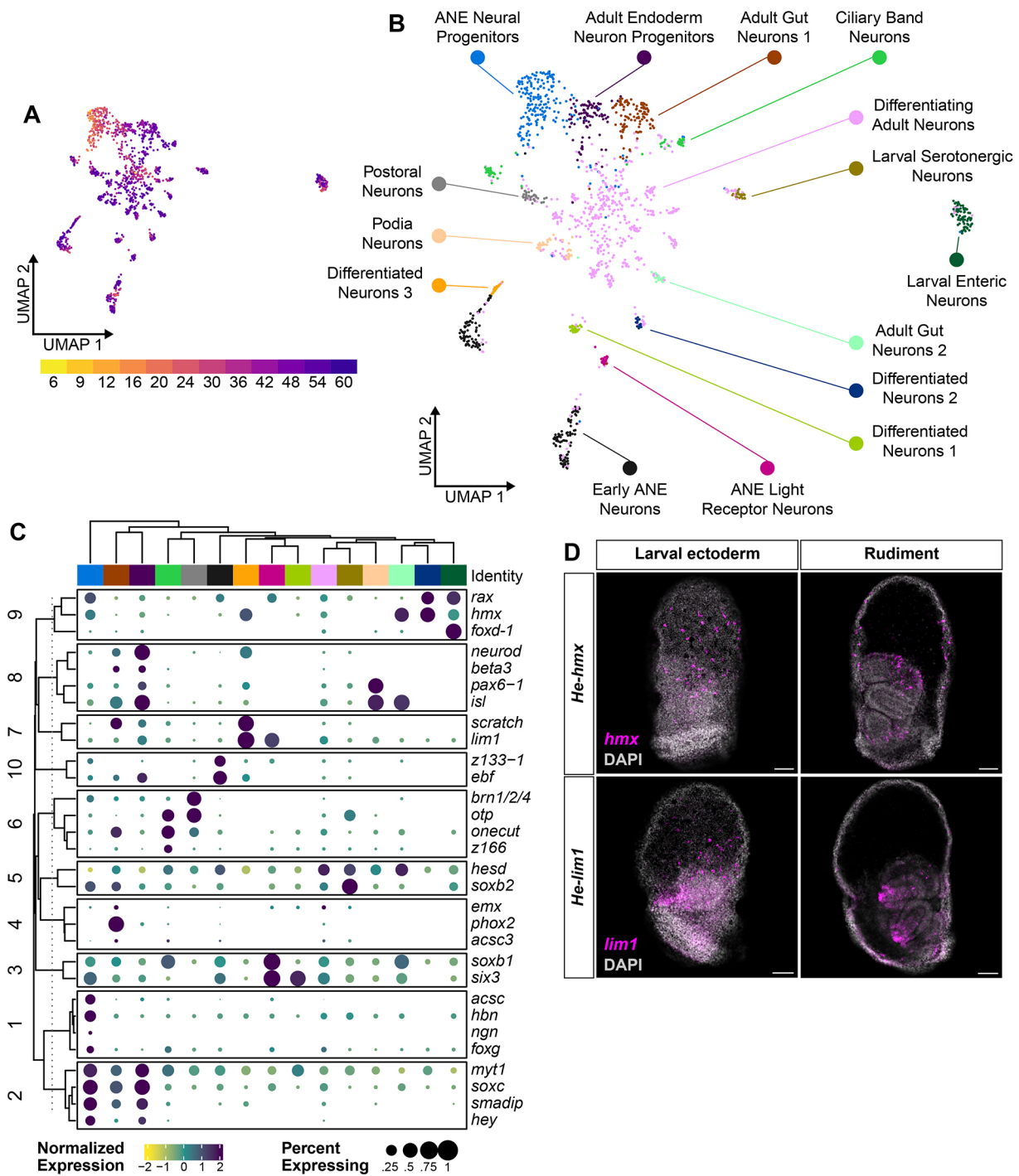


Fig. 6. *He* larvae retain ancestral sea urchin larval neural cell types as well as acquire the expression of several putative adult neural cell types.

(A) UMAP of the sub-clustered neural cells, colored by sample time point. (B) UMAP of the sub-clustered neural cells, colored by neural cell-type identities that were assigned based on expression of known larval and adult sea urchin neural patterning genes. Note that several clusters are labeled only as 'differentiated neurons' because they matched no known neural cell-type signature. (C) Clustered dot plot of the expression levels of 30 neural TFs in each of the *He* neural cell types. Clustering was performed using k-means clustering with $k=10$. (D) HCR micrographs in 56 hpf *He* larvae for *hmx* and *lim1*, two TFs that are expressed in neurons in *He*. The first column depicts expression in neurons in the larval ectoderm. The second column shows a more interior view of the larva, where these genes have expression in the adult rudiment. Scale bars: 50 μ m.

several have roles in patterning the larval nervous system in planktotrophs, such as *scratch* and *isl* (Fig. 8B) (Slota et al., 2019).

Because embryonic GRN genes appear to have lower expression levels at later stages, we hypothesized that adult rudiment development in *He* may be controlled by a new set of transcriptional regulators. We moved to generate a list of

candidate TFs that could play a regulatory role in this process. We returned to the scRNA-seq dataset and focused on the rudiment ectoderm and posterior ectoderm cell clusters, which are the primary cell types in the scRNA-seq atlas with verified expression patterns in the rudiment (Fig. 8D). Of the 234 TFs that had late expression peaks, we filtered for genes that had high expression in the

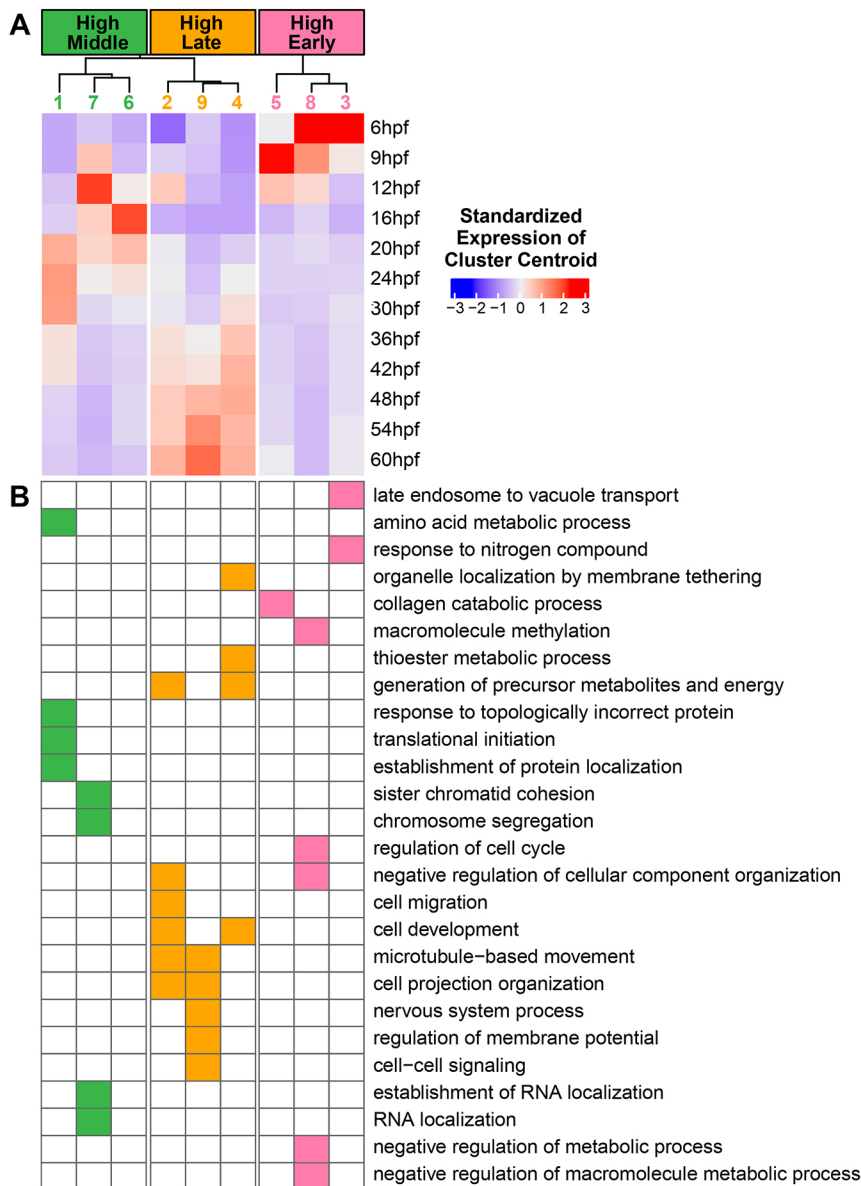


Fig. 7. Genes with late expression peaks in the scRNA-seq time course appear to participate in a unique set of biological processes. (A) Heatmap showing the centroids of expression profiles from the Mfuzz analysis of pseudobulked scRNA-seq data at each time point. (B) Heatmap showing GO term enrichments for each of the Mfuzz cluster profiles. A filled square indicates that there was a statistically significant enrichment of genes that perform the indicated biological process in the Mfuzz cluster.

aforementioned cell types and low expression elsewhere. This resulted in a set of 33 unique TFs (Fig. 8C, red; Fig. S6; Table S3) that are potentially important regulators of adult rudiment development. Some of these genes are already known to pattern the rudiment in *He* (*msx*; Fig. 8E) (Koop et al., 2017; Wilson et al., 2005), but others appear to be previously unappreciated transcriptional regulators of sea urchin development. For example, *giant* and *itl* showed strong localized expression in the rudiment ectoderm clusters (Fig. 8E). *lox*, which is known to pattern the hindgut in planktotrophic larvae (Cole et al., 2009), was expressed in the posterior ectoderm cluster, suggesting it may have an additional role in rudiment patterning (Fig. 8E). These 33 genes represent a list of candidate TFs to experimentally manipulate in future studies to understand the regulation of adult rudiment development.

DISCUSSION

In this paper, we present a comprehensive scRNA-seq atlas of embryonic and larval development in *He*. Building on the insights gained from previous datasets in sea urchins with planktotrophic

development (Foster et al., 2020; Massri et al., 2021; Paganos et al., 2021) and *He* (Massri et al., 2024 preprint), the present analysis provides single-cell transcriptomic data for late stages of sea urchin larval development that has lacked coverage in prior studies. Research on cell-type trajectories in later larval stages has historically been difficult owing to the long periods required to raise planktotrophic larvae through metamorphosis. In contrast, the accelerated developmental process of *He* makes it a useful model for studying the full time frame of larval developmental events, as well as the initial stages of adult morphogenesis (Wray, 2022). By using HCRs to map marker genes from the scRNA-seq dataset onto discrete tissues, we present some of the first transcriptomic data that is specific to the adult rudiment. These data are an important resource for studying the fates of larval cell types during adult development and allow us to start uncovering the GRN that controls this process. We also observed two important patterns regarding the biphasic lifecycle of *He*. First, several larval cell-fate trajectories have been modified compared to those in planktotrophs. Second, the larval developmental program appears to be distinct from the adult program, potentially allowing for each stage to evolve in a modular fashion.

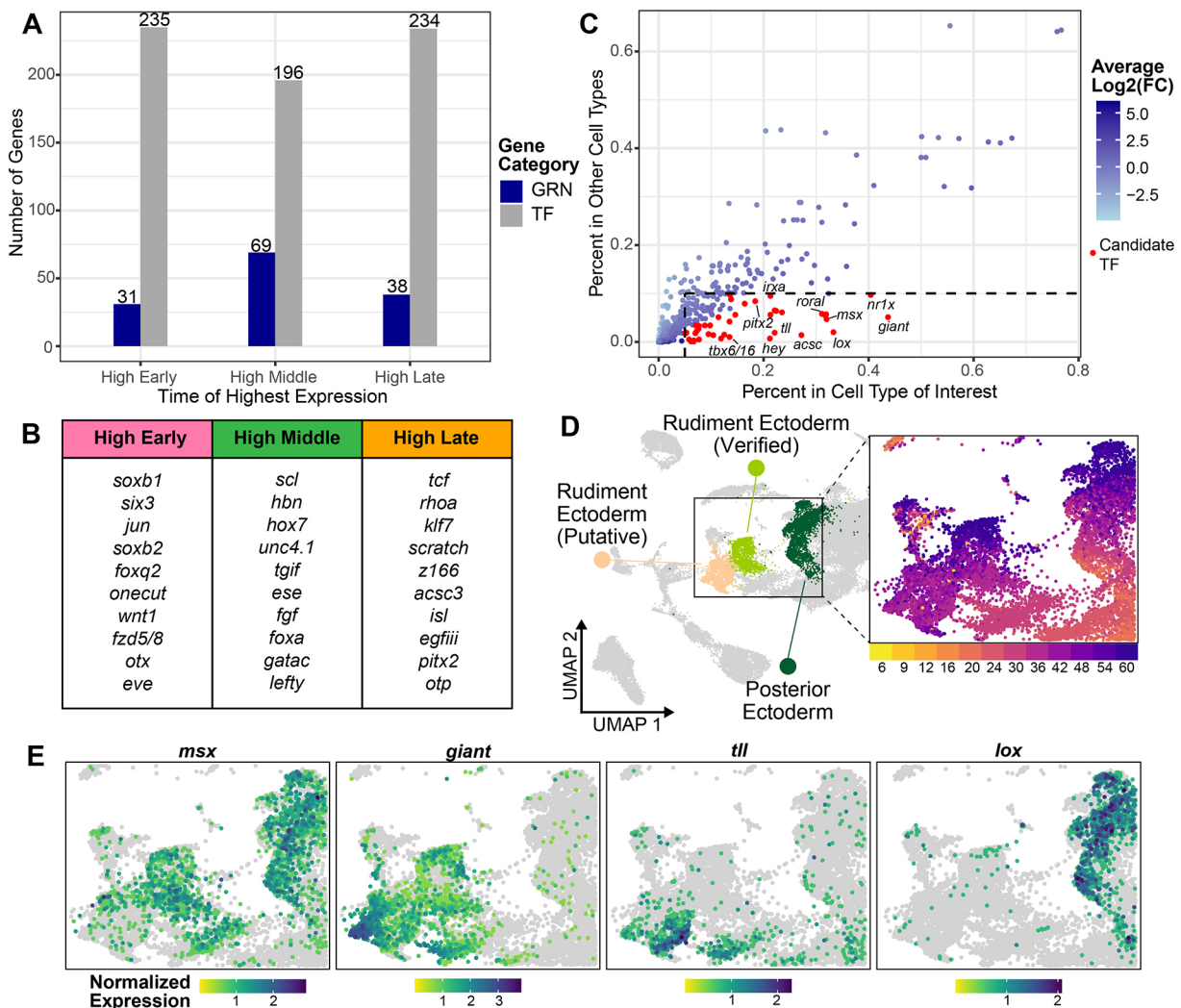


Fig. 8. Different gene regulatory suites appear to control the development of the larval and adult body plans. (A) Chart showing the number of TFs (gray) and embryonic GRN genes (blue) in the *He* genome with expression peaks in each temporal category. (B) The top 10 expressed embryonic GRN genes with expression peaks in each temporal category. (C) Plot showing the percentage of cells showing expression of TFs with late peaks in cell types of interest (rudiment ectoderm and posterior ectoderm) versus all other cell types, colored by their average log2-fold change in the cell types of interest (as a measure of enrichment). Points colored red indicate genes with average log2-fold changes >0.9 and that were expressed in >5% of cells in the cell type of interest and <10% of cells in the other cell types. (D) UMAP of *He* development highlighting the adult rudiment clusters. (E) UMAPs of *He* development zoomed in on the adult rudiment clusters, colored by expression levels for TFs potentially controlling adult rudiment morphogenesis.

Modifications to larval cell-type trajectories during the evolution of lecithotrophy

Prior research in *He* embryos revealed that several long-conserved cell fate trajectories have been reprogrammed during the evolution of lecithotrophy (Davidson et al., 2022; Parks et al., 1988; Wray and Raff, 1990). We observed similar trends with larval cell types in our scRNA-seq atlas. First, there was a notable change in the expression profiles of SKCs. In *He*, our co-expression analyses indicate that the SKCs lack expression of larva-specific skeletal genes but have gained expression of genes known to mark SMCs, which are likely the progenitors of the adult SKC population in planktotrophs (Figs 3 and 4) (Gao and Davidson, 2008; Tate et al., 2024; Yajima, 2007). This aligns with research in *He* showing that SKCs are no longer specified early in embryonic development and have lost expression of key components of the ancestral GRN (Davidson et al., 2022; Massri et al., 2024 preprint). The most likely explanation is that the ancestral larva-specific SKC population has been lost in *He*. This raises the question of what is the origin of the cell population that

constructs *He*'s vestigial larval skeleton. One possibility is that early-arising adult SKCs fulfill this role. A few of the cells that are intended to produce skeletal structures in the adult rudiment may instead respond to leftover patterning cues in the ectoderm that evolved to guide larval SKCs. Another possibility is that the larval skeleton may be produced by SMC-like cells that undergo transiting toward skeletogenic fates. Removal of the SKC progenitors in the embryos of species with planktotrophic larvae is known to trigger a subset of SMCs to activate the downstream SKC GRN (Ettensohn et al., 2007; Ettensohn and McClay, 1988; Sharma and Ettensohn, 2011). *He* embryos may have retained this transiting capability for the purposes of normal development. In either case, the elimination of the need to specify larval SKCs in the early embryo may have facilitated the acceleration of larval development during the evolution of lecithotrophy. Alternatively, the loss of most larval SKCs may be a consequence of the acceleration.

We also report the emergence of an undifferentiated cell cluster present in the *He* embryo and early larva. It does not appear that

these cells are a classical stem cell population, given their failure to express known pluripotency or germline-related genes (Fig. 5D; Table S1). We can also rule out that the cluster is the hypothesized ‘set aside’ cell population that is thought to supply most of the cells needed during metamorphosis (Davidson et al., 1995; Peterson et al., 1997), as few cells are present in the cluster at later time points (Fig. 5B). A more likely possibility is that these cells represent a previously unidentified cell state that is held back from differentiating into terminal cell fates early in development. It is already known that cell fate specification is delayed in *He* relative to species with planktotrophic development (Davidson et al., 2022), even though the overall larval developmental period is accelerated. These cells may be a dedicated population that is reserved for rapid differentiation once initial patterning events take place in the embryo. This is supported by the fact that the undifferentiated cluster groups with the early ectoderm cluster (which mostly consists of undifferentiated cells from the 6 hpf time point) in the cell type phylogenetic tree (Fig. S3), despite containing cells from several later time points. Although the exact fates of cells in this cluster are unknown, there is no analogous cell population that appears in scRNA-seq atlases of planktotrophs (Foster et al., 2020; Massri et al., 2021; Paganos et al., 2021) that arises early in development and persists over multiple time points. This cell state could therefore be a feature unique to lecithotrophs and may be a consequence of altered cell specification patterns in the *He* embryo.

A third *He* larval cell type that showed a notable shift was the neural cells. Work from the past couple of decades has led to a thorough understanding of neurogenesis early sea urchin larvae (reviewed by McClay, 2022). Many of these neural cell types were captured by our scRNA-seq atlas, including the larval serotonergic neurons and postoral neurons (Fig. 6). At the same time, we observed the emergence of several putative adult neural cell types at the later developmental time points in the atlas. Early appearance of adult neural cell types is likely associated with the acceleration of larval development in *He*. We also noticed that several TFs expressed in the putative adult neurons are expressed in clusters corresponding to adult rudiment tissues, suggesting that the processes of patterning the adult nervous system and broader pentaradial adult body plan are tightly connected (Ferkowicz and Raff, 2001; Sly et al., 2002). This aligns with recent work showing how anterior-posterior patterning genes, which primarily pattern the neuroectoderm in hemichordates and chordates, have been re-deployed to pattern the pentaradial adult body plan of sea stars (Formery et al., 2023). Nervous systems and body plans have frequently been observed to evolve hand-in-hand (Holland, 2016; Holland et al., 2015), so further research on the development of the sea urchin adult nervous system could help explain the transition from bilateral to pentaradial symmetry in echinoderm evolution.

The emergence of a new regulatory network governing adult body plan development in sea urchins

Our data also allow us to address questions about the regulation of adult morphogenesis in sea urchins. It has previously been unclear how organisms with biphasic lifecycles code for two different body plans in the same genome (Formery and Lowe, 2023). Our data suggest that a largely new set of regulatory genes controls the adult developmental process in sea urchins. A key finding is that most genes from the sea urchin embryonic GRN show the highest expression in the embryo and early larvae of *He* and have lower expression during later larval time points (Fig. 8A). To a certain extent, this is not entirely surprising because the GRN was constructed based on early developmental time points (Davidson

et al., 2002). However, many of the genes in the embryonic GRN are part of signaling systems used in a diverse array of developmental patterning events in metazoans, such as Wnt/Fzd and TGF β factors (Clevers, 2006; Wu and Hill, 2009), so the reduced role of some of these genes in adult patterning events is notable. This lack of overlap may partially be due to the drastic switch from bilateral to pentaradial symmetry during metamorphosis in echinoderms (Peterson et al., 2000; Smith, 2008). Not all of the genes that pattern the bilateral larva may be suited for patterning the pentaradial adult, requiring the introduction of a new suite of developmental regulators. Thus, the embryonic GRN appears to play a reduced role in regulating adult morphogenesis, at least in lecithotrophic *He* larvae.

In place of the embryonic GRN, we generated a candidate list of 33 TFs that may fulfill the regulatory requirements of patterning the pentaradial adult rudiment (Fig. 8E; Fig. S5; Table S3). Unfortunately, the scRNA-seq expression patterns of these genes alone did not reveal the upstream mechanism responsible for specifying the initial pentaradial pattern. To address this question, we can turn to previous embryological and gene expression studies in *He*. It appears that signaling between endomesodermal tissues and the overlying larval ectoderm is necessary for the invagination of the vestibule (Minsuk et al., 2005; Minsuk and Raff, 2002), and these interactions may also establish pentaradial symmetry in both tissues (Minsuk et al., 2009). Koop et al. (2017) followed up on this work by examining the expression of genes in the Nodal-BMP signaling network in early *He* larvae, which is known to pattern the dorsal-ventral and left-right axes of sea urchin embryos (reviewed by Molina et al., 2013). Of the genes they examined, *bmp2/4* had the earliest pentaradial expression pattern in the rudiment, suggesting that the Nodal-BMP network could be responsible for the initial establishment of the adult body plan. This led them to claim that the embryonic GRN may play a major role in patterning the rudiment, contradicting our data. One potential explanation for this discrepancy is that key signaling systems such as the Nodal-BMP network play a role in the upstream establishment of pentaradial symmetry, but the downstream TFs and effector genes that pattern specific adult tissues may lie outside the scope of the embryonic GRN. Extending scRNA-seq atlases to later larval stages in other sea urchin species, along with targeted perturbation studies, will allow us to construct a GRN specific to sea urchin adult morphogenesis.

The evolution of biphasic lifecycles

In summary, our analysis is valuable for understanding the switch in developmental strategy that occurred during the evolution of *He*, as well as the process of adult body plan construction in general across sea urchins. By examining a scRNA-seq time course that covers most of larval development in *He*, we assessed how larval cell types are established in sea urchins and how these may be involved in the transition to the adult body plan. Although our data do not capture gene expression events that occur immediately before and during metamorphosis, this work does provide a major advance in our knowledge of how late larval development is regulated in sea urchins. Furthermore, the fact that sea urchins may use two different GRNs to pattern the larval and adult body plans could address uncertainties about the origins of biphasic lifecycles in bilaterians (Davidson et al., 1995; Nielsen, 2013; Raff, 2008; Valentine and Collins, 2000). Instead of larval patterning mechanisms being extensively co-opted for patterning the adult, or vice versa, the two processes appear distinct. A scenario whereby the larval or adult phase was secondarily added to the lifecycle by simply co-opting

pre-existing regulatory networks (Sly et al., 2003) appears unlikely. Answering this question with more confidence will require comparative studies with bilaterian species in other phyla.

Our results also allow us to start to develop a model for biphasic lifecycle evolution. The acceleration of development in sea urchins that have lecithotrophic larvae could mean that adult development in *He* follows different patterns compared to species that have planktotrophic larvae. Even so, the adults of *He* and its sister species *Ht*, which has planktotrophic larvae, have remarkably similar morphologies (Byrne and O'Hara, 2017). Although it is possible that each species relies on a different developmental program to arrive at its adult form, a more parsimonious explanation is that *He* has retained many of the mechanisms found in species with planktotrophic development. At the same time, the divergence of larval patterning strategies in *He* compared to planktotrophs indicates that, under the appropriate selective conditions, developmental processes can be highly labile. We hypothesize that the larval and adult portions of biphasic lifecycles can be viewed as semi-independent modules, given that they may be controlled by largely distinct sets of regulatory factors. Taken together, these patterns suggest that the simultaneous conservation and modification of different stages of biphasic lifecycles may have contributed to the diversity of metazoan developmental strategies seen today.

MATERIALS AND METHODS

scRNA-seq analysis

Data retrieval

Heliocidaris erythrogramma embryos and larvae raised at 23°C were sampled at 12 time points from 6 to 60 hpf (6, 9, 12, 16, 20, 24, 30, 36, 42, 48, 54 and 60 hpf) to yield scRNA-seq libraries. All libraries were prepared and sequenced during data collection for the work described by Massri et al. (2024 preprint). However, only samples from the 6-30 hpf time points were used in that analysis. For the present analysis, data for the same 6-30 hpf time points, as well as five additional time points up to 60 hpf, were analyzed with a focus on larval and adult rudiment development. Gene expression counts tables resulting from processing the sequencing data using Cell Ranger (Zheng et al., 2017) were retrieved for all 12 time points and were used as input into downstream analysis steps.

Data filtering and normalization

After obtaining the counts tables, the R package Seurat v4.3.0 was used for downstream filtering and normalization steps (Hao et al., 2021). The counts tables for each time point were merged into the same Seurat object, which was filtered for high-quality cells with $nFeature_RNA > 200$, $nCount_RNA < 10,000$ and $nFeature_RNA < 4000$. In addition, a new column 'percent.Rb' was added to the metadata to identify the proportion of transcripts in each cell that have names matching the regular expression pattern '\\b\\w*Rp[sl]\\w*\\b', which identifies genes that encode ribosomal proteins. The function 'SCTransform' was used with the 'glmGamPoi' method and 6000 variable features to normalize, rescale, and identify variable features in the dataset (Choudhary and Satija, 2022). The parameter `vars.to.regress='percent.Rb'` was also passed to the 'SCTransform' function to regress out expression signals from ribosomal proteins.

Dimensionality reduction, clustering and visualization

The 'RunPCA' function was then used to perform principal component analysis and dimensional reduction on the Seurat object, using 200 principal components in total. Clustering of different cell types was then performed using the 'FindNeighbors' function with the first 195 principal component dimensions followed by running the 'FindClusters' function with `resolution=3`. This detected 57 distinct cell clusters across all 12 sample time points, which was a number that remained relatively robust even at higher resolution values. Each of the 57 cell clusters was assigned a cell-type identity based on expression of known sea urchin developmental GRN genes and *in situ* hybridization patterns (Fig. 2C-H; Fig. S2; Table S1). To

visualize the cell clusters in two-dimensional space, the UMAP algorithm was applied using the 'RunUMAP' function. Unless otherwise noted, all gene expression UMAPs were generated using the 'FeaturePlot' function from Seurat or the 'FeaturePlot_scCustom' wrapper function from the R package scCustomize v.2.0.1 (<https://zenodo.org/records/10161832>).

Neural cell analysis

The clusters that were annotated as 'Neural' in the full Seurat object were isolated into a separate Seurat object using the subset function. The neural-specific object was re-clustered using the 'FindNeighbors' function with the first 195 principal component dimensions and the 'FindClusters' function with `resolution=0.8`. This resulted in the identification of 17 neural cell types, which was a number that remained relatively robust to increases in the resolution value. Published data on neural marker genes were used to assign each of these clusters a neural cell type identity (Fig. S4; Table S2). Finally, the UMAP algorithm was applied to the neural-specific object using 'RunUMAP' to re-project the neural cells into UMAP space.

To evaluate the TFs used to pattern neural cell types, the list of neural TFs identified in the *Strongylocentrotus purpuratus* genome was retrieved from Burke et al. (2006a). The orthologs of these genes in the *He* genome were identified using the annotations published by Davidson et al. (2022). The top 30 expressed neural TFs were selected based on highest average expression levels across all the neural cell types. The expression levels of these genes in each neural cell type were plotted in Fig. 5C using the 'Clustered_DotPlot' function from scCustomize. K-means clustering was used to group the expression patterns of these genes in neural cells into ten clusters.

Co-expression analyses for skeletal and set aside genes

The Seurat function 'WhichCells' was used to highlight cells in atlas that co-expressed two different genes of interest. Cells with expression levels > 0.5 for both genes were designated as 'low co-expression', and cells with expression levels > 1 for both genes were designated as 'high co-expression'. Co-expressing cells were plotted in two-dimensional UMAP space.

scRNA-seq data for *Lytechinus variegatus* (*Lv*) embryos and larvae, a species with planktotrophic development, was retrieved from the scRNA-seq atlas presented by Massri et al. (2021). This atlas contains hourly or bi-hourly scRNA-seq time points of development in *Lv* from 2 to 24 hpf. The same steps used for analyzing SKC co-expression patterns in the *He* scRNA-seq atlas were applied to the *Lv* scRNA-seq atlas.

Waddington-OT trajectory inference

Waddington-OT trajectory inference (v.1.0.8) was performed as previously described (Massri et al., 2024preprint, 2021; Schiebinger et al., 2019), with only a few minor modifications for the full 6-60 hpf time course. Briefly, the full Seurat object was used to generate a 'SCTransform'-normalized expression matrix showing gene expression for each cell. Separate files with cell-type annotation and UMAP embedding assignments for each cell were also generated. Cell division rates for each cell type were estimated based on the expected number of cells that are supposed to be present at each developmental time point. These rates, along with the expression matrix, were used to compute transport maps between each of the time points, with parameters set at `epsilon=0.05`, `lambda1=1`, `lambda2=50`, and `growth_iterations=20`. The transition_table function was used to generate the plot in Fig. 5C.

Pseudobulk analysis and gene expression profile clustering

To perform pseudobulking of the scRNA-seq data, raw gene expression count data were extracted from the Seurat object. Gene expression counts were aggregated across the cells at each time point, resulting in a matrix with the 12 time points as columns and genes as rows. The counts table was input into the 'DESeq' function from the R package DESeq2 v.1.38.3 (Love et al., 2014) to normalize the gene expression data using the median of ratios method. The normalized counts data were then exported from the DESeq object for input into the Mfuzz pipeline.

The R package Mfuzz v.2.58.0 (Futschik and Carlisle, 2005; Kumar and Futschik, 2007) was used to perform fuzzy c-means clustering of temporal

gene expression profiles. First, genes with normalized expression counts <5 at any time point were filtered out of the dataset. The data were filtered to exclude genes with undetected expression values and then standardized (for a total of 19,585 genes). The fuzzifier parameter m was estimated using the ‘mestimate’ function. Finally, the normalized and standardized counts data, along with the estimate for m , were provided to the ‘mfuzz’ function to cluster the gene expression profiles. The optimal number of clusters (value of c) was determined as described by Israel et al. (2016). Briefly, the number of clusters was chosen to be as high as possible without having the correlation between any pair of cluster centroids exceed 0.85. This was to ensure that the data were not over clustered and that each cluster showed a relatively distinct pattern from the others. Nine clusters ($c=9$) were selected as the optimal value. The centroids of each cluster were plotted (Fig. 6A) using the R package ComplexHeatmap v.2.18.0 (Gu et al., 2016). This package was also used to perform k-means clustering on the Mfuzz cluster profiles based on Euclidean distance with $k=3$.

Protein functional annotation and GO analysis

The Linux package InterProScan v.5.64-96.0 (Jones et al., 2014) was used to assign functional annotations to the protein-coding sequences in the *He* genome. Gene models from *He* were retrieved from Davidson et al. (2022) and translated into peptide sequences. The main InterProScan analysis pipeline was run on these peptide sequences to assign GO terms using all the functional annotation databases contained in the standard InterProScan distribution.

GO over-representation analysis was performed using the R package clusterProfiler v.4.10.0 (Wu et al., 2021). The ‘compareCluster’ function was used to compare over-represented GO terms between genes assigned to each of the nine Mfuzz clusters. ‘Biological Process’ GO terms were specifically analyzed, with $pvalueCutoff=0.05$, $pAdjustMethod='BH'$ and $qvalueCutoff=0.2$. The ‘pairwise_termsim’ and ‘treemap’ functions from the R package enrichplot v.1.22.0 (<https://yulab-smu.top/biomedical-knowledge-mining-book/>) were used to plot the heatmap in Fig. 6D.

GO over-representation analysis was also used to identify the top ‘Biological Process’ GO terms for genes enriched in the undifferentiated cell cluster. The Seurat ‘FindMarkers’ function was used to identify the genes significantly enriched in this cluster ($n=342$), with $m.pct=0.25$, $logfc.threshold=0.25$ and $P_value_adj<0.05$. The ‘enrichGO’ function from clusterProfiler was run to compare these genes against all other genes in the genome as a background set, with $pvalueCutoff=0.05$, $pAdjustMethod='BH'$ and $qvalueCutoff=0.2$.

Embryonic gene regulatory network analysis

A list of genes in the ancestral sea urchin embryonic GRN (192 total) was retrieved from Davidson et al. (2022). In order to access gene expression information across all analyses, only embryonic GRN genes that were retained in the Mfuzz and Seurat analyses were analyzed (138 total). Based on the Mfuzz cluster to which each gene was assigned, embryonic GRN genes were classified as belonging to the ‘High Early’, ‘High Middle’ or ‘High Late’ expression profile groups, as described above.

TF analysis

When assembling a list of all the predicted TFs in the *He* genome, the goal was to evaluate as many genes as possible, rather than relying on a restricted set of verified genes. Thus, lists of predicted TFs from two sources were combined along with a list of zinc finger proteins, which are frequently transcription regulators (Klug, 2010). First, the list of TFs and zinc finger proteins used in a previous genomic analysis of *He* (Devens et al., 2023) was retrieved. This was combined with a larger list of TFs predicted from their associated GO terms. Genes with the following terms were included: GO:0003700 (‘DNA-binding transcription factor activity’), GO:0001216 (‘DNA-binding transcription activator activity’), GO:0001227 (‘DNA-binding transcription repressor activity, RNA polymerase II-specific’), GO:0001228 (‘DNA-binding transcription activator activity, RNA polymerase II-specific’), GO:0000981 (‘DNA-binding transcription factor activity, RNA polymerase II-specific’) and GO:0001217 (‘DNA-binding transcription repressor activity’) (Gene Ontology Consortium et al., 2023).

In total, the combination of these two lists resulted in 1400 transcripts in the *He* genome that are predicted to be TFs, some of which may be duplicated copies of the same gene. Only TFs that were retained in the Mfuzz analysis ($n=665$) were used for downstream steps.

To identify the key TFs responsible for patterning adult rudiment structures in *He*, the TF list was filtered for genes that had temporal expression profiles belonging to the ‘High Late’ category ($n=234$). The Seurat function ‘FoldChange’ was used to identify genes from this list that were enriched in the ‘Rudiment Ectoderm (Verified)’, ‘Rudiment Ectoderm (Putative)’ and ‘Posterior Ectoderm’ clusters in the scRNA-seq atlas. TF genes were retained that had $pct.2<0.1$, $pct.1>0.05$, and $avg_log2FC>0.9$ in any one of these cell type categories. This resulted in a list of 33 unique genes that showed relatively high, enriched expression in adult tissues in *He* larvae (Fig. 8E; Fig. S5; Table S3).

All computational analyses were conducted using R v.4.2.0 or v.4.3.0 (depending on package compatibility) and Python v.3.8.

In situ gene expression analysis

Embryo fixation

He embryos and larvae from desired developmental stages were fixed overnight in 4% paraformaldehyde in artificial sea water (ASW) at 4°C. Embryos were washed once with ASW and once with ice-cold methanol. These were then moved to a fresh aliquot of methanol and stored at -20°C until used for later steps.

HCR probe design

DNA probes for HCR were designed by Molecular Instruments or using the *insitu_probe_generator* python script from Kuehn et al. (2022). For the non-Molecular Instruments probes, the script selected the maximum number of probes to hybridize with the exon sequence of the gene of interest without overlapping, and the probes were designed with attachments for either B1, B2 or B3 hairpins for the HCR amplification step. Probes were synthesized using the oPools Oligo Pools service from Integrated DNA Technologies. Upon receipt, the probes were diluted to 1 μM in RNase-free water and stored at -20°C . See Table S4 for the probe sequences.

HCR

Whole-mount fluorescent *in situ* hybridization was performed on fixed *He* embryos and larvae using the v.3.0 HCR method, following the sea urchin-specific protocol published by Choi et al. (2016) with some modifications. Unless otherwise noted, all washes were performed at room temperature. Fixed samples were rehydrated stepwise from methanol into PBST (1 $\times$ PBS and 0.1% Tween 20). To permeabilize the embryo membranes, the samples were incubated for 30 min in a detergent solution containing 1% SDS, 0.5% Tween 20, 50 mM Tris-HCl, 1 mM EDTA, and 150 mM NaCl. Samples were washed in PBST, and a post-fixation step was performed by incubating the embryos in 4% paraformaldehyde in PBST for 25 min. The fixative was removed with five washes of PBST.

The samples were then pre-incubated in Probe Hybridization Buffer (Molecular Instruments) for 30 min at 37°C. Probe solutions were prepared by adding 2 μl each of DNA probes for one or two genes of interest to Probe Hybridization Buffer for each embryo sample (final concentration, 10 nM), adding buffer such that the final volume was 200 μl . The samples were moved into the probe solution and incubated for 36–48 h at 37°C to ensure that the probes fully penetrated the large *He* embryos. The probe solution was removed by washing 4 $\times$ 15 min in Probe Wash Buffer (Molecular Instruments) at 37°C. Samples were washed 2 $\times$ 5 min in 5 $\times$ SSC (5 $\times$ SSC and 0.1% Tween 20).

Samples were then pre-incubated in Amplification Buffer (Molecular Instruments) for 30 min at room temperature. During this incubation step, DNA hairpins with Alexa Fluor 488, Alexa Fluor 546 or Alexa Fluor 647 fluorophores (Molecular Instruments) and B1, B2 or B3 binding regions were heated to 95°C for 90 s and then stored in the dark at room temperature for 30 min. Hairpin solutions were prepared by adding 4 μl of hairpin h1 and 4 μl of hairpin h2 that match the B1, B2 or B3 probe attachment to Amplification Buffer for each embryo sample (final concentration, 60 nmol), adding buffer such that the final volume was 200 μl . Samples

were moved into the hairpin solution and incubated for 16–24 h in the dark at room temperature. Hairpin solution was then removed by washing 3×20 min in 5× SSCT.

Samples were incubated for 3–6 h in 1:500 DAPI in 500 mM NaCl solution (500 mM NaCl in 1× PBS) in the dark at room temperature. The embryos were then cleared in glycerol solution by incubating for 1 h in 50% glycerol in 1× PBS, and then moving to 70% glycerol in 1× PBS. Samples were mounted for imaging on slides in the 70% glycerol solution.

To verify probe specificity, the HCR protocol was run with hairpins only (no gene-specific probes), as well as with no probes and no hairpins. No localized fluorescence patterns were observed under either condition (Fig. S7).

Microscopy and image analysis

Mounted embryos and larvae were visualized using a Zeiss 880 Airyscan inverted confocal microscope paired with a 10×/0.30 EC Plan-Neofluar objective. Images were analyzed using the Fiji distribution of ImageJ (Schindelin et al., 2012). Only linear adjustments, including the ‘minimum’ and ‘brightness’ settings for each color channel, were made to remove background fluorescence. All images presented in this article are single slices of the original multi-slice z-stacks.

Acknowledgements

We thank Esther Miranda, Zachary Pracher, Hannah Devens, Jane Swart, Carl Manner, Emma Wallace, and the other members of the Wray lab for their helpful advice during this project. We also thank the staff of the Sydney Institute for Marine Sciences for supporting our data collection efforts in Australia. We acknowledge the use of the Duke University Light Microscopy core facility for preparing micrographs of the HCR samples presented in this paper, as well as the Duke University Sequencing and Genomic Technologies core facility for sequencing the single cell libraries.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: B.D.M., G.A.W.; Methodology: B.D.M., A.J.M., G.A.W.; Software: B.D.M., A.B.; Validation: B.D.M.; Formal analysis: B.D.M., A.B.; Investigation: B.D.M., A.J.M., M.B.; Resources: M.B., D.R.M., G.A.W.; Data curation: B.D.M.; Writing - original draft: B.D.M.; Writing - review & editing: B.D.M., A.J.M., A.B., M.B., D.R.M., G.A.W.; Visualization: B.D.M.; Supervision: G.A.W.; Project administration: G.A.W.; Funding acquisition: G.A.W.

Funding

This project was funded by the National Science Foundation (IOS-1929934 to G.A.W.). Other authors received funding support from the National Institutes of Health (RO1 HD14483 to D.R.M. and A.B.; Predoctoral T32 HD040372 to A.J.M.), the National Science Foundation (Predoctoral DGE-1644868 to A.J.M.), the Australian Research Council (to M.B.), and the Duke University Undergraduate Research Support Office (to B.D.M.). Open Access funding provided by Directorate for Biological Sciences. Deposited in PMC for immediate release.

Data availability

All code used to conduct the analyses and generate the figures in this paper can be found on GitHub (https://github.com/brennandmcdonald/He_Larval_scRNA-seq). The scRNA-seq data for *Heliocidaris erythrogramma* have been deposited in Gene Expression Omnibus under accession numbers [GSE277517](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE277517) (for the 6–30 hpf scRNA-seq time points) and [GSE277518](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE277518) (for the 36–60 hpf scRNA-seq time points). The scRNA-seq data for *Lytechinus variegatus* have been deposited in Gene Expression Omnibus under accession number [GSE184538](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE184538).

Peer review history

The peer review history is available online at <https://journals.biologists.com/dev/lookup/doi/10.1242/dev.203015.reviewer-comments.pdf>

Special Issue

This article is part of the Special Issue ‘Uncovering developmental diversity’, edited by Cassandra Extavour, Liam Dolan and Karen Sears. See related articles at <https://journals.biologists.com/dev/issue/151/20>.

References

Bisgrove, B. W., Andrews, M. E. and Raff, R. A. (1991). Fibropellins, products of an Egf repeat-containing gene, form a unique extracellular matrix structure that

- surrounds the sea urchin embryo. *Dev. Biol.* **146**, 1, 89–99. doi:10.1016/0012-1606(91)90449-D
- Burke, R. D. (1980). Podial sensory receptors and the induction of metamorphosis in echinoids. *J. Exp. Mar. Biol. Ecol.* **47**, 223–234. doi:10.1016/0022-0981(80)90040-4
- Burke, R. D., Angerer, L. M., Elphick, M. R., Humphrey, G. W., Yaguchi, S., Kiyama, T., Liang, S., Mu, X., Agca, C., Klein, W. H. et al. (2006a). A genomic view of the sea urchin nervous system. *Dev. Biol.* **300**, 434–460. doi:10.1016/j.ydbio.2006.08.007
- Burke, R. D., Osborne, L., Wang, D., Murabe, N., Yaguchi, S. and Nakajima, Y. (2006b). Neuron-specific expression of a synaptotagmin gene in the sea urchin. *J. Comp. Neurol.* **496**, 2, 244–251. doi:10.1002/cne.20939
- Byrne, M. and O'Hara, T. D. (2017). *Australian Echinoderms: Biology, Ecology and Evolution*. CSIRO Publishing. doi:10.1071/9781486307630
- Byrne, M., Koop, D., Morris, V. B., Chui, J., Wray, G. A. and Cisternas, P. (2018). Expression of genes and proteins of the pax-six-eya-dach network in the metamorphic sea urchin: Insights into development of the enigmatic echinoderm body plan and sensory structures. *Dev. Dyn.* **247**, 239–249. doi:10.1002/dvdy.24584
- Calestani, C., Rast, J. P. and Davidson, E. H. (2003). Isolation of pigment cell specific genes in the sea urchin embryo by differential macroarray screening. *Development* **130**, 4587–4596. doi:10.1242/dev.00647
- Choi, H. M. T., Calvert, C. R., Husain, N., Huss, D., Barsi, J. C., Deverman, B. E., Hunter, R. C., Kato, M., Lee, S. M., Abelin, A. C. T. et al. (2016). Mapping a multiplexed zoo of mRNA expression. *Development* **143**, 3632–3637. doi:10.1242/dev.140137
- Choudhary, S. and Satija, R. (2022). Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol.* **23**, 27. doi:10.1186/s13059-021-02584-9
- Clevers, H. (2006). Wnt/ β -catenin signaling in development and disease. *Cell* **127**, 469–480. doi:10.1016/j.cell.2006.10.018
- Cole, A. G., Rizzo, F., Martinez, P., Fernandez-Serra, M. and Arnone, M. I. (2009). Two Parahox genes, SpLox and SpCdx, interact to partition the posterior endoderm in the formation of a functional gut. *Development* **136**, 541–549. doi:10.1242/dev.029959
- Davidson, E. H., Peterson, K. J. and Cameron, R. A. (1995). Origin of bilaterian body plans: evolution of developmental regulatory mechanisms. *Science* **270**, 1319–1325. doi:10.1126/science.270.5240.1319
- Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.-H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C. et al. (2002). A genomic regulatory network for development. *Science* **295**, 1669–1678. doi:10.1126/science.1069883
- Davidson, P. L., Guo, H., Swart, J. S., Massri, A. J., Edgar, A., Wang, L., Berrio, A., Devens, H. R., Koop, D., Cisternas, P. et al. (2022). Recent reconfiguration of an ancient developmental gene regulatory network in *Heliocidaris* sea urchins. *Nat. Ecol. Evol.* **6**, 1907–1920. doi:10.1038/s41559-022-01906-9
- Devens, H. R., Davidson, P. L., Byrne, M. and Wray, G. A. (2023). Hybrid epigenomes reveal extensive local genetic changes to chromatin accessibility contribute to divergence in embryonic gene expression between species. *Mol. Biol. Evol.* **40**, 11. doi:10.1093/molbev/msad222
- Emlet, R. B. (1995). Larval spicules, cilia, and symmetry as remnants of indirect development in the direct-developing sea-urchin *Heliocidaris erythrogramma*. *Dev. Biol.* **167**, 405–415. doi:10.1006/dbio.1995.1037
- Ettensohn, C. A. and McClay, D. R. (1988). Cell lineage conversion in the sea-urchin embryo. *Dev. Biol.* **125**, 396–409. doi:10.1016/0012-1606(88)90220-5
- Ettensohn, C. A., Illies, M. R., Oliveri, P. and De Jong, D. L. (2003). Alx1, a member of the Cart1/Alx3/Alx4 subfamily of Paired-class homeodomain proteins, is an essential component of the gene network controlling skeletogenic fate specification in the sea urchin embryo. *Development* **130**, 2917–2928. doi:10.1242/dev.00511
- Ettensohn, C. A., Kitazawa, C., Cheers, M. S., Leonard, J. D. and Sharma, T. (2007). Gene regulatory networks and developmental plasticity in the early sea urchin embryo: alternative deployment of the skeletogenic gene regulatory network. *Development* **134**, 3077–3087. doi:10.1242/dev.009092
- Ferkowicz, M. J. and Raff, R. A. (2001). Wnt gene expression in sea urchin development: Heterochronies associated with the evolution of developmental mode. *Evol. Dev.* **3**, 24–33. doi:10.1046/j.1525-142x.2001.00084.x
- Formery, L. and Lowe, C. J. (2023). Integrating complex life cycles in comparative developmental biology. *Annu. Rev. Genet.* **57**, 321–339. doi:10.1146/annurev-genet-071719-020641
- Formery, L., Peluso, P., Kohnle, I., Malnick, J., Thompson, J. R., Pitel, M., Uhlinger, K. R., Rokhsar, D. S., Rank, D. R. and Lowe, C. J. (2023). Molecular evidence of anteroposterior patterning in adult echinoderms. *Nature* **623**, 555–561. doi:10.1038/s41586-023-06669-2
- Foster, S., Oulhen, N. and Wessel, G. (2020). A single cell RNA sequencing resource for early sea urchin development. *Development* **147**, 17. doi:10.1242/dev.191528
- Futschik, M. E. and Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. *J. Bioinform. Comput. Biol.* **3**, 965–988. doi:10.1142/S0219720005001375

- Gao, F. and Davidson, E. H. (2008). Transfer of a large gene regulatory apparatus to a new developmental address in echinoid evolution. *Proc. Natl. Acad. Sci. USA* **105**, 6091-6096. doi:10.1073/pnas.0801201105
- Gene Ontology Consortium, Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N. L. et al. (2023). The Gene Ontology knowledgebase in 2023. *Genetics* **224**, 1. doi:10.1093/genetics/iyad031
- Gonzalez, P., Uhlinger, K. R. and Lowe, C. J. (2017). The adult body plan of indirect developing hemichordates develops by adding a Hox-patterned trunk to an anterior larval territory. *Curr. Biol.* **27**, 87-95. doi:10.1016/j.cub.2016.10.047
- Gu, Z., Eils, R. and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849. doi:10.1093/bioinformatics/btw313
- Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., III, Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M. et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587. doi:10.1016/j.cell.2021.04.048
- Hart, M. W., Abt, C. H. J. and Emlet, R. B. (2011). Molecular phylogeny of echinometrid sea urchins: more species of *Heliocidaris* with derived modes of reproduction. *Invertebr. Biol.* **130**, 175-185. doi:10.1111/j.1744-7410.2011.00231.x
- Holland, N. D. (2016). Nervous systems and scenarios for the invertebrate-to-vertebrate transition. *Philos. Trans. R. Soc. B* **371**, 1685. doi:10.1098/rstb.2015.0047
- Holland, N. D., Holland, L. Z. and Holland, P. W. H. (2015). Scenarios for the making of vertebrates. *Nature* **520**, 450-455. doi:10.1038/nature14433
- Israel, J. W., Martik, M. L., Byrne, M., Raff, E. C., Raff, R. A., McClay, D. R. and Wray, G. A. (2016). Comparative developmental transcriptomics reveals rewiring of a highly conserved gene regulatory network during a major life history switch in the sea urchin genus *Heliocidaris*. *PLoS Biol.* **14**, e1002391. doi:10.1371/journal.pbio.1002391
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G. et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236-1240. doi:10.1093/bioinformatics/btu031
- Juliano, C. E., Voronina, E., Stack, C., Aldrich, M., Cameron, A. R. and Wessel, G. M. (2006). Germ line determinants are not localized early in sea urchin development, but do accumulate in the small micromere lineage. *Dev. Biol.* **300**, 406-415. doi:10.1016/j.ydbio.2006.07.035
- Klug, A. (2010). The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem.* **79**, 213-231. doi:10.1146/annurev-biochem-010909-095056
- Koop, D., Cisternas, P., Morris, V. B., Strbenac, D., Yang, J. Y. H., Wray, G. A. and Byrne, M. (2017). Nodal and BMP expression during the transition to pentamerism in the sea urchin *Heliocidaris erythrogramma*: insights into patterning the enigmatic echinoderm body plan. *BMC Dev. Biol.* **17**, 4. doi:10.1186/s12861-017-0145-1
- Kuehn, E., Clausen, D. S., Null, R. W., Metzger, B. M., Willis, A. D. and Özpolat, B. D. (2022). Segment number threshold determines juvenile onset of germline cluster expansion in *Platynereis dumerilii*. *J. Exp. Zool. B Mol. Dev. Evol.* **338**, 225-240. doi:10.1002/jez.b.23100
- Kumar, L. and Futschik, M. E. (2007). Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* **2**, 5-7. doi:10.6026/97320630002005
- Lacalli, T. C. (2005). Protochordate body plan and the evolutionary role of larvae: old controversies resolved? *Can. J. Zool.* **83**, 216-224. doi:10.1139/z04-162
- Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. doi:10.1186/s13059-014-0550-8
- Martin-Zamora, F. M., Liang, Y., Guynes, K., Carrillo-Baltodano, A. M., Davies, B. E., Donnellan, R. D., Tan, Y., Moggioli, G., Seudre, O., Tran, M. et al. (2023). Annelid functional genomics reveal the origins of bilaterian life cycles. *Nature* **615**, 105-110. doi:10.1038/s41586-022-05636-7
- Massri, A. J., Greenstreet, L., Afanassiev, A., Berrio, A., Wray, G. A., Schiebinger, G. and McClay, D. R. (2021). Developmental single-cell transcriptomics in the *Lytechinus variegatus* sea urchin embryo. *Development* **148**, 19. doi:10.1242/dev.198614
- Massri, A. J., Berrio, A., Afanassiev, A., Greenstreet, L., Pipho, K., Byrne, M., Schiebinger, G., McClay, D. R. and Wray, G. A. (2024). Single-cell transcriptomics reveals evolutionary reconfiguration of embryonic cell fate specification in the sea urchin *Heliocidaris erythrogramma*. *bioRxiv*. doi:10.1101/2024.04.30.591752
- Materna, S. C., Ransick, A., Li, E. H. and Davidson, E. H. (2013). Diversification of oral and aboral mesodermal regulatory states in pregastrular sea urchin embryos. *Dev. Biol.* **375**, 92-104. doi:10.1016/j.ydbio.2012.11.033
- McClay, D. R. (2022). Development of a larval nervous system in the sea urchin. *Curr. Top. Dev. Biol.* **146**, 25-48. doi:10.1016/bs.ctdb.2021.10.002
- McEdward, L. R. and Janies, D. A. (1997). Relationships among development, ecology, and morphology in the evolution of Echinoderm larvae and life cycles. *Biol. J. Linn. Soc.* **60**, 381-400. doi:10.1111/j.1095-8312.1997.tb01502.x
- McEdward, L. R. and Miner, B. G. (2001). Larval and life-cycle patterns in echinoderms. *Can. J. Zool.* **79**, 1125-1170. doi:10.1139/z00-218
- Minsuk, S. B. and Raff, R. A. (2002). Pattern formation in a pentamerous animal: induction of early adult rudiment development in sea urchins. *Dev. Biol.* **247**, 335-350. doi:10.1006/dbio.2002.0704
- Minsuk, S. B., Andrews, M. E. and Raff, R. A. (2005). From larval bodies to adult body plans: patterning the development of the presumptive adult ectoderm in the sea urchin larva. *Dev. Genes Evol.* **215**, 383-392. doi:10.1007/s00427-005-0486-9
- Minsuk, S. B., Turner, F. R., Andrews, M. E. and Raff, R. A. (2009). Axial patterning of the pentaradial adult echinoderm body plan. *Dev. Genes Evol.* **219**, 89-101. doi:10.1007/s00427-009-0270-3
- Molina, M. D., de Crozé, N., Haillet, E. and Lepage, T. (2013). Nodal: master and commander of the dorsal-ventral and left-right axes in the sea urchin embryo. *Curr. Opin. Genet. Dev.* **23**, 445-453. doi:10.1016/j.gde.2013.04.010
- Moran, N. A. (1994). Adaptation and constraint in the complex life-cycles of animals. *Annu. Rev. Ecol. Syst.* **25**, 573-600. doi:10.1146/annurev.es.25.110194.003041
- Nielsen, C. (2009). How did indirect development with Planktotrophic larvae evolve? *Biol. Bull.* **216**, 203-215. doi:10.1086/BBLv216n3p203
- Nielsen, C. (2013). Life cycle evolution: was the eumetazoan ancestor a holopelagic, planktotrophic gastraea? *BMC Evol. Biol.* **13**, 171. doi:10.1186/1471-2148-13-171
- Paganos, P., Voronov, D., Musser, J. M., Arendt, D. and Arnone, M. I. (2021). 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* **10**, e70416. doi:10.7554/eLife.70416
- Parks, A. L., Parr, B. A., Chin, J. E., Leaf, D. S. and Raff, R. A. (1988). Molecular analysis of heterochronic changes in the evolution of direct developing sea-urchins. *J. Evolution. Biol.* **1**, 27-44. doi:10.1046/j.1420-9101.1988.1010027.x
- Peter, I. S. and Davidson, E. H. (2011). A gene regulatory network controlling the embryonic specification of endoderm. *Nature* **474**, 635-639. doi:10.1038/nature10100
- Peterson, K. J., Cameron, R. A. and Davidson, E. H. (1997). Set-aside cells in maximal indirect development: evolutionary and developmental significance. *BioEssays* **19**, 623-631. doi:10.1002/bies.950190713
- Peterson, K. J., Arenas-Mena, C. and Davidson, E. H. (2000). The A/P axis in echinoderm ontogeny and evolution: evidence from fossils and molecules. *Evol. Dev.* **2**, 93-101. doi:10.1046/j.1525-142x.2000.00042.x
- Popodi, E., Kissinger, J. C., Andrews, M. E. and Raff, R. A. (1996). Sea urchin Hox genes: Insights into the ancestral Hox cluster. *Mol. Biol. Evol.* **13**, 1078-1086. doi:10.1093/oxfordjournals.molbev.a025670
- Poustka, A. J., Kühn, A., Radosavljevic, V., Wellenreuther, R., Lehrach, H. and Panopoulou, G. (2004). On the origin of the chordate central nervous system: expression of *oncut* in the sea urchin embryo. *Evol. Dev.* **6**, 227-236. doi:10.1111/j.1525-142X.2004.04028.x
- Raff, R. A. (1992). Direct-developing sea urchins and the evolutionary reorganization of early development. *BioEssays* **14**, 211-218. doi:10.1002/bies.950140403
- Raff, R. A. (2008). Origins of the other metazoan body plans: the evolution of larval forms. *Philos. Trans. R. Soc. B* **363**, 1473-1479. doi:10.1098/rstb.2007.2237
- Rafiq, K., Cheers, M. S. and Ettensohn, C. A. (2012). The genomic regulatory control of skeletal morphogenesis in the sea urchin. *Development* **139**, 579-590. doi:10.1242/dev.073049
- Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., Gould, J., Liu, S., Lin, S., Berube, P. et al. (2019). Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell* **176**, 928-943. doi:10.1016/j.cell.2019.01.006
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682. doi:10.1038/nmeth.2019
- Sethi, A. J., Angerer, R. C. and Angerer, L. M. (2009). Gene regulatory network interactions in sea urchin endomesoderm induction. *PLoS Biol.* **7**, 248-264. doi:10.1371/journal.pbio.1000029
- Sharma, T. and Ettensohn, C. A. (2011). Regulative deployment of the skeletogenic gene regulatory network during sea urchin development. *Development* **138**, 2581-2590. doi:10.1242/dev.065193
- Slota, L. A., Miranda, E. M. and McClay, D. R. (2019). Spatial and temporal patterns of gene expression during neurogenesis in the sea urchin. *Evodevo* **10**, 2. doi:10.1186/s13227-019-0115-8
- Sly, B. J., Hazel, J. C., Popodi, E. M. and Raff, R. A. (2002). Patterns of gene expression in the developing adult sea urchin central nervous system reveal multiple domains and deep-seated neural pentamerism. *Evol. Dev.* **4**, 189-204. doi:10.1046/j.1525-142X.2002.02002.x
- Sly, B. J., Snoko, M. S. and Raff, R. A. (2003). Who came first - larvae or adults? Origins of bilaterian metazoan larvae. *Int. J. Dev. Biol.* **47**, 623-632.
- Smith, A. B. (2008). Deuterostomes in a twist: the origins of a radical new body plan. *Evol. Dev.* **10**, 493-503. doi:10.1111/j.1525-142X.2008.00260.x
- Smith, M. J., Boom, J. D. G. and Raff, R. A. (1990). Single-copy DNA distance between 2 congeneric sea-urchin species exhibiting radically different modes of development. *Mol. Biol. Evol.* **7**, 315-326.

- Strathmann, R. R. (1978). The evolution and loss of feeding larval stages of marine invertebrates. *Evolution* **32**, 894-906. doi:10.2307/2407502
- Strathmann, R. R. (1985). Feeding and nonfeeding larval development and life-history evolution in marine invertebrates. *Annu. Rev. Ecol. Syst.* **16**, 339-361. doi:10.1146/annurev.es.16.110185.002011
- Su, Y.-H., Li, E., Geiss, G. K., Longabaugh, W. J. R., Krämer, A. and Davidson, E. H. (2009). A perturbation model of the gene regulatory network for oral and aboral ectoderm specification in the sea urchin embryo. *Dev. Biol.* **329**, 410-421. doi:10.1016/j.ydbio.2009.02.029
- Sumrall, C. D. and Wray, G. A. (2007). Ontogeny in the fossil record: diversification of body plans and the evolution of 'aberrant' symmetry in Paleozoic echinoderms. *Paleobiology* **33**, 149-163. doi:10.1666/06053.1
- Tate, H. M., Barone, V., Schrankel, C. S., Hamdoun, A. and Lyons, D. C. (2024). Localization and origins of juvenile skeletogenic cells in the sea urchin *Lytechinus pictus*. *Dev. Biol.* **514**, 12-27. doi:10.1016/j.ydbio.2024.05.012
- Telmer, C. A., Karimi, K., Chess, M. M., Agalakov, S., Arshinoff, B. I., Lotay, V., Wang, D. Z., Chu, S., Pells, T. J., Vize, P. D. et al. (2024). Echinobase: a resource to support the echinoderm research community. *Genetics* **227**, iyae002. doi:10.1093/genetics/iyae002
- Thorson, G. (1950). Reproductive and larval ecology of marine bottom invertebrates. *Biol. Rev.* **25**, 1-45. doi:10.1111/j.1469-185X.1950.tb00585.x
- Ullrich-Lüter, E. M., Dupont, S., Arboleda, E., Hausen, H. and Arnone, M. I. (2011). Unique system of photoreceptors in sea urchin tube feet. *Proc. Natl. Acad. Sci. USA* **108**, 8367-8372. doi:10.1073/pnas.1018495108
- Valentine, J. W. and Collins, A. G. (2000). The significance of moulting in Ecdysozoan evolution. *Evol. Dev.* **2**, 152-156. doi:10.1046/j.1525-142X.2000.00043.x
- Williams, D. H. C. and Anderson, D. (1975). The reproductive system, embryonic development, larval development and metamorphosis of the sea urchin *Heliocidaris erythrogramma* (Val.) (Echinoidea : Echinometridae). *Aust. J. Zool.* **23**, 371-403. doi:10.1071/ZO9750371
- Wilson, K. A., Andrews, M. E., Turner, F. R. and Raff, R. A. (2005). Major regulatory factors in the evolution of development: the roles of goosecoid and Msx in the evolution of the direct-developing sea urchin *Heliocidaris erythrogramma*. *Evol. Dev.* **7**, 416-428. doi:10.1111/j.1525-142X.2005.05046.x
- Wray, G. A. (1996). Parallel evolution of nonfeeding larvae in echinoids. *Syst. Biol.* **45**, 308-322. doi:10.1093/sysbio/45.3.308
- Wray, G. A. (2022). Extreme phenotypic divergence and the evolution of development. *Curr. Top. Dev. Biol.* **146**, 79-112. doi:10.1016/bs.ctdb.2021.10.004
- Wray, G. A. and Raff, R. A. (1990). Novel origins of lineage founder cells in the direct-developing sea-urchin *heliocidaris-erythrogramma*. *Dev. Biol.* **141**, 41-54. doi:10.1016/0012-1606(90)90100-W
- Wray, G. A. and Raff, R. A. (1991). The evolution of developmental strategy in marine invertebrates. *Trends Ecol. Evol.* **6**, 45-50. doi:10.1016/0169-5347(91)90121-D
- Wu, M. Y. and Hill, C. S. (2009). TGF- β superfamily signaling in embryonic development and homeostasis. *Dev. Cell* **16**, 329-343. doi:10.1016/j.devcel.2009.02.012
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L. et al. (2021). clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**, 100141. doi:10.1016/j.xinn.2021.100141
- Yaguchi, S., Yaguchi, J., Angerer, R. C. and Angerer, L. M. (2008). A Wnt-FoxQ2-nodal pathway links primary and secondary axis specification in sea urchin embryos. *Dev. Cell* **14**, 97-107. doi:10.1016/j.devcel.2007.10.012
- Yajima, M. (2007). A switch in the cellular basis of skeletogenesis in late-stage sea urchin larvae. *Dev. Biol.* **307**, 272-281. doi:10.1016/j.ydbio.2007.04.050
- Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J. et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049. doi:10.1038/ncomms14049
- Zigler, K. S., Raff, E. C., Popodi, E., Raff, R. A. and Lessios, H. A. (2003). Adaptive evolution of bindin in the genus *Heliocidaris* is correlated with the shift to direct development. *Evolution* **57**, 2293-2302. doi:10.1111/j.0014-3820.2003.tb00241.x