Single cell RNA sequencing of the *Strongylocentrotus purpuratus* larva reveals the blueprint of major cell types and nervous system of a non-chordate deuterostome

Periklis Paganos[1], Danila Voronov[1], Jacob Musser[2], Detlev Arendt[2] and Maria I. Arnone[1*]

[1]Stazione Zoologica Anton Dohrn, Department of Biology and Evolution of Marine Organisms, Villa Comunale, 80121, Naples, Italy

[2]European Molecular Biology Laboratory, Developmental Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany

*corresponding author: miarnone@szn.it

Abstract

Identifying the molecular fingerprint of organismal cell types is key for understanding their function and evolution. Here, we use single cell RNA sequencing (scRNA-seq) to survey the cell types of the sea urchin early pluteus larva, representing an important developmental transition from non-feeding to feeding larva. We identified 21 distinct cell clusters, representing cells of the digestive, skeletal, immune, and nervous systems. Further subclustering of these revealed a highly detailed portrait of cell diversity across the larva, including the identification of 12 distinct neuronal cell types. Moreover, we corroborated co-expression of key regulatory genes previously shown to drive sea urchin gene regulatory networks, and revealed additional domains in which these regulatory networks are likely to function within the larva. Lastly, we recovered a neuronal cell type co-expressing *Pdx-1* and *Brn1/2/4*, which had previously been shown to share similar gene expression with vertebrate pancreas. Our results extend this finding, revealing twenty transcription factors shared by this population of neurons in sea urchin and vertebrate pancreatic cells. Using differential expression results from Pdx-1 knockdown experiments, we generate a draft of the Pdx-1 regulatory network in these cells, and hypothesize this network was present in an ancestral deuterostome neuron before being co-opted into the pancreas developmental lineage in vertebrates.

Introduction

Multicellular organisms consist of numerous cell types, specialized in performing different tasks that guide all aspects of their growth and survival. During embryonic development, cells go through rounds of proliferation, specification and differentiation into cell types with distinct function. The information for this developmental diversification lies in the genome and the spatio-temporal expression of regulatory genes that specifies the molecular fingerprint of a given cell type (Fu et al., 2017).

The identity of each cell type is established, controlled and maintained by distinct Gene Regulatory Networks (GRNs). GRNs are logical maps of the regulatory inputs and outputs active in a cell at a given place and time, and are enacted by transcription factors, signaling molecules and terminal differentiation genes (Davidson et al., 2003, Davidson and Erwin, 2006). GRNs have been studied in a variety of organisms ranging from plants to animals in order to analyze the gene interactions at a specific time and place during the life of an organism (Krouk et al., 2013), and have been used for understanding the relationship between genome and development (Davidson and Erwin, 2006). Therefore, understanding the genetic mechanisms that provide cell types with a specific identity, and the conservation of this identity across animal taxa, is essential for understanding cell type function and evolutionary history (Arendt, 2008, Arnone et al., 2016).

Until recently, most approaches for comparing cell types relied on the identification of molecular markers, perturbation of gene expression and fate mapping. However, technological advances in microfluidics and nucleic acid barcoding now allow the high-through-put recognition of an organism's cell types at a single cell level. In particular, single cell RNA sequencing (scRNA-seq) technology, developed during the last decade, is a powerful method used to unravel the transcriptional content of individual cells, resulting in the identification of distinct cell types in an unbiased manner (Tang et al., 2009, Klein et al., 2015). ScRNA-seq involves dissociation of an organism, organ or tissue into single cells, isolation and capture of the single cells into droplets, specific barcoding of individual mRNAs, and sequencing of transcriptomic content of each cell. Computational analysis can then identify putative cell types by clustering cells with similar transcriptional profiles.

Echinoderms are a member of the phylogenetic sister group to deuterostomes, making them an ideal model for understanding the origin and diversification of deuterostome and chordate cell types. Sea urchin embryos and larvae have also been extensively used to unravel the general mechanisms of cell type specification and differentiation during development (Cameron and Davidson, 1991, Davidson et al., 1998, McClay, 2011, Lyons et al., 2012, McClay et al., 2020). The main reason for this lays in the ease with which different cell types and biological processes can be observed in the optically transparent embryos and larvae. Among the most well-studied sea urchin cell types are those comprising the nervous (Bisgrove and Burke, 1987, Burke et al., 2006a, McClay et al., 2018), immune (Rast et al., 2006, Ho et al., 2017) and digestive systems (Annunziata et al., 2014, Annunziata and Arnone, 2014, Perillo and Arnone, 2014, Perillo et al., 2016), and of both musculature (Andrikou et al., 2013, Andrikou et al., 2015) and skeleton (Okazaki, 1965, Duloquin et al., 2007, Rafiq et al., 2012, Sun and Ettensohn, 2017). For these, the developmental origins and gene regulatory wiring has been described in great detail, making the sea urchin an ideal model for GRN comparative analyses in development and evolution (Cary et al., 2020).

Here we take advantage of the detailed characterization of the sea urchin cell types performed over the years, the available cell type specific molecular markers, and the ease with which the sea urchin larvae are dissociated into single cells, to perform scRNA-seq and generate a comprehensive atlas of sea urchin larval cell types. Our findings suggest that the larva consists of 21 genetically-distinct cell clusters, each representing a distinct cell type or set of closely related subtypes (Shekhar and Menon, 2019), which we validate

using fluorescent *in situ* hybridization (FISH) and immunohistochemistry (IHC). Based on previous studies tracing developmental lineage, we assigned cell types to specific germ layers, revealing that most transcription factors are expressed pleiotropically in cells derived from multiple germ layers, yet tend to be cell type-specific within a germ layer. In addition, we illustrate how single-cell data can help corroborate previously studied GRNs, and also reveal novel cellular domains where these GRNs are likely also activated. Lastly, we investigate neuronal diversity in the sea urchin larva, identifying 12 distinct neuronal cell types. Notably, we also recovered a unique neurosecretory type that expressed *Sp-Pdx1* and *Sp-Brn1/2/4*, which had been described previously as exhibiting a pancreatic-like gene expression signature (Perillo et al., 2018). Our results confirm and extend this pancreatic-like signature, suggesting that an ancestral neuron in early deuterostomes may have given rise to the endocrine cells in the vertebrate pancreas. Supporting this, knockdown of *Sp-Pdx1* shows it is necessary for differentiation of this pancreatic-like neuronal endocrine population, indicating it has an evolutionary conserved role as a mediator of endocrine fate.

Results

Building a cell type atlas of the sea urchin larva with single-cell transcriptomics

Sea urchin early pluteus larvae were cultured and collected at 3 dpf. We performed single-cell RNA sequencing on six samples from four independent biological replicates. Individual samples were dissociated into single cells using a gentle enzyme-free dissociation protocol and using the 10x Chromium scRNA-seq system (Figure 1A). In total, transcriptomes from 19,699 cells were included in the final analysis. To identify sea urchin larval cell types, we used Louvain graph clustering as implemented in the Seurat pipeline (see methods). This revealed 21 genetically-distinct cell clusters (Figure 1B, supplement 1A and B), each representing an individual cell type or set of closely related cell types in the early pluteus larva.

Next, we set out to explore the identity of our initial 21 cell clusters. We first assigned preliminary identities to each cluster based on the expression of previously described cell type markers: ciliary band (*Btub2*) (Harlow and Nemer, 1987), apical plate (*Hbn*) (Burke et al., 2006a), aboral ectoderm (*Spec2a*) (Yuh et al., 2001), lower oral ectoderm (*Bra*) (Wei et al., 2012), upper oral ectoderm (*Gsc*) (Wei et al., 2012), neurons (*SynB*) (Burke et al., 2006a), esophageal muscles (*Mhc*) (Andrikou et al., 2013), coelomic pouches (*Nan2*) (Juliano et al., 2010), blastocoelar cells (185/333) (Ho et al., 2017), immune cells (*Gcm*) (Materna et al., 2013), skeleton (*Msp130*) (Harkey et al., 1992), anus (*Hox11/13b*), intestine (*Cdx*), pyloric sphincter (*Pdx-1*), different stomach domains (*Chp, ManrC1a, Endo16*) (Annunziata and Arnone, 2014), exocrine pancreas-like domain (*Ptf1a*) (Perillo et al., 2016), cardiac sphincter (*Trop1*) (Yaguchi et al., 2017) and esophagus (*Brn1/2/4*) (Cole and Arnone, 2009). Further, we grouped putative cell types according to embryonic germ layer origin (Figure 1C) using knowledge from previous lineage tracing experiments (Angerer and Davidson, 1984, Cameron et al., 1987).

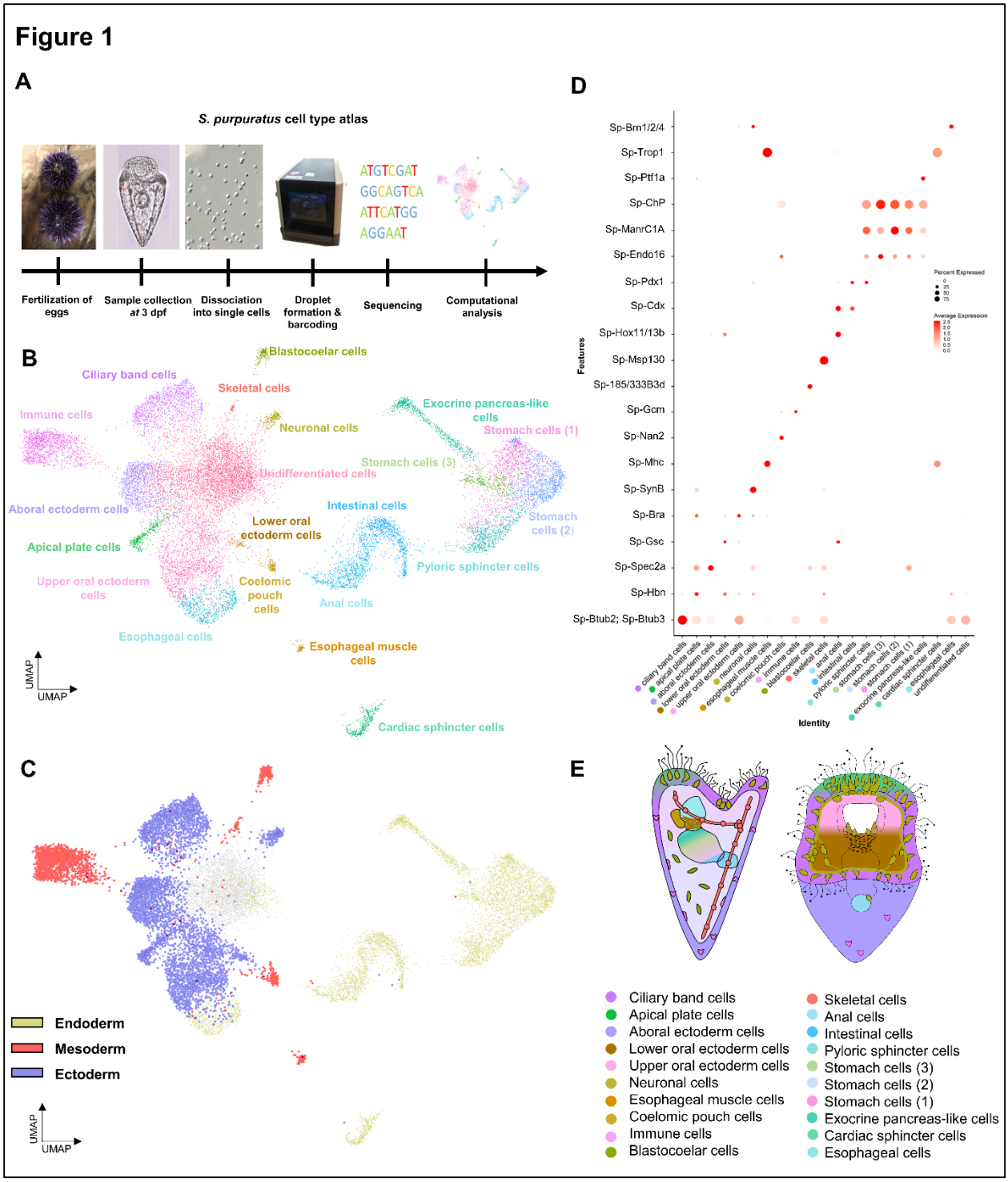


Figure 1. Cell type atlas of the 3 dpf *S. purpuratus* larva. A) Single-cell RNA sequencing pipeline from gamete fertilization to computational analysis. **B)** UMAP showing 3 dpf larval cells colored by their assignment to initial set of 21 distinct cell clusters. **C)** UMAP with cells colored by germ layer they are derived from: Endoderm (yellow), mesoderm (red), and ectoderm (blue). **D)** Dotplot of gene markers specific to cell clusters. **E)** Illustration depicting location of cell types on different larval domains. Color-code is the same as in Figure 1B.

To validate these identities, we identified all genes expressed in each cell type, totaling 15,578 genes, and performed *in situ* hybridization on a selected set of these together in combination with previously described markers. Based on this, we mapped 20 of the 21 clusters to distinct larval domains and confirmed their identity (Figure 1E). Notably, the resulting expression patterns validated the initial predictions (Figure 2- Figure supplement 2), verifying the high quality nature of the single-cell dataset. Importantly, this approach also confirmed new markers for each cell type, including *Sp-FbsL_2* (ciliary band; Figure 2A1*)*, *Sp-hypp_2386* (skeletal cells; Figure 2A9), and *Sp-Serp2/3* (exocrine pancreas-like cells; Figure 2A14). The 21st cluster, which had a poorly-defined molecular signature and lacked specific localization, likely represents undifferentiated cells (Figure 1- Figure Supplement 1C-E).

Our scRNA-seq analysis and *in situ* hybridization protocol also allowed us to identify novel expression domains for several previously described cell type markers. For instance, the transcription factors *Sp-SoxC* and *Sp-Hbn*, previously described in early neuronal specification (Garner et al., 2016, Wei et al., 2016, Yaguchi et al., 2016), were predicted by our scRNA-seq analysis to also be expressed in skeletal cells (Figure 2B & 2C1-3, 2C4-6). Confirming this, we found *Sp-Hbn* expressed in skeletal cells using *In situ* hybridization with immunostaining of PMCs along the skeletal rods (Harkey et al., 1992) (Figure2- Figure supplement 3). Likewise, the FGF signaling ligand, *Sp-Fgf9/16/20*, is known to be involved in skeletal formation and is expressed in specific populations of PMCs (Adomako-Ankomah and Ettensohn, 2014). ScRNA-seq indicates it is also expressed in oral ectoderm, cardiac sphincter, intestine and anus (Figure 2B & 2C7-9).

Lastly, we compared the limits of detection by *in situ* hybridization versus single cell RNA sequencing, using the coelomic pouch cell cluster as a case study. The coelomic pouch is derived from the mesoderm and gives rise to the rudiment and juvenile sea urchin after metamorphosis (Strathmann, 1987, Smith et al., 2008). The formation of the coelomic pouch is complex, and includes contributions from the small micromeres, a mesodermal cell population that is set aside during early development (Pehrson and Cohen, 1986, Strathmann, 1987). In an attempt to characterize this population, a study by Juliano et. al. analyzed the expression of genes involved in germ line determination and maintenance in a variety of organisms. While some well known germ line specific transcripts and proteins were found exclusively expressed in the small micromeres and the coelomic pouch of the sea urchin embryo (Juliano et al., 2006), the majority of the genes tested by *in situ* hybridization were not enriched in this cell type. Interestingly, plotting the Juliano and coauthors' gene list alongside to previously described coelomic pouch specific genes (Luo and Su, 2012, Martik and McClay, 2015) we found all candidates to be enriched in the same cell cluster (Figure 2- Figure supplement 3) suggesting the higher detection sensitivity of single cell RNA sequencing compared to the *in situ* hybridization as well as adding crucial missing information on the molecular fingerprint of such a complex cell type.

Figure 2

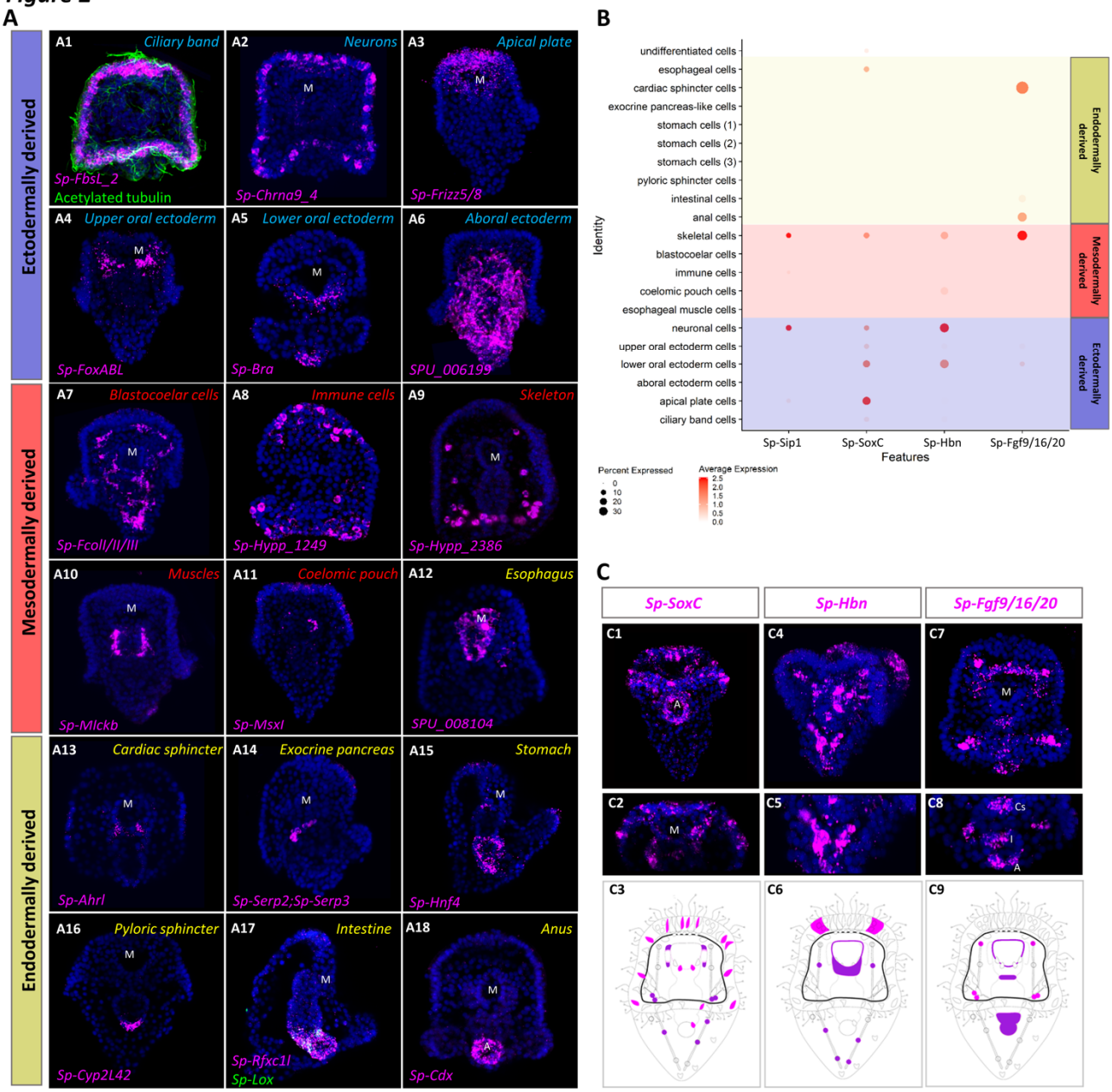


Figure 2. Validation of scRNA-seq predictions and novel expression domains.
A) FISH of *S. purpuratus* 3 dpf larvae with antisense probes for *Sp-FbsL_2* (A1), *Sp-Chrna9_4* (A2), *Sp-Frizz5/8* (A3), *Sp-FoxABL* (A4), *Sp-Bra* (A5), *SPU_006199* (A6), *Sp-Fcoll/II/III* (A7), *Sp-Hypp_1249* (A8), *Sp-Hypp_2386* (A9), *Sp-Mlckb* (A10), *Sp-Msxl* (A11), *SPU_008104* (A12), *Sp-Ahrl* (A13), *Sp-Serp2; Sp-Serp3* (A14), *Sp-Hnf4* (A15), *Sp-Cyp2L42* (A16), *Sp-Rfxc1l* (A17), *Sp-Pdx1* (A17) and *Sp-Cdx* (A18). Color-code indicates germ layer embryonic origin: endoderm (yellow), mesoderm (red), ectoderm (blue). Immunofluorescent detection of acetylated tubulin in ciliary band (green). **B)** Dotplot of *Sp-Sip1*, *Sp-SoxC*, *Sp-Hbn* and *Sp-Fgf9/16/20* expression showing previously described and novel expression domains. **C)** FISH of *S. purpuratus* 3 dpf larvae with antisense probes for *Sp-SoxC* (C1-C2), *Sp-Hbn* (C4-C5) and *Sp-Fgf9/16/20* (C7-C8). Illustrations depicting all expression domains of *Sp-SoxC* (C3), *Sp-Hbn* (C6) and *Sp-Fgf9/16/20* (C9); previously described expression domains are in magenta, newly identified ones are in

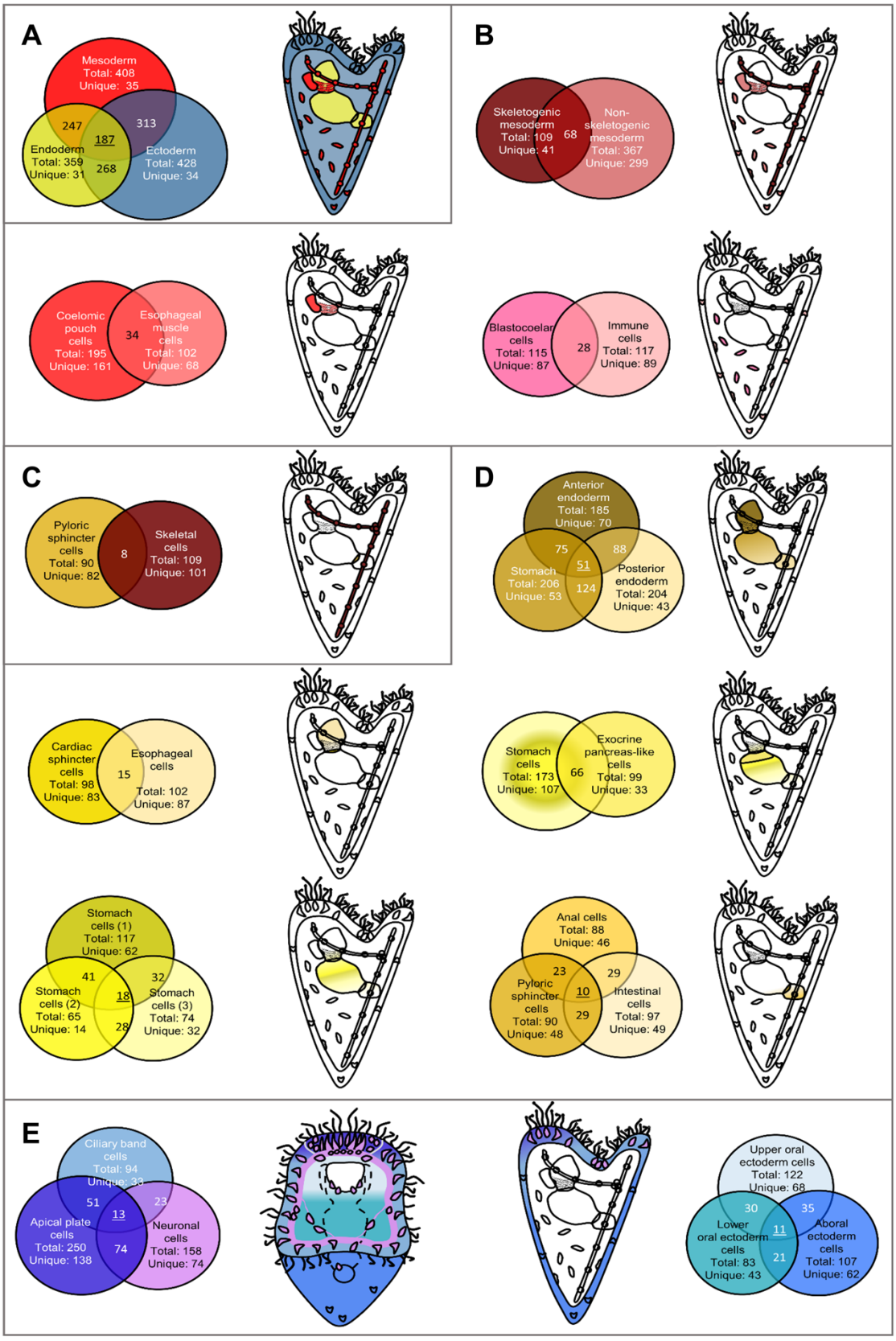
purple. Nuclei are labelled with DAPI (in blue). All images are stacks of merged confocal Z sections. A, Anus; Cs, Ca rdiac sphincter; M, Mouth; I, Intestine; Ps, Pyloric sphincter; St, Stomach.

Lastly, to determine which cells in the larva were undergoing active proliferation, we plotted expression of cell division markers in sea urchin, including *pcna*, DNA polymerases, DNA ligases, condensins, and centromere proteins (Perillo et al., 2020). The majority of cell proliferation genes were found to be enriched in the ciliary band, apical plate, coelomic pouch, immune and skeletal cell clusters (Figure 2- Figure supplement 4A). We also observed Cdk genes enriched in several endodermally derived cell types (Figure 2-Supplement 4A). Validating this, we observed S-phase cells in endodermal and skeletal cells using Edu pulse labelling (Figure 2- Figure supplement 4B). In contrast, we did not observe Edu fluorescence in cell clusters that lacked expression of proliferation markers, such as aboral ectoderm.

Identifying the different regulatory states of the larval cell types

Transcription factors, and their cooperative interactions, play a critical role in establishing and maintaining cell type identity. In order to unravel the regulatory states of early larval cell types, we performed a comparison of the transcription factors expressed in early larval cell types. First, we observed that expression programs of cell types derived from the same germ layer tend to share greater similarity with each other compared with cell types derived from other germ layers (Figure 3- Figure supplement 1). This is consistent with the finding that cell type expression programs often retain information about their developmental lineage (Sladitschek et al., 2020). Based on this, we initially compared transcription factor expression by germ layer of origin, artificially merging clusters by germ layer and excluding the cluster of undifferentiated cells. This revealed that few transcription factors are specific to cells derived from a single germ layer. Rather, a majority of regulators are expressed in derivatives of more than one germ layer (Figure 3A), and nearly one third (n = 187) are shared by cell types derived from of all three layers. Notably, mesodermal cell types share expression of more transcription factors with ectodermal than endodermal cell types, with which they are closely linked in development. Depending on the comparison, cell types derived from the same germ layer may share a great amount of TFs (1/3 in most cases), although in most comparisons the majority of transcription factors are cell type-specific in a given comparison (Figure 3B, 3D & 3E). In general, neighboring cell types and those with common developmental origins share a larger number of TFs (Figure 3D), compared to cell types with different developmental histories, such as skeletal and pyloric sphincter cells, which share only seven transcription factors in common (Figure 3C).

Figure 4



221

Figure 3. Regulatory states of the 3 dpf *S. purpuratus* larva. A) Comparison of the transcription factor content per germ layer. Venn diagram showing the shared and unique transcription factors per germ layer. Ectodermally derived cell types are shown in blue, mesodermally derived in red, and endodermally derived in yellow. **B)** Comparison of the transcription factor content across mesodermal lineages and cell types. Venn diagram showing the shared and unique transcription factors per comparison. **C)** Transcription factor content comparison of pyloric sphincter (endodermally derived) and skeletal cells (mesodermally derived), used as a negative control of our comparison. **D)** Comparison of the transcription factor content per endodermal lineage and endodermally derived cell types. Venn diagram showing the shared and unique transcription factors per comparison. **E)** TF signature comparison of ectodermally derived cell types. Venn diagram showing the shared and unique transcription factors per comparison. Cartoons indicated the relative position of each cell type/lineage. Mesodermal cell types/lineages are shown in shades of red, endodermal ones in shades of yellow and endodermal ones in shades of blue.

To further characterize the regulatory profile of larval cell types we set out to identify the expression domains of members of major TF families (Figure 4 and Figure 4 supplement 1). The *S. purpuratus* Homeobox transcription factors were first identified in a study by Howard-Ashby et al., which found most were expressed by the gastrula stage (2 dpf), and with several members expressed in domains derived from all three germ layers (Howard-Ashby et al., 2006). Our single cell analysis, although at a later developmental time point, supports these findings, and further refines our understanding of their expression to specific cell types. In the early pluteus larva, most Homeobox class transcription factors are enriched in ectodermally derived cell types, such as the apical plate and neurons. In contrast, ANTP Class and HNF class transcription factors are enriched in endodermal derivatives (Figure 4). Other major transcription factor families, such as the Forkhead, Ets, and Zinc-finger families, members of which are expressed throughout sea urchin embryogenesis (Tu et al., 2006, Rizzo et al., 2006, Materna et al., 2006), are also expressed across a spectrum of cell types. Forkhead and zinc-finger transcription factors are highly expressed in specific cell types of all three germ layer derivatives, whereas Ets family TFs are enriched in ectodermal and mesodermal derivatives (Figure 4).

The active regulatory state of a given cell type is an immediate consequence of the gene regulatory network active at this time point. Previous research in sea urchin has described in detail many regulatory networks active during embryonic and larval development. Our scRNA-seq data broadly corroborates previous studies, yet also identifies new domains and cell types in which these regulatory networks may be active. For instance, we plotted all transcription factors active in specifying coelomic pouch cells. Our data confirmed their co-expression in coelomic pouch, but also revealed they were co-expressed in the apical plate (Figure 5A). Similarly, when plotting genes involved in the aboral ectoderm gene regulatory network (Ben-Tabou de-Leon et al., 2013), we found all genes in both the aboral ectoderm cluster as well as in the apical plate cells (Figure 5B). On the other hand, plotting members of the pre-gastrula skeletogenic mesoderm regulatory network revealed most were still active in the pluteus larva and specific to skeletal cells (Figure 5C). Finally, our scRNA-seq recreates a nearly identical 3 dpf endoderm expression pattern atlas as that published previously by our group using more traditional methods (Annunziata et al., 2014), providing additional information on each gene's average expression and the percentage of cells expressing each marker (Figure 5- Figure Supplement 1).

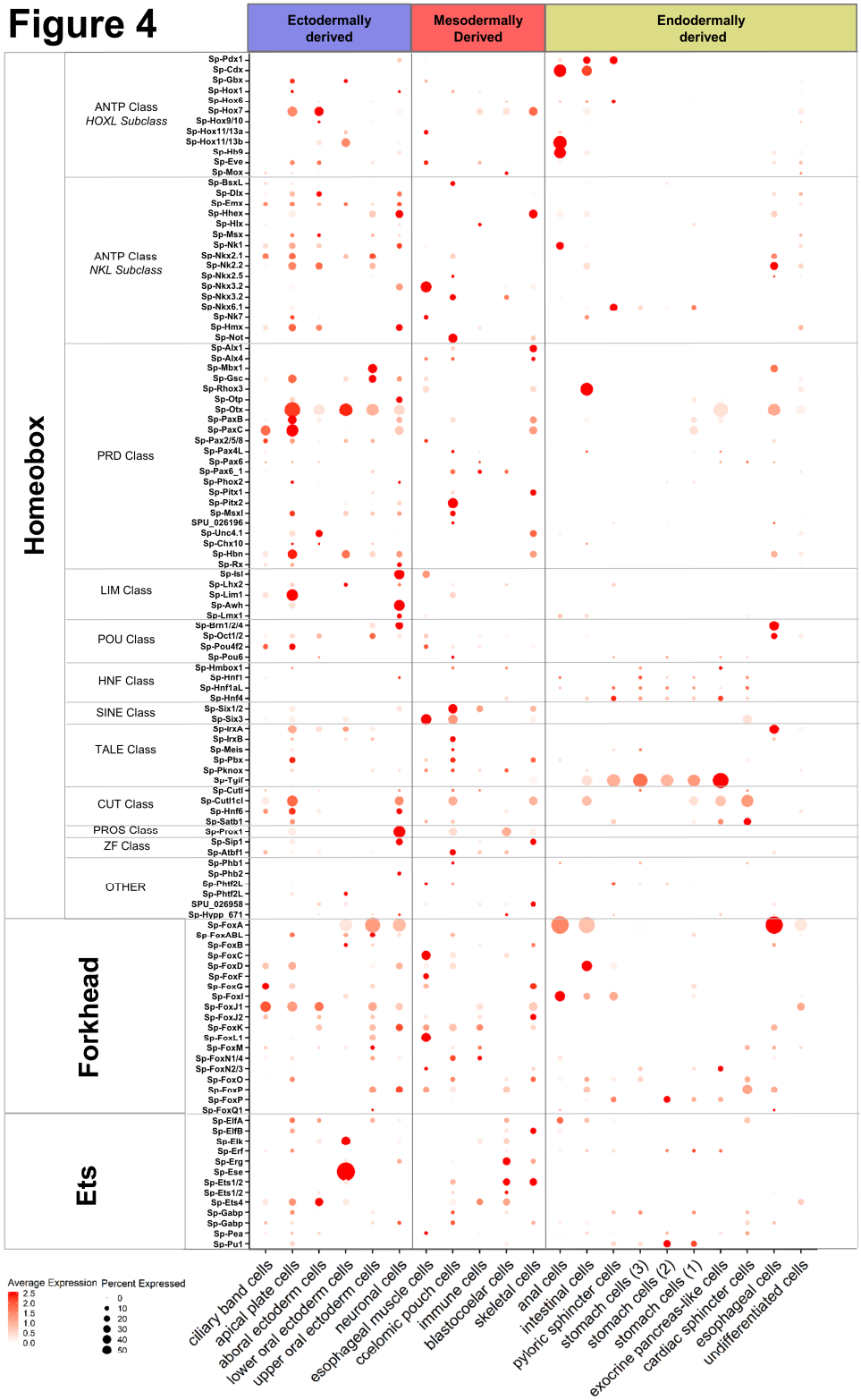
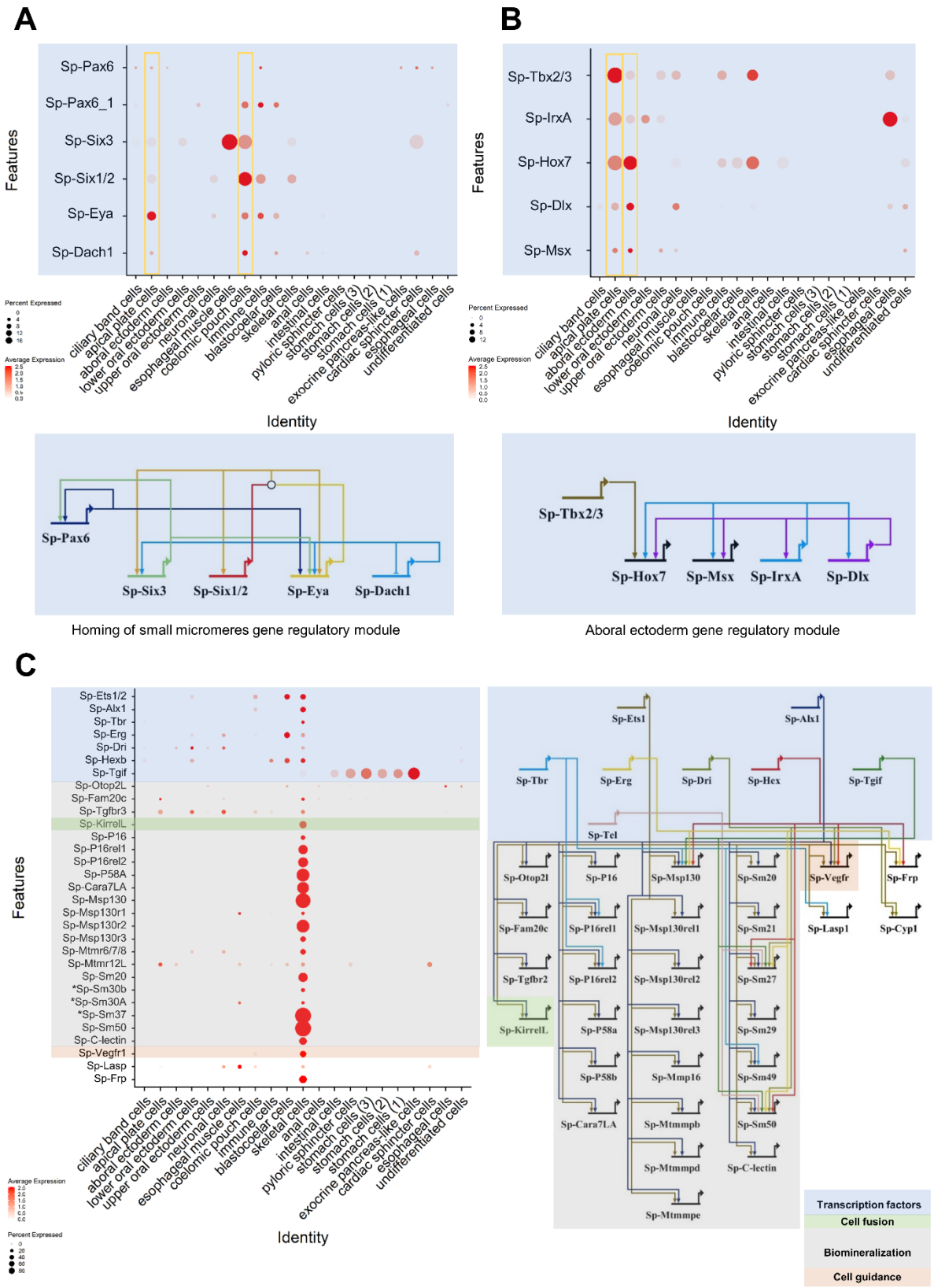


Figure 4. Localization of major transcription factor family members. Dotplot showing the average expression of members of the Homeobox, Forkhead and Ets transcription factor families. The developmental origins of each cell type are shown in blue for ectodermally derived, red for mesodermally derived and yellow for endodermally derived ones.

Figure 5



273

Figure 5. Validation of preexisting GRNs and putative novel function of specific gene regulatory modules. A) Dotplot showing the mRNA localization of genes involved in the homing of small micromeres to the coelomic pouch and novel apical plate domain. **B)** Dotplot of aboral ectoderm regulatory module genes showing novel apical plate expression. **C)** Pre-gastrula gene regulatory network enriched in skeletal cells of the sea urchin pluteus larva. Asterisks indicate larval genes involved in biomineralization, putative members of this GRN.

Unravelling the neuronal diversity and molecular signature of the nervous system

The classification of neuronal cell type diversity is an important step for unravelling the evolution and function of the nervous system. The sea urchin free swimming larva is equipped with a nervous system consisting of interconnecting ganglia (Burke et al., 2006a) that allows the animal to respond to environmental stimuli and coordinate its swimming ability (Soliman, 1983, Katow et al., 2010). Several neuronal types, including apical and ciliary band neurons, as well as neurons along the digestive tube, have been previously identified and their specification described in detail (Burke et al., 2006a, Burke et al., 2006b, Wei et al., 2009, Wei et al., 2011, Garner et al., 2016, Wei et al., 2016, McClay et al., 2018, Perillo et al., 2018, Wood et al, 2018).

Our initial clustering analysis resolved single clusters for neuronal cells, as well as for PMCs and immune cells. However, expression of known markers suggested the presence of distinct subclusters in each of these cell type groups. In order to investigate this, we independently performed subclustering and re-analysis of the neuronal, immune, and PMC cells. Subclustering of each of these initial major clusters revealed 12 neuronal, 8 immune, and 5 PMC subclusters, each likely representing distinct cell types (Figure 6A and Figure 6- Figure Supplement 1). Two of the immune subclusters expressed polyketide synthase 1 (*Sp-Pks1*), suggesting these represent sea urchin pigment cells populations (Calestani and Rogers, 2010). We also found a subcluster of immune cells that expresses the membrane attack complex/perforin family gene (*Sp-MacpfA2*), suggesting this corresponds to immune system globular cells (Figure 6-Supplement 2). Notably, our finding of 5 PMC subclusters corroborates previous reports showing five distinct groups of PMC cells along the syncytium (Sun and Ettensohn, 2014) (Figure 6-Figure Supplement 1).

To identify the 12 neuronal cell types revealed via subclustering we took advantage of the extensive previous work investigating neurogenesis and neuronal differentiation in sea urchin. Plotting known neuronal markers, we resolved unique molecular signatures for each subcluster and assigned each a putative identity and location in the larva (Figures 6B-D). To validate this, we conducted in-situ hybridization experiments for gene markers labeling these specific neuronal populations (Figure 6D), including genes encoding transcription factors (*SoxC*, *Delta*, *Ngn*, *Prox1*, *Isl*, *Hbn*, *SoxB2*, *Otx*, *NeuroD1*, *Six3*), and members of neurotransmitter (*Ddc*, *Nacha6*, *Chrna9_4*, *Tph*), and neuropeptidergic signaling pathways (*An*, *Salmfap*, *Trh*).

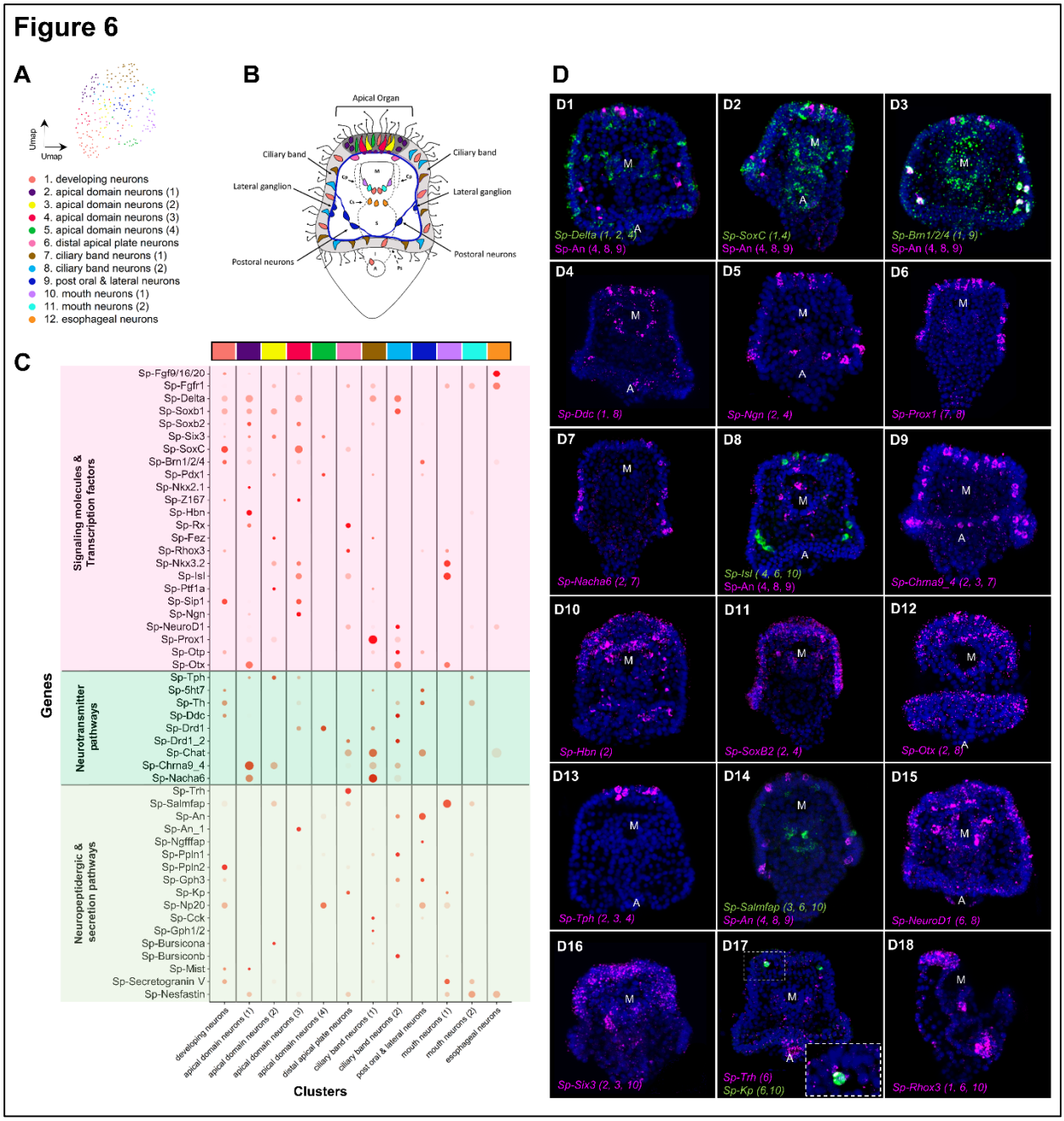


Figure 6. Neuronal complexity of the 3 dpf *S. purpuratus* larva. A) UMAP showing 12 distinct neuronal subclusters. **B)** Schematic representation of the 3 dpf pluteus larva showing the localization of neuronal subclusters (colors as in A). **C)** Dotplot of signaling molecules, transcription factors, and neurotransmitters involved in sea urchin neuronal function and neurogenesis (colors as in A). **D)** FISH of *S. purpuratus* 3 dpf larvae with antisense probes for the neuronal genes *Sp-Delta* (D1), *Sp-SoxC* (D2), *Sp-Brn1/2/4* (D3), *Sp-Ddc* (D4), *Sp-Ngn* (D5), *Sp-Prox1* (D6), *Sp-Nacha6* (D7), *Sp-Isl* (D8), *Sp-An* (D8 and D14), *Sp-Chrna9_4* (D9), *Sp-Hbn* (D10), *Sp-SoxB2* (D11), *Sp-Otx* (D12), *Sp-Tph* (A13), *Sp-Salmfap* (D14), *Sp-NeuroD1*(D15), *Sp-Six3* (D16), *Sp-Trh* (D17), *Sp-Kp* (D17) and *Sp-Rhox3* (D18). FISH shown in figures D1-3 are paired with immunohistochemical detection of the neuropeptide Sp-An. Nuclei are labelled with DAPI (in blue). All images are stacks of merged confocal Z sections. A, anus; M, mouth.

The sea urchin larva neuronal differentiation proceeds via stepwise differentiation, including transient expression of the Notch ligand Delta, followed by expression of the transcription factors *SoxC* and *Brn1/2/4* (Garner et al., 2016). During final stages of neurogenesis, the transcription factors *Sip1*, *Z167*, *Ngn* and *Otp* regulate differentiation of diverse neuronal populations, including apical and ciliary band neurons (Wei et al., 2016, McClay et al., 2018). In our data, we observed *Sp-Delta*, *Sp-SoxC* and *Sp-Brn1/2/4,* as well different combinations of the transcription factors mentioned above, co-localize in three neuronal populations (subclusters 1, 2 and 4), indicating neuronal differentiation is taking place in those three subclusters (Figure 6C). Interestingly, in one of these populations (subcluster 2) we found expression of the transcription factors *Sp-Rx*, *Sp-Hbn* (Figure 6D10) and *Sp-Six3* (Figure 6D16), which are known to be expressed in the periphery of the larva's apical domain (Burke et al., 2006a, Wei et al., 2009). This suggests that this population is located in the periphery of the apical plate and not within the apical organ. In the apical domain, we also detected a subcluster (number 6), which coexpress *Sp-Trh* and *Sp-Salmfap* neuropeptides (Wood et al, 2018), as well as *Sp-Kp* (Kissepeptin) (Figure 6D17). In total, we identified three neuronal subclusters located in the apical domain (subclusters 2, 3 and 4) of the larva that express Tryptophan hydroxylase (*Tph*), which encodes a key enzyme in the serotonin biosynthesis pathway, suggesting these represent serotonergic neurons in the larva. Within the ciliary band, which comprises the larva's peripheral nervous system (Slota et al., 2020), we identified two distinct cholinergic subclusters (7 and 8) expressing the enzyme involved in acetylcholine biosynthesis (*Sp-Chat*) (Figure 6C), one of which (subcluster 8) expresses also two nicotinic acetylcholine receptors (*Nacha6*, *Chrna9*). Moreover, we identified a neuronal subcluster in close proximity with the ciliary band, which corresponds to the lateral and post-oral neurons (subcluster 9). This population has been previously characterized by our group and was found to co-express *Sp-Pdx1*, *Sp-Brn1/2/4*, and the neuropeptide Sp-An (Perillo et al., 2018). Using gene markers that mark differentiated neurons expressed in the rim of the larva's mouth, including *Sp-Nkx3.2* (Wei et al., 2011), *Sp-Isl* (Perillo et al., 2018), the neuropeptide *Sp-Salmfap* (Wood et al, 2018), and the enzyme Tyrosine hydroxylase (*Sp-Th*) involved in the dopaminergic pathway, we identified two distinct mouth neurons subtypes (subclusters 10 and 11; Figure 6C). Lastly, we found one neuronal population that, based on its molecular signature, is associated with endodermal structures such as the esophagus (subcluster 12). Overall, our subclustering analysis increases the resolution of the different neuronal subtypes present at this developmental stage, describing new neuronal subtypes and providing novel markers and gene candidates for future studies of these cell types and their gene regulatory networks.

Characterizing a neuroendocrine neuronal population controlled by Pdx-1

Previous studies from our group suggested that the nervous system of the sea urchin larva displays a strong pre-pancreatic signature, with neurons expressing genes that are involved in endocrine cell differentiation in the vertebrate pancreas (Perillo et al., 2018). To investigate this, we focused on the post oral and lateral neuron subcluster, which co-expresses *Sp-Pdx1*, *Sp-Brn1/2/4* and *Sp-An* (Figure 6). Double immunohistochemical staining of the neuronal marker 1E11 and Sp-An shows that these neurons lie in close proximity to the ciliary band, and project axons towards both the apical plate and ciliary

band (Figure 7 A1-A3). Double FISH of *Sp-An* and *Sp-FbsL_2,* a ciliary band marker revealed by this study*,* as well as double IHC of Sp-An and acetylated tubulin (labeling cilia), further indicate their distribution relative to the ciliary band (Figure 7 A4-A6). Moreover, we found that the post-oral Sp-An neurons are found in close proximity to cells of both ciliary band subclusters (Figure 7 A7-A9), and project axons to the cell bodies of the *Sp-Prox1* positive neurons (Figure 7 A9). We also observed close proximity between An positive neurons in the post oral arms with the cells of the ventral-lateral cluster of PMCs (Figure 7 A10), and with immune globular cells (Figure 7 A11). Next, we set out to validate whether novel genes predicted by our single cell analysis to be expressed in this neuronal population can be validated *in vivo*. Among the genes predicted to be expressed in this population are the transcription factors *Sp-Nk1* and *Sp-Otp,* as well as the catecholaminergic and cholinergic neuronal markers Sp-Th and Sp-Chat (Slota and McClay, 2018), respectively. Double fluorescent *in situ* hybridization of Sp-Nk1 and the neuropeptide Sp-An, as well as fluorescent *in situ* hybridization of *Sp-Otp* combined with the immunohostochemical detection of Sp-An, reveal co-localization of these three genes in the post-oral and lateral neuronal population (Figure 7 A13-A14), verifying the single-cell data. Additionally, double immunostainings of the anti-Sp-An with anti-Th (Figure 7A9) and anti-Chat antibodies suggest that these two key enzymes, involved in different neurotransmitter biosynthesis pathways, are co-produced in the Sp-*Pdx1/Sp-Brn1/2/4* neurons (Figure 7 A12 and A15).

The vertebrate orthologue of *Sp-Pdx1* is essential for pancreas development, β-cell differentiation, and maintaining mature β-cell function (Kaneto et al., 2007). As previously described by our group, knockdown of the pancreatic transcription factor Sp-Pdx1 results in severe downregulation of the Sp-An neuropeptide, compromising the neuroendocrine fate of this neuronal type (Perillo et al., 2018). To further characterize the *Sp-Pdx1*/*Sp-Brn1/2/4* neuronal population, we performed a comprehensive analysis of genes involved in pancreatic development and β-cell differentiation, as well as gene markers related to neuroendocrine fate. We identified a total of 20 transcription factors, all involved in the formation and proper function of vertebrate endocrine pancreas, which are differentially enriched in the *Sp-Pdx1*/*Sp-Brn1/2/4* neurons (Figure 7B). Next, we intersected our scRNA-seq with bulk RNA sequencing data derived from Pdx1 morphants assayed at the same developmental stage in a previous study (Annunziata and Arnone, 2014). By coupling knowledge of cell type-specific expression programs with genes differentially expressed in the Pdx1 knockdown mutants, we were able to identify and refine likely gene targets specific to the *Sp-Pdx1*/*Sp-Brn1/2/4* neuronal cell type. In total, we found 249 genes belonging to the *Sp-Pdx1*/*Sp-Brn1/2/4* neuron subcluster (9) that were differentially expressed in the *Pdx1* knockdown dataset, with 65% of the targets being downregulated (Figure 7C). Among the downregulated genes, we found key transcription factors involved in neuronal differentiation, including *Sp-Brn1/2/4* and *Sp-Otp*, as well as terminal differentiation genes important in neuronal signaling, such as *Sp-An*, *Sp-Ngfffap*, *Sp-Th* and *Sp-Chat* (Figure 7C). Based on this, we reconstructed a provisional GRN of *Sp-Pdx1*/*Sp-Brn1/2/4* neurons, reflecting the potential role of *Sp-Pdx1* as an activator of neuroendocrine fate. Future studies are needed in order to verify these gene interactions and thus the actual connectivity of the regulatory network, although our approach highlights the power of integrating single-cell RNAseq data with data from gene knockdowns.

Figure 7

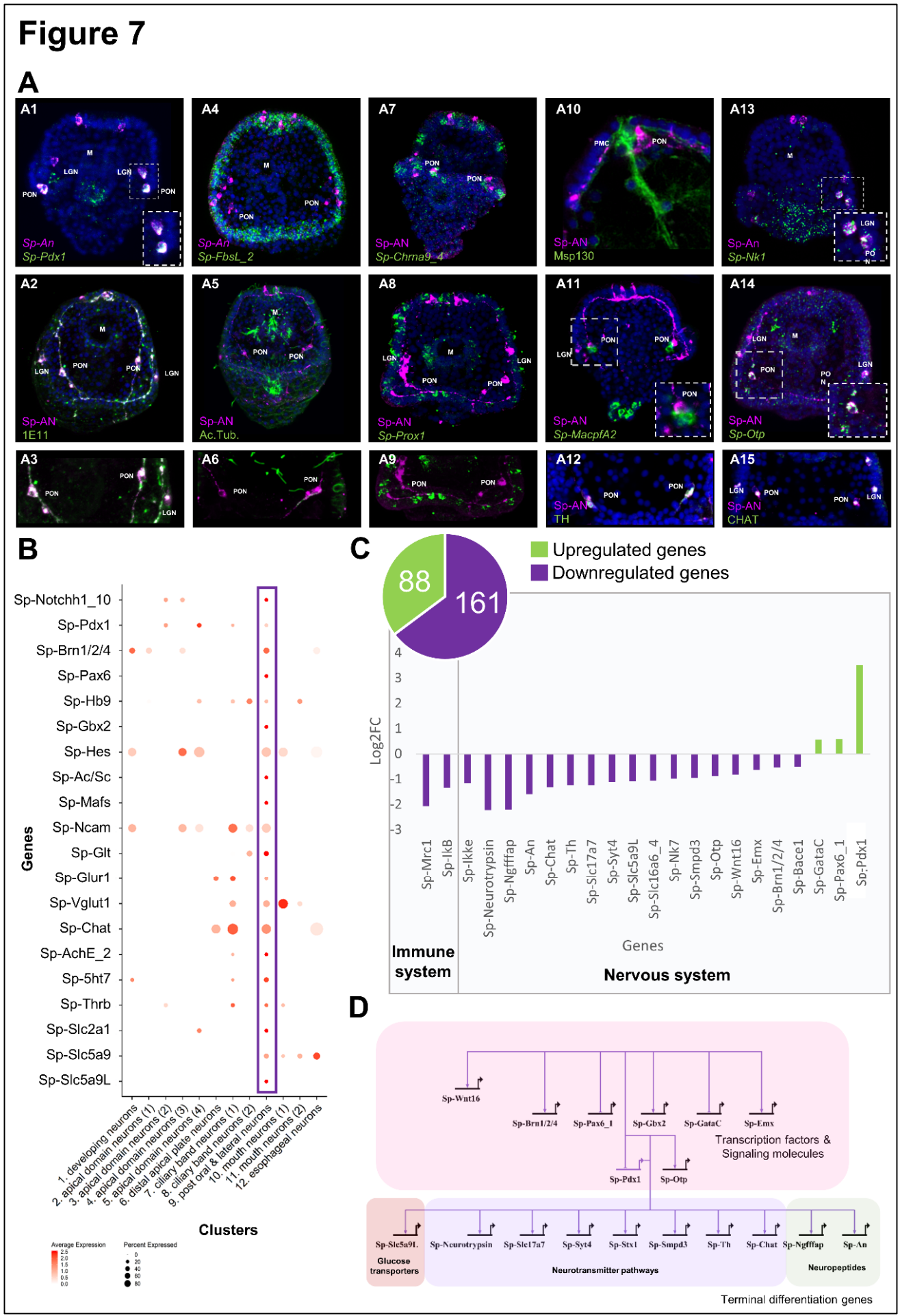


Figure 7. ScRNA-seq reveals a Pdx-1 dependent neuroendocrine cell type. **A)** Molecular characterization of a *Sp-Pdx1*/*Sp-Brn1/2/4* double positive neuronal population. A1) Double FISH of *S. purpuratus* 3 dpf larvae with specific antisense probes for *Sp-Pdx1* and *Sp-An*. A2) Double immunohistochemical detection of the Sp-An and Synaptotagmin (1E11) proteins. A3) Close up caption of the Sp-An PON neurons shown in A2. A4) Double FISH of *S. purpuratus* 3 dpf larvae with specific antisense probes for *Sp-FbsL_2* and Sp-An. Double immunohistochemical detection of the Sp-An and acetylated tubulin proteins. A6) Close up caption of the Sp-An PON neurons shown in A5. A7) FISH of *S. purpuratus* 3 dpf larvae with a specific antisense probe for *Sp-Chrna9_4* paired with immunodetection of Sp-An. A8) FISH of *S. purpuratus* 3 dpf larvae with a specific antisense probe for *Sp-Prox1* paired with immunohistochemical detection of Sp-An. A9) Close up caption of the Sp-An PON neurons shown in A8. A10) Double immunohistochemical staining for the neuropeptide Sp-An and the skeletal cells marker Msp130. A11) FISH of *S. purpuratus* 3 dpf larvae with a specific antisense probe for *Sp-MacpfA2* paired with immunohistochemical detection of Sp-An. A12) Double immunohistochemical staining for the neuropeptide Sp-An and the enzyme Sp-TH. A13) Double FISH of *S. purpuratus* 3 dpf larvae with specific antisense probes for *Sp-Nk1* and Sp-An. A14) FISH of *S. purpuratus* 3 dpf larvae with a specific antisense probe for *Sp-Otp* paired with immunohistochemical detection of Sp-An. Double immunohistochemical staining for the neuropeptide Sp-An and the enzyme Sp-Chat. Nuclei are labelled with DAPI (in blue). All images are stacks of merged confocal Z sections. LGN, lateral ganglion neurons; M, Mouth; PON, Post-oral neurons. **B)** Dotplot of genes important in endocrine pancreas differentiation and function in vertebrates. **C)** Bar plot of selected Sp-Pdx1 target genes in the *Sp-Pdx1*/*Sp-Brn1/2/4* positive population as revealed by differential RNA sequencing analysis of *Sp-Pdx1* knockdown larvae. **D)** Provisional GRN of the *Sp-Pdx1*/*Sp-Brn1/2/4* neuronal population as revealed by the combination of scRNA-seq and differential RNA-seq analysis after *Sp-Pdx1* knockdown.

Discussion

Cell type identity is determined by the differential use of genomic information among cells. Unraveling the distinct transcriptomic signatures of cell types yields valuable insight into their function, as well their evolutionary and developmental origins. In recent years, single cell transcriptomics has emerged as a powerful and unbiased approach for characterizing cell type diversity across a wide variety of animal taxa, with studies spanning insects (Davie et al., 2018, Severo et al., 2018, Cho et al., 2020), cnidarians (Sebe-Pedros et al., 2018) and sea squirts (Sharma et al., 2019, Cao et al., 2019), as well as vertebrates such as zebrafish (Wagner et al., 2018, Chestnut et al., 2020), mice (Nestorowa et al., 2016, Jung et al., 2019, Ximerakis et al., 2019, Yu et al., 2019, Qi et al., 2020) and humans (Yu et al., 2019, Qi et al., 2020, Esaulova et al., 2020, Zhao et al., 2020).

The sea urchin embryo has served as a valuable model for understanding cell type molecular specification and differentiation via gene regulatory networks. Despite this, knowledge of later stages of development, including larval cell types, is limited. Here we used single cell RNA sequencing to generate a detailed atlas of early cell types of the pluteus larva, and to unravel the neuronal diversity at this critical stage that marks the end of embryogenesis and the beginning of the larval life cycle.

Cellular diversity of the *S. purpuratus* pluteus larva

Conducting scRNA-seq on isolated *S. purpuratus* early pluteus cells we initially identified 21 genetically distinct cell clusters from 19.699 cells (Figure 1 A), expressing in total 15.578 genes. Notably, the computationally identified number of cells per cluster did not correlate well with their actual distribution in the larva. This could be a bias linked to the dissociation process, as for instance skeletal structures are the last to dissociate and hence cells can get trapped within the debris and thus be underestimated in our datasets. Nonetheless, the cell types identified in our study include all known larval cell types, suggesting we obtained a sufficient number of cells to comprehensively survey cell diversity at this developmental stage. Further, the total number of genes expressed in our data was 15.578, in relatively close agreement to the 16.500 genes expressed at the end of *S. purpuratus* embryogenesis (Tu et al., 2014).

Our results reveal a rich tapestry of cell types within the early pluteus larva. In particular, our study reveals that a majority of transcriptional diversity among cells of the pluteus larva relates to feeding and digestion. This includes two distinct oral ectoderm cell types, as well as distinct cell expression programs for the esophagus, cardiac and pyloric sphincters, exocrine pancreas-like cells, three distinct stomach cell types, intestine and anus. We also identified the mesodermally derived muscle cell type that ensures the proper function of the digestive apparatus by regulating the flow of the water containing food within the different compartments of the gut.

Beyond our initial cell clusters, we used subclustering to uncover diversity among neuronal, PMC, and immune cells in the pluteus larva. Among immune cells, the presence of two pigment cell subclusters is in line with findings by Perillo and colleagues that revealed two such populations (Perillo et al., 2020). However, in our study we found that only one cluster is *Sp-Gcm* positive, in contrast to their findings which found both of their clusters expressed *Sp-Gcm*. This difference could be a result of the different approaches used to identify the different cell types or due to transient expression of *Sp-Gcm* in the additional pigment cell type.

The larval nervous system has been among the first echinoderm cell types to be characterized at a molecular level and yet the exact number of neuronal subtypes is still not clear. Extensive work has been done on identifying the molecular pathways guiding neuronal specification and the genes active during neuronal differentiation, however this is limited to describing general neuronal categories. Most of the current information on the different neuronal types relies on detection of specific neuropeptides, neurotransmitters and enzymes involved in their biosynthesis. The most recent estimate of neuronal diversity used neuropeptidergic content to identify seven distinct neuronal types (Wood et al, 2018). Our study supports and refines this earlier work by providing a comprehensive and unbiased survey of neuronal diversity in the early pluteus larva of *S. purpuratus.*

Neurons in *S. purpuratus* arise from two ectodermal neurogenic regions (ciliary band and apical domain) and from the anteriormost part of the foregut (Garner et al., 2016, McClay et al., 2018, Wei et al., 2011), which is derived from endoderm. Different neuronal types from these domains arise at different developmental time points, although by 3 dpf most larval neurons are thought to be present and patterning diverse larval domains. Notably,

our initial clustering analysis recovered neurons as a single cluster, suggesting they share a common molecular signature regardless of their developmental origin. Subclustering of these revealed twelve distinct neuronal cell types, which we were able to trace back on the larva. By doing so, we identified one cell type apparently undergoing differentiation, four associated with the apical domain, one matching the distal apical plate neurons, two corresponding to ciliary band neurons, two located in the rim of the mouth, one associated with esophageal structures, and one corresponding to the post-oral and lateral neurons. Thus, our study greatly enhances our knowledge of neuronal diversity in the *S. purpuratus* larva, nearly doubling the number of known neuronal cell types at this developmental stage.

Lastly, we also found one major cell cluster that could not be traced back to a specific domain, and we speculate that it consists of non-differentiated cells, owing to their weak transcriptomic identity compared to the rest of the cell types (Figure1- Figure supplement 1D). Overall, this cluster exhibited greatest transcriptional similarity to cell types derived from the ectoderm (Figure 1- Figure supplement 1E, Figure 3 Figure Supplement 1), suggesting it may also be of ectodermal origin. Consistent with this, a recent study by Perillo et. al. revealed a similar uncharacterized ectodermal cell type (Perillo et al., 2020), further suggesting this cell population exists and is not solely an artifact of our analysis. Taking into account the great plasticity and regeneration capability of the sea urchin larva, it is also possible that this non-differentiated ectodermal cell type is a progenitor population in stasis, waiting to being activated. Future studies are necessary to validate its identity, function, and origin during larval development.

Sp-Pdx1 as a regulator of neuroendocrine fate

Morphogenesis and organogenesis rely on the hierarchical control of gene expression as encompassed in the GRNs. During evolution, gene regulatory elements were co-opted and incorporated to different developmental or morphogenetic programs to give rise to diverse cell types (Monteiro, 2012, Preger-Ben Noon and Frankel, 2015, Martik and McClay, 2015, Hu et al., 2018, Morgulis et al., 2019, McQueen and Rebeiz, 2020, Cary et al., 2020). It has been previously hypothesized that β pancreatic cells arose during evolution by co-option of a preexisting neuronal cell type program into the pancreas developmental lineage (Arntfield and van der Kooy, 2011, Perillo et al., 2018) based on the many physiological, morphological and molecular features endocrine pancreatic cells share with neurons (Alpert et al., 1988, Eberhard, 2013).

Although sea urchins diverged from chordates prior to the origin of the pancreas, we have previously demonstrated that the neurogenic and neuronal territories of the sea urchin embryo and larva have a strong pancreatic-like molecular signature (Perillo et al., 2018). Interestingly, one neuronal population was found to express *Sp-Pdx1* and *Sp-Brn1/2/4*, as well as the echinoderm-specific neuropeptide An. *Pdx1* in mammals is essential for proper pancreatic formation, β-cell differentiation and regulation of the mature β-cells physiology and function (Hui and Perfetti, 2002). In mice, *Pdx1* is on top of the gene regulatory hierarchy and its knockout leads to endocrine dysfunctions and absence of pancreas (Kaneto et al., 2008). In sea urchins, *Sp-Pdx1* is also found to be expressed in multiple posterior gut cell types, where it is essential for the digestive tube

compartmentalization and thus proper function (Cole et al., 2009). On the other hand, the vertebrate orthologue of *Sp-Brn1/2/4*, *Brn4* is expressed in the α pancreatic cells, where it acts as a key differentiation factor of this lineage (Hussain et al., 2002).

The gene regulatory cascade leading to the endocrine pancreas formation in vertebrates has been described in great detail and several transcription factors have been characterized as essential for this process (Zaret and Grompe, 2008, Tritschler et al., 2017). The initial steps of the differentiation of a pancreatic progenitor cell includes Notch signaling, which is involved in determining whether a pancreatic cell remains in a progenitor state or adopting an endocrine or exocrine fate. The signaling cascade results in transcriptional activation of *Hes/hairy* and enhancer of split and *Hes1* that repress *Ascl1* (Ishibashi et al., 1995, de la Pompa et al., 1997, Iso et al., 2003). The transcription factor *Pax6*, a master-gene involved in several biological processes, is a crucial element in the pancreatic cell differentiation cascade and is important for maintaining the differentiated state of the mature β cell (Hart et al., 2013, Mitchell et al., 2017, Buckle et al., 2018). On the other hand, loss of the homeobox transcription factor *Hb9* in zebrafish results in inhibition of insulin production, while loss of its mice homologue results in abolishment of the pancreatic differentiation program (Li et al., 1999, Arkhipova et al., 2012). Similarly, the transcription factor *Mafs* negatively regulates β-cell function by competing with *MafA*, a transcription factor crucial for insulin synthesis, while Gbx2 has been found to be expressed in the insulin-producing MIN6 cell lines (Mizusawa et al., 2004), in which its role remains unknown.

Here, we dissected the molecular fingerprint of the *Sp-Pdx1*/*Sp-Brn1/2/4* neuronal type and identified the presence of genes involved in pancreatic development as well as of genes known to be expressed in both endocrine pancreatic cells and neurons. From our scRNA-seq analysis it is evident that *Sp-Notch*, *Sp-Pax6*, *Sp-Hb9*, *Sp-Hes*, *Sp-Ac/Sc* (the orthologue of ASCL1), *Sp-Mafs* and the recently re-annotated *Sp-Gbx2* (previously annotated as Sp-Nk7- https://new.echinobase.org) are all expressed in the *Sp-Pdx1*/*Sp-Brn1/2/4* neurons. This suggests that these neurons of a non-chordate deuterostome have a gene regulatory machinery similar to the endocrine pancreas cells, consistent with the hypothesis that β pancreatic cells evolved from gut progenitors adopting a preexisting neuronal cell type program.

It has also been demonstrated that both endocrine pancreas and neuronal cells share similar features and are able to produce and respond to several neuronal genes and neurotransmitters. Interestingly, we were able to identify these shared components in our *Sp-Pdx1*/*Sp-Brn1/2/4* neurons. For instance, the neural cell adhesion molecule *Ncam* known to be produced in the nervous system and endocrine cells of the rat is also expressed in this cluster (Langley et al., 1989). Moreover, these neurons also express genes encoding members of the glutamate signaling pathway (*Sp-Glt*: glutamate synthase; *Sp-Vglut1*: glutamate transporter; *Sp-Glur1*: glutamate receptor), which in mammals are involved in glucose-responsive insulin secretion (Gonoi et al., 1994, Maechler and Wollheim, 1999), and tyrosine hydroxylase (*Th*), the rate-limiting enzyme of catecholamine biosynthesis, which is present in the endocrine pancreas of multiple species (Teitelman et al., 1993, Iturriza and Thibault, 1993). Furthermore, *Sp-Pdx1*/*Sp-Brn1/2/4* neurons express choline acetyltransferase (*Chat*), which has been found to be

highly expressed in human pancreatic islets and may be essential for the stimulation of insulin secretion by the neighboring β-cells (Rodriguez-Diaz et al., 2011).

Additionally, serotonergic signaling is believed to be involved in the regulation of insulin secretion as several serotonin receptors have been found to be expressed in human pancreatic islets (Amisten et al., 2015). Transcripts of the sea urchin serotonin receptor *Sp-5ht7* are present in *Sp-Pdx1*/*Sp-Brn1/2/4* neuronal population suggesting the serotonin might regulate the neurotransmitter/neuropeptide secretion in a similar way as for insulin. Similarly, transcripts of the Thyroid hormone receptor B are also found in this population, suggesting that thyroid hormone signaling may play a similar role to its differentiation similar to the one in murine endocrine pancreas differentiation (Aiello et al., 2014). The hormone secretion mediated by pancreatic endocrine cells depends on their ability to detect changes in extracellular glucose levels. To this end, they are equipped with Glucose transporters and co-transporters (Navale and Paranjape, 2016, Berger and Zdzieblo, 2020). Our analysis revealed that the *Sp-Pdx1*/*Sp-Brn1/2/4* neurons produce transcripts of three glucose co-transporter genes (*Sp-Slc2a1*, *Sp-Slc5a9* and *Sp-Slc5a9L*) proposing that they are able to detect such changes in glucose levels similarly to the endocrine pancreas cells.

Based on the significant role of all those genes in regulation, production and secretion of Insulin the question arises of whether these pancreatic-like cells in *S. purpuratus* are able to produce insulin. As previously demonstrated by our group (Perillo and Arnone, 2014) the sea urchin genome contains two genes encoding two Insulin-like peptides, Sp-ILP1 and Sp-ILP2. The gene structure of *Sp-ILP1* is evolutionarily conserved with vertebrate Insulin, whereas *Sp-ILP2* has diverged more substantially. Transcripts of *Sp-ILP1* were found localized in the gut starting from the 10 dpf pluteus larva, whereas transcripts of the divergent *Sp-ILP2* were found enriched in the coelomic pouch and esophageal structures of the 3 dpf pluteus larva (Perillo and Arnone, 2014). However, at this developmental time point (3 dpf) we were not able to detect transcripts for any of the insulin like genes in these neurons. One hypothesis is that these neurons produce Insulin only later in development, which is in line with the previous observation that Sp-Ilp1 is only found to be expressed at 10 dpf pluteus larva and onwards (Perillo and Arnone, 2014). Another hypothesis is that the *Sp-Pdx1*/*Sp-Brn1/2/4* neurons do not produce Insulin and that the Insulin regulating machinery is reutilized for different functions. One of them could be the regulation of the production and secretion of different hormones and peptides such as growth factors, neuropeptides and neurotransmitters. Based on the fact that *Sp-Pdx1* is necessary for the differentiation of these neurons as demonstrated by both the study by Perillo et al., 2018 and this study, and that these neurons are able to produce several neuromodulators (Figure 6C and Figure 7B), we favor the hypothesis that this gene regulatory machinery is used to regulate the differentiation of those neurons as well as the activity and production of neuromodulators (Sp-An, Sp-Ngfffap, Dopamine, Acetylcholine). Nonetheless, future studies are needed to shed light on whether these pancreatic-like neurons in sea urchin produce Insulin, and the exact and function and regulatory connections of pancreatic genes in the differentiation cascade of these cells.

Taken together, our data show that the *Sp-Pdx1/Sp-Brn1/2/4* neurons express key genes that are necessary for the endocrine pancreas differentiation and function. The presence of such genes in the neurons of a non-chordate deuterostome that lacks a pancreas as a

distinct organ strengthens the argument that pancreatic cells arose by redeploying a pre-existing neuronal cell type into pancreatic development. The shared features of *Sp-Pdx1/Sp-Brn1/2/4* and pancreatic cells suggest these represent features of a cell type present in the deuterostome ancestor.

Materials and Methods

1. Animal husbandry and culture of embryos

Adult *Strongylocentrotus purpuratus* individuals were obtained from Patrick Leahy (Kerckhoff Marine Laboratory, California Institute of Technology, Pasadena, CA, USA) and maintained in circulating seawater aquaria at Stazione Zoologica Anton Dohrn in Naples. Gametes were obtained by vigorous shaking of the animals. Embryos and larvae were cultured at 15°C in filtered Mediterranean Sea water diluted 9:1 with de-ionized water.

2. Larvae dissociation

Dissociation of the 3 dpf *Strongylocentrotus purpuratus* plutei into single cells was performed according to adaptation of several protocols (McClay, 1986, McClay, 2004, Juliano et al., 2014). Larvae were collected, concentrated using a 40 µm Nitex mesh filter and spun down at 500 g for 5 min. Sea water was removed and larvae were resuspended in Ca^{2+} Mg^{2+}free artificial sea water. Larvae were spun down at 500 g for 5 min and resuspended in dissociation buffer containing 1M glycine and 0.02 M EDTA in Ca^{2+} Mg^{2+}Free artificial sea water. Larvae were incubated for 10 min on ice and mixed gently via pipette aspiration every 2 min. From that point and onwards the progress of dissociation was monitored. Dissociated cells were spun down at 700 g for 5 min and washed several times with Ca^{2+} Mg^{2+}Free artificial sea water. Cell viability was assessed via using Propidium Iodide and Fluorescein diacetate and only specimens with cell viability ≥ 90% were further processed. Single cells were counted using a hemocytometer and diluted according to the manufacturer's protocol (10x Genomics). Throughout this procedure samples were kept at 4°C.

3. Single cell RNA sequencing

Single cell RNA sequencing was performed using the 10x Genomics single cell capturing system. Specimens from four independent biological replicates, ranging from 6000-20.000 cells, were loaded on the 10X Genomics Chromium Controller. Single cell cDNA libraries were prepared using the Chromium Single Cell 3' Reagent Kit (Chemistries v2 and v3). Libraries were sequenced by GeneCore (EMBL, Heidelberg, Germany) for 75 bp paired-end reads (Illumina NextSeq 500), resulting in a mean of 88M reads. Cell Ranger Software Suite 3.0.2 (10x Genomics) was used for the alignment of the single-cell RNA-seq output reads and generation of feature, barcode and matrices. The genomic index was made in Cell

Ranger using the S. purpuratus genome version 3.1 (Sea Urchin Genome Sequencing et al., 2006, Kudtarkar and Cameron, 2017). Cell Ranger output matrices for four biological and two technical replicates were used for further analysis in Seurat v3.0.2 R package (Stuart et al., 2019). The analysis was performed according to the Seurat scRNA-seq R package documentation (Butler et al., 2018, Stuart et al., 2019). Genes that are transcribed in less than three cells and cells that have less than a minimum of 200 transcribed genes were excluded from the analysis. The cutoff number of transcribed genes was determined based on feature scatter plots and varies depending on the replicate. 19.699 cells out of the 29,130 cells estimated by Cell Ranger passed the quality checks and were further analyzed. Datasets were normalized and variable genes were found using the vst method with a maximum of 2000 variable features. Data integration was performed via identification of anchors between the six different objects. Next the datasets were scaled and principal component (PCA) analysis was performed. Nearest Neighbor (SNN) graph was computed with 20 dimensions (resolution 1.0) to identify the clusters. Uniform Manifold Approximate and Projection (UMAP) was used to perform clustering dimensionality reduction. Cluster markers were found using the genes that are detected in at least 0.01 fraction of min.pct cells in the two clusters. Transcripts of all genes per cell type were identified by converting a Seurat DotPlot with all these transcripts as features into a table (ggplot2 3.2.0 R package). Subclustering analysis was performed by selecting a cell type of interest and performing similar analysis as described above. All resulting tables containing the genes transcribed within different cell types were further annotated adding PFAM terms (Trapnell et al., 2010, Finn et al., 2014) for associated proteins, gene ontology terms and descriptions from Echinobase (Kudtarkar and Cameron, 2017). Further details about the steps of the computational analysis can be found in the "Sp3dpf_clustering_analysis. Rmd" R Markdown object.

4. Whole mount RNA Fluorescent *in situ* hybridization

Fluorescent *in situ* hybridization was performed as outlined in (Perillo et al., 2021). Fluorescent signal was developed via using fluorophore conjugated tyramide technology (Perkin Elmer, Cat. #NEL752001KT). Antisense probes were transcribed form linearized DNA and labeled either during transcription via using digoxigenin-11-UTP nucleotides, or post-transcriptionally by using Fluorescein (Mirus Bio, Cat. #MIR3200) or DNP (Mirus Bio, Cat. #MIR3825) following the manufacturer's instructions. Probes for *Sp-Pdx1, Sp-Cdx, Sp-ManrC1A, Sp-Six1/2, Sp-Fgf9/16/20, Sp-Brn1/2/4, Sp-Ngn, Sp-Isl, Sp-NeuroD1, Sp-Pks1, Sp-Soxb2, Sp-An, Sp-Trh and Sp-Salmfap* were produced as previously published [*Sp-Pdx1, Sp-Cdx* (Cole et al., 2009), *Sp-ManrC1A* (Annunziata et al., 2014), *Sp-Six1/2, Sp-Fgf9/16/20* (Andrikou et al., 2015), *Sp-Brn1/2/4* (Cole and Arnone, 2009), *Sp-Ngn, Sp-Isl, Sp-NeuroD1* (Perillo et al., 2018), *Sp-Pks1* (Perillo et al, 2020), *Sp-SoxB2* (Anishchenko et al., 2018), *Sp-An, Sp-Trh, Sp-Salmfap* (Wood et al, 2018)]. Primer sequences used for cDNA isolation and probes synthesis are

in Supplementary material. Specimens were imaged using a Zeiss LSM 700 confocal microscope.

5. Immunohistochemistry (IHC)

Immunohistochemical staining or IHC paired with FISH was performed as described in (Perillo et al., 2021). Briefly 3 dpf plutei were fixed in 4 % paraformaldehyde (PFA) in filtered sea water (FSW) for 15 min at room temperature (RT). FSW was removed and samples were incubated in 100% methanol for 1 min at RT, washed multiple times with phosphate buffer saline with 0.1% Tween 20 (PBST) and incubated blocking solution containing 1 mg/ml Bovine Serum Albumin (BSA) and 4% sheep in PBST for 1h. Primary antibodies were added in the appropriate dilution and incubated for 1h and 30 min at 37°C. Anti-acetylated alpha tubulin (Sigma-Aldrich T67930) was used to label cilia and microtubules (1:200), Anti-Msp130 (gift from Dr. David R. McClay) to label skeletogenic cells (undiluted), 1E11 (gift from Dr. Robert Burke) to mark the nervous system (undiluted), 5c7 (gift from Dr. David R. McClay) to label the endoderm (undiluted), Sp-An to label the post-oral and lateral neurons (1:250), Sp-Th (Sigma-Aldrich AB152) to label catecholaminergic neurons (1:100) and Sp-Chat (GeneTex GXGTX113164S) to label cholinergic neurons. Specimens were washed multiple times with PBST and incubated for 1h with the appropriate secondary antibody (AlexaFluor) diluted 1:1000 in PBST. Larvae were washed several times with PBST and imaged using a Zeiss LSM 700 confocal microscope.

6. EdU labelling paired with immunohistochemistry

In order to understand the spatial distribution of proliferating cells across the putative broad cell types cell proliferation assays were carried out using Click-It EdU Cell Proliferation Kit for Imaging Alexa Flour^{TM} 647 (Thermo Fisher Scientific). Larvae were treated with EdU at a final concentration of 10 mM in FSW and let to grow for 2 hours. Samples were fixed in 4% PFA in FSW for 15 min (RT) and washed several times with PBST. PBST was removed, replaced by 100% Methanol for 1 min (RT) and followed by several washes with PBST. After this step one can continue with either developing the EdU signal or performing immunohistochemistry as described above. In order to develop the EdU signal the Click-iT ^{TM} reaction mix was prepared according to the manufacturer's guidelines. PBST was removed and the reaction mix was added to the samples for 30 min (RT). Larvae were washed several times with PBST, mounted and imaged using a Zeiss LSM 700 confocal microscope

7. Gene regulatory network draft

Gene regulatory modules and networks were drafted using the interactive tool for building and visualizing GRNs BioTapestry (Longabaugh, 2012).

Competing interests

The authors declare no competing interests, or appearance of competing interests, in the production and dissemination of this work.

Acknowledgements

The authors would like to thank Prof. Paola Oliveri (UCL) for providing critical input on the analysis and structure of the manuscript. We are grateful to Drs. David McClay and Robert Burke for kindly providing antibodies and Dr. Francesco Lamanna (ZMBH) for helping with the computational analysis. We also would like to thank Davide Caramiello for taking care of the adult sea urchins, Dr. Giovanna Benvenuto (SZN) for microscopy assistance and the Arnone lab members Maria Cocorullo and Inés Fournon Berodia for their help in gene cloning and preparation of RNA probes. We also thank Drs. Vladimir Benes and Bianka Baying (GeneCore, Heidelberg, Germany) for sequencing of our single cell datasets. Also we are grateful to the animal technician Emily Savage (EMBL, Heidelberg, Germany) for her precious assistance.

Funding

This project has been supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 766053 (EvoCELL: grant to MIA and DA, fellowship to PP) and the Advanced grant 'NeuralCellTypeEvo' 788921 by the European Comission (grant to DA and JM).

References

ADOMAKO-ANKOMAH, A. & ETTENSOHN, C. A. 2014. Growth factors and early mesoderm morphogenesis: insights from the sea urchin embryo. *Genesis,* 52**,** 158-72.

AIELLO, V., MORENO-ASSO, A., SERVITJA, J. M. & MARTIN, M. 2014. Thyroid hormones promote endocrine differentiation at expenses of exocrine tissue. *Exp Cell Res,* 322**,** 236-48.

AMISTEN, S., NEVILLE, M., HAWKES, R., PERSAUD, S. J., KARPE, F. & SALEHI, A. 2015. An atlas of G-protein coupled receptor expression and function in human subcutaneous adipose tissue. *Pharmacol Ther,* 146**,** 61-93.

ANDRIKOU, C., IOVENE, E., RIZZO, F., OLIVERI, P. & ARNONE, M. I. 2013. Myogenesis in the sea urchin embryo: the molecular fingerprint of the myoblast precursors. *Evodevo,* 4**,** 33.

ANDRIKOU, C., PAI, C. Y., SU, Y. H. & ARNONE, M. I. 2015. Logics and properties of a genetic regulatory program that drives embryonic muscle development in an echinoderm. *Elife,* 4.

ANGERER, R. C. & DAVIDSON, E. H. 1984. Molecular indices of cell lineage specification in sea urchin embryos. *Science,* 226**,** 1153-60.

ANISHCHENKO, E., ARNONE, M. I. & D'ANIELLO, S. 2018. SoxB2 in sea urchin development: implications in neurogenesis, ciliogenesis and skeletal patterning. *Evodevo,* 9**,** 5.

ANNUNZIATA, R., ANDRIKOU, C., PERILLO, M., CUOMO, C. & ARNONE, M. I. 2019. Development and evolution of gut structures: from molecules to function. *Cell Tissue Res,* 377**,** 445-458.

ANNUNZIATA, R. & ARNONE, M. I. 2014. A dynamic regulatory network explains ParaHox gene control of gut patterning in the sea urchin. *Development,* 141**,** 2462-72.

ANNUNZIATA, R., PERILLO, M., ANDRIKOU, C., COLE, A. G., MARTINEZ, P. & ARNONE, M. I. 2014. Pattern and process during sea urchin gut morphogenesis: the regulatory landscape. *Genesis,* 52**,** 251-68.

ARENDT, D. 2008. The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet,* 9**,** 868-82.

ARKHIPOVA, V., WENDIK, B., DEVOS, N., EK, O., PEERS, B. & MEYER, D. 2012. Characterization and regulation of the hb9/mnx1 beta-cell progenitor specific enhancer in zebrafish. *Dev Biol,* 365**,** 290-302.

ARNONE, M. I., ANDRIKOU, C. & ANNUNZIATA, R. 2016. Echinoderm systems for gene regulatory studies in evolution and development. *Curr Opin Genet Dev,* 39**,** 129-137.

ARNTFIELD, M. E. & VAN DER KOOY, D. 2011. beta-Cell evolution: How the pancreas borrowed from the brain: The shared toolbox of genes expressed by neural and pancreatic endocrine cells may reflect their evolutionary relationship. *Bioessays,* 33**,** 582-7.

BEN-TABOU DE-LEON, S., SU, Y. H., LIN, K. T., LI, E. & DAVIDSON, E. H. 2013. Gene regulatory control in the sea urchin aboral ectoderm: spatial initiation, signaling inputs, and cell fate lockdown. *Dev Biol,* 374**,** 245-54.

BERGER, C. & ZDZIEBLO, D. 2020. Glucose transporters in pancreatic islets. *Pflugers Arch,* 472**,** 1249-1272.

BISGROVE, B. W. & BURKE, R. D. 1987. Development of the Nervous-System of the Pluteus Larva of Strongylocentrotus-Droebachiensis. *Cell and Tissue Research,* 248**,** 335-343.

BUCKLE, A., NOZAWA, R. S., KLEINJAN, D. A. & GILBERT, N. 2018. Functional characteristics of novel pancreatic Pax6 regulatory elements. *Hum Mol Genet,* 27**,** 3434-3448.

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., BRANDHORST, B. P., ROWE, M., WILSON, K., CHURCHER, A. M., TAYLOR, J. S., CHEN, N., MURRAY, G., WANG, D., MELLOTT, D., OLINSKI, R., HALLBOOK, F. & THORNDYKE, M. C. 2006a. A genomic view of the sea urchin nervous system. *Dev Biol,* 300**,** 434-60.

BURKE, R. D., MOLLER, D. J., KRUPKE, O. A. & TAYLOR, V. J. 2014. Sea urchin neural development and the metazoan paradigm of neurogenesis. *Genesis,* 52**,** 208-21.

BURKE, R. D., OSBORNE, L., WANG, D., MURABE, N., YAGUCHI, S. & NAKAJIMA, Y. 2006b. Neuron-specific expression of a synaptotagmin gene in the sea urchin Strongylocentrotus purpuratus. *J Comp Neurol,* 496**,** 244-51.

BUTLER, A., HOFFMAN, P., SMIBERT, P., PAPALEXI, E. & SATIJA, R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol,* 36**,** 411-420.

CALESTANI, C. & ROGERS, D. J. 2010. Cis-regulatory analysis of the sea urchin pigment cell gene polyketide synthase. *Dev Biol,* 340**,** 249-55.

CAMERON, R. A. & DAVIDSON, E. H. 1991. Cell type specification during sea urchin development. *Trends Genet,* 7**,** 212-8.

CAMERON, R. A., HOUGH-EVANS, B. R., BRITTEN, R. J. & DAVIDSON, E. H. 1987. Lineage and fate of each blastomere of the eight-cell sea urchin embryo. *Genes Dev,* 1**,** 75-85.

CAO, C., LEMAIRE, L. A., WANG, W., YOON, P. H., CHOI, Y. A., PARSONS, L. R., MATESE, J. C., WANG, W., LEVINE, M. & CHEN, K. 2019. Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature,* 571**,** 349-354.

CARY, G. A., MCCAULEY, B. S., ZUEVA, O., PATTINATO, J., LONGABAUGH, W. & HINMAN, V. F. 2020. Systematic comparison of sea urchin and sea star developmental gene regulatory networks explains how novelty is incorporated in early development. *Nat Commun,* 11**,** 6235.

CHESTNUT, B., CASIE CHETTY, S., KOENIG, A. L. & SUMANAS, S. 2020. Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle. *Nat Commun,* 11**,** 2796.

CHO, B., YOON, S. H., LEE, D., KORANTENG, F., TATTIKOTA, S. G., CHA, N., SHIN, M., DO, H., HU, Y., OH, S. Y., LEE, D., VIPIN MENON, A., MOON, S. J., PERRIMON, N., NAM, J. W. & SHIM, J. 2020. Single-cell transcriptome maps of myeloid blood cell lineages in Drosophila. *Nat Commun,* 11**,** 4483.

COLE, A. G. & ARNONE, M. I. 2009. Fluorescent in situ hybridization reveals multiple expression domains for SpBrn1/2/4 and identifies a unique ectodermal cell type that co-expresses the ParaHox gene SpLox. *Gene Expr Patterns,* 9**,** 324-8.

COLE, A. G., RIZZO, F., MARTINEZ, P., FERNANDEZ-SERRA, M. & 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**,** 541-9.

DAVIDSON, E. H., CAMERON, R. A. & RANSICK, A. 1998. Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. *Development,* 125**,** 3269-90.

DAVIDSON, E. H. & ERWIN, D. H. 2006. Gene regulatory networks and the evolution of animal body plans. *Science,* 311**,** 796-800.

DAVIDSON, E. H., MCCLAY, D. R. & HOOD, L. 2003. Regulatory gene networks and the properties of the developmental process. *Proc Natl Acad Sci U S A,* 100**,** 1475-80.

DAVIE, K., JANSSENS, J., KOLDERE, D., DE WAEGENEER, M., PECH, U., KREFT, L., AIBAR, S., MAKHZAMI, S., CHRISTIAENS, V., BRAVO GONZALEZ-BLAS, C., POOVATHINGAL, S., HULSELMANS, G., SPANIER, K. I., MOERMAN, T., VANSPAUWEN, B., GEURS, S., VOET, T., LAMMERTYN, J., THIENPONT, B., LIU, S., KONSTANTINIDES, N., FIERS, M., VERSTREKEN, P. & AERTS, S.

2018. A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. *Cell,* 174**,** 982-998 e20.

DE LA POMPA, J. L., WAKEHAM, A., CORREIA, K. M., SAMPER, E., BROWN, S., AGUILERA, R. J., NAKANO, T., HONJO, T., MAK, T. W., ROSSANT, J. & CONLON, R. A. 1997. Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development,* 124**,** 1139-48.

DULOQUIN, L., LHOMOND, G. & GACHE, C. 2007. Localized VEGF signaling from ectoderm to mesenchyme cells controls morphogenesis of the sea urchin embryo skeleton. *Development,* 134**,** 2293-302.

ESAULOVA, E., CANTONI, C., SHCHUKINA, I., ZAITSEV, K., BUCELLI, R. C., WU, G. F., ARTYOMOV, M. N., CROSS, A. H. & EDELSON, B. T. 2020. Single-cell RNA-seq analysis of human CSF microglia and myeloid cells in neuroinflammation. *Neurol Neuroimmunol Neuroinflamm,* 7.

ETTENSOHN, C. A. 2020. The gene regulatory control of sea urchin gastrulation. *Mech Dev***,** 103599.

FINN, R. D., BATEMAN, A., CLEMENTS, J., COGGILL, P., EBERHARDT, R. Y., EDDY, S. R., HEGER, A., HETHERINGTON, K., HOLM, L., MISTRY, J., SONNHAMMER, E. L., TATE, J. & PUNTA, M. 2014. Pfam: the protein families database. *Nucleic Acids Res,* 42**,** D222-30.

FU, X., HE, F., LI, Y., SHAHVERANOV, A. & HUTCHINS, A. P. 2017. Genomic and molecular control of cell type and cell type conversions. *Cell Regen,* 6**,** 1-7.

GARNER, S., ZYSK, I., BYRNE, G., KRAMER, M., MOLLER, D., TAYLOR, V. & BURKE, R. D. 2016. Neurogenesis in sea urchin embryos and the diversity of deuterostome neurogenic mechanisms. *Development,* 143**,** 286-97.

GONOI, T., MIZUNO, N., INAGAKI, N., KUROMI, H., SEINO, Y., MIYAZAKI, J. & SEINO, S. 1994. Functional neuronal ionotropic glutamate receptors are expressed in the non-neuronal cell line MIN6. *J Biol Chem,* 269**,** 16989-92.

HARKEY, M. A., WHITELEY, H. R. & WHITELEY, A. H. 1992. Differential expression of the msp130 gene among skeletal lineage cells in the sea urchin embryo: a three dimensional in situ hybridization analysis. *Mech Dev,* 37**,** 173-84.

HARLOW, P. & NEMER, M. 1987. Coordinate and selective beta-tubulin gene expression associated with cilium formation in sea urchin embryos. *Genes Dev,* 1**,** 1293-304.

HART, A. W., MELLA, S., MENDRYCHOWSKI, J., VAN HEYNINGEN, V. & KLEINJAN, D. A. 2013. The developmental regulator Pax6 is essential for maintenance of islet cell function in the adult mouse pancreas. *PLoS One,* 8**,** e54173.

HO, E. C., BUCKLEY, K. M., SCHRANKEL, C. S., SCHUH, N. W., HIBINO, T., SOLEK, C. M., BAE, K., WANG, G. & RAST, J. P. 2017. Perturbation of gut bacteria induces a coordinated cellular immune response in the purple sea urchin larva. *Immunol Cell Biol,* 95**,** 647.

HOWARD-ASHBY, M., MATERNA, S. C., BROWN, C. T., CHEN, L., CAMERON, R. A. & DAVIDSON, E. H. 2006. Identification and characterization of homeobox transcription factor genes in Strongylocentrotus purpuratus, and their expression in embryonic development. *Dev Biol,* 300**,** 74-89.

HU, Y., SCHMITT-ENGEL, C., SCHWIRZ, J., STROEHLEIN, N., RICHTER, T., MAJUMDAR, U. & BUCHER, G. 2018. A morphological novelty evolved by co-

941 option of a reduced gene regulatory network and gene recruitment in a beetle.
942 *Proc Biol Sci,* 285.
943 HUI, H. & PERFETTI, R. 2002. Pancreas duodenum homeobox-1 regulates pancreas
944 development during embryogenesis and islet cell function in adulthood. *Eur J*
945 *Endocrinol,* 146**,** 129-41.
946 HUSSAIN, M. A., MILLER, C. P. & HABENER, J. F. 2002. Brn-4 transcription factor
947 expression targeted to the early developing mouse pancreas induces ectopic
948 glucagon gene expression in insulin-producing beta cells. *J Biol Chem,* 277**,**
949 16028-32.
950 ISHIBASHI, M., ANG, S. L., SHIOTA, K., NAKANISHI, S., KAGEYAMA, R. &
951 GUILLEMOT, F. 1995. Targeted disruption of mammalian hairy and Enhancer of
952 split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors,
953 premature neurogenesis, and severe neural tube defects. *Genes Dev,* 9**,** 3136-
954 48.
955 ISO, T., KEDES, L. & HAMAMORI, Y. 2003. HES and HERP families: multiple effectors
956 of the Notch signaling pathway. *J Cell Physiol,* 194**,** 237-55.
957 ITURRIZA, F. C. & THIBAULT, J. 1993. Immunohistochemical investigation of tyrosine-
958 hydroxylase in the islets of Langerhans of adult mice, rats and guinea pigs.
959 *Neuroendocrinology,* 57**,** 476-80.
960 JULIANO, C., SWARTZ, S. Z. & WESSEL, G. 2014. Isolating specific embryonic cells of
961 the sea urchin by FACS. *Methods Mol Biol,* 1128**,** 187-96.
962 JULIANO, C. E., VORONINA, E., STACK, C., ALDRICH, M., CAMERON, A. R. &
963 WESSEL, G. M. 2006. Germ line determinants are not localized early in sea
964 urchin development, but do accumulate in the small micromere lineage. *Dev Biol,*
965 300**,** 406-15.
966 JULIANO, C. E., YAJIMA, M. & WESSEL, G. M. 2010. Nanos functions to maintain the
967 fate of the small micromere lineage in the sea urchin embryo. *Dev Biol,* 337**,** 220-
968 32.
969 JUNG, M., WELLS, D., RUSCH, J., AHMAD, S., MARCHINI, J., MYERS, S. R. &
970 CONRAD, D. F. 2019. Unified single-cell analysis of testis gene regulation and
971 pathology in five mouse strains. *Elife,* 8.
972 KANETO, H., MIYATSUKA, T., KAWAMORI, D., YAMAMOTO, K., KATO, K.,
973 SHIRAIWA, T., KATAKAMI, N., YAMASAKI, Y., MATSUHISA, M. & MATSUOKA,
974 T. A. 2008. PDX-1 and MafA play a crucial role in pancreatic beta-cell
975 differentiation and maintenance of mature beta-cell function. *Endocr J,* 55**,** 235-
976 52.
977 KANETO, H., MIYATSUKA, T., SHIRAIWA, T., YAMAMOTO, K., KATO, K., FUJITANI,
978 Y. & MATSUOKA, T. A. 2007. Crucial role of PDX-1 in pancreas development,
979 beta-cell differentiation, and induction of surrogate beta-cells. *Curr Med Chem,*
980 14**,** 1745-52.
981 KATOW, H., SUYEMITSU, T., OOKA, S., YAGUCHI, J., JIN-NAI, T., KUWAHARA, I.,
982 KATOW, T., YAGUCHI, S. & ABE, H. 2010. Development of a dopaminergic
983 system in sea urchin embryos and larvae. *J Exp Biol,* 213**,** 2808-19.
984 KLEIN, A. M., MAZUTIS, L., AKARTUNA, I., TALLAPRAGADA, N., VERES, A., LI, V.,
985 PESHKIN, L., WEITZ, D. A. & KIRSCHNER, M. W. 2015. Droplet barcoding for
986 single-cell transcriptomics applied to embryonic stem cells. *Cell,* 161**,** 1187-1201.

KROUK, G., LINGEMAN, J., COLON, A. M., CORUZZI, G. & SHASHA, D. 2013. Gene regulatory networks in plants: learning causality from time and perturbation. *Genome Biology,* 14.

KUDTARKAR, P. & CAMERON, R. A. 2017. Echinobase: an expanding resource for echinoderm genomic information. *Database (Oxford),* 2017.

LI, H., ARBER, S., JESSELL, T. M. & EDLUND, H. 1999. Selective agenesis of the dorsal pancreas in mice lacking homeobox gene Hlxb9. *Nat Genet,* 23**,** 67-70.

LONGABAUGH, W. J. 2012. BioTapestry: a tool to visualize the dynamic properties of gene regulatory networks. *Methods Mol Biol,* 786**,** 359-94.

LUO, Y. J. & SU, Y. H. 2012. Opposing nodal and BMP signals regulate left-right asymmetry in the sea urchin larva. *PLoS Biol,* 10**,** e1001402.

LYONS, D. C., KALTENBACH, S. L. & MCCLAY, D. R. 2012. Morphogenesis in sea urchin embryos: linking cellular events to gene regulatory network states. *Wiley Interdiscip Rev Dev Biol,* 1**,** 231-52.

MAECHLER, P. & WOLLHEIM, C. B. 1999. Mitochondrial glutamate acts as a messenger in glucose-induced insulin exocytosis. *Nature,* 402**,** 685-9.

MARTIK, M. L. & MCCLAY, D. R. 2015. Deployment of a retinal determination gene network drives directed cell migration in the sea urchin embryo. *Elife,* 4.

MASSRI, A. J., GREENSTREET, L., AFANASSIEV, A., ESCOBAR, A. B., WRAY, G. M., SCHIEBINGER, G. & MCCLAY, D. R. 2020. Developmental Single-cell transcriptomics in the Lytechinus variegatus Sea Urchin Embryo. *bioRxiv***,** 2020.11.12.380675.

MATERNA, S. C., HOWARD-ASHBY, M., GRAY, R. F. & DAVIDSON, E. H. 2006. The C2H2 zinc finger genes of Strongylocentrotus purpuratus and their expression in embryonic development. *Dev Biol,* 300**,** 108-20.

MATERNA, S. C., RANSICK, A., LI, E. & DAVIDSON, E. H. 2013. Diversification of oral and aboral mesodermal regulatory states in pregastrular sea urchin embryos. *Dev Biol,* 375**,** 92-104.

MCCLAY, D. R. 1986. Embryo dissociation, cell isolation, and cell reassociation. *Methods Cell Biol,* 27**,** 309-23.

MCCLAY, D. R. 2004. Methods for embryo dissociation and analysis of cell adhesion. *Methods Cell Biol,* 74**,** 311-29.

MCCLAY, D. R. 2011. Evolutionary crossroads in developmental biology: sea urchins. *Development,* 138**,** 2639-48.

MCCLAY, D. R., MIRANDA, E. & FEINBERG, S. L. 2018. Neurogenesis in the sea urchin embryo is initiated uniquely in three domains. *Development,* 145.

MCCLAY, D. R., WARNER, J., MARTIK, M., MIRANDA, E. & SLOTA, L. 2020. Gastrulation in the sea urchin. *Curr Top Dev Biol,* 136**,** 195-218.

MCQUEEN, E. & REBEIZ, M. 2020. On the specificity of gene regulatory networks: How does network co-option affect subsequent evolution? *Curr Top Dev Biol,* 139**,** 375-405.

MITCHELL, R. K., NGUYEN-TU, M. S., CHABOSSEAU, P., CALLINGHAM, R. M., PULLEN, T. J., CHEUNG, R., LECLERC, I., HODSON, D. J. & RUTTER, G. A. 2017. The transcription factor Pax6 is required for pancreatic beta cell identity, glucose-regulated ATP synthesis, and Ca(2+) dynamics in adult mice. *J Biol Chem,* 292**,** 8892-8906.

MIZUSAWA, N., HASEGAWA, T., OHIGASHI, I., TANAKA-KOSUGI, C., HARADA, N., ITAKURA, M. & YOSHIMOTO, K. 2004. Differentiation phenotypes of pancreatic islet beta- and alpha-cells are closely related with homeotic genes and a group of differentially expressed genes. *Gene,* 331**,** 53-63.

MONTEIRO, A. 2012. Gene regulatory networks reused to build novel traits: co-option of an eye-related gene regulatory network in eye-like organs and red wing patches on insect wings is suggested by optix expression. *Bioessays,* 34**,** 181-6.

MORGULIS, M., GILDOR, T., ROOPIN, M., SHER, N., MALIK, A., LALZAR, M., DINES, M., BEN-TABOU DE-LEON, S., KHALAILY, L. & BEN-TABOU DE-LEON, S. 2019. Possible cooption of a VEGF-driven tubulogenesis program for biomineralization in echinoderms. *Proc Natl Acad Sci U S A,* 116**,** 12353-12362.

NAVALE, A. M. & PARANJAPE, A. N. 2016. Glucose transporters: physiological and pathological roles. *Biophys Rev,* 8**,** 5-9.

NESTOROWA, S., HAMEY, F. K., PIJUAN SALA, B., DIAMANTI, E., SHEPHERD, M., LAURENTI, E., WILSON, N. K., KENT, D. G. & GOTTGENS, B. 2016. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood,* 128**,** e20-31.

OKAZAKI, K. 1965. Skeleton formation of sea urchin larvae. V. Continuous observation of the process of matrix formation. *Exp Cell Res,* 40**,** 585-96.

PEHRSON, J. R. & COHEN, L. H. 1986. The fate of the small micromeres in sea urchin development. *Dev Biol,* 113**,** 522-6.

PERILLO, M. & ARNONE, M. I. 2014. Characterization of insulin-like peptides (ILPs) in the sea urchin Strongylocentrotus purpuratus: insights on the evolution of the insulin family. *Gen Comp Endocrinol,* 205**,** 68-79.

PERILLO, M., OULHEN, N., FOSTER, S., SPURRELL, M., CALESTANI, C. & WESSEL, G. 2020. Regulation of dynamic pigment cell states at single-cell resolution. *Elife,* 9.

PERILLO, M., PAGANOS, P., MATTIELLO, T., COCURULLO, M., OLIVERI, P. & ARNONE, M. I. 2018. New Neuronal Subtypes With a "Pre-Pancreatic" Signature in the Sea Urchin Stongylocentrotus purpuratus. *Front Endocrinol (Lausanne),* 9**,** 650.

PERILLO, M., PAGANOS, P., SPURRELL, M., ARNONE, M. I. & WESSEL, G. M. 2021. Methodology for Whole Mount and Fluorescent RNA In Situ Hybridization in Echinoderms: Single, Double, and Beyond. *Methods Mol Biol,* 2219**,** 195-216.

PERILLO, M., WANG, Y. J., LEACH, S. D. & ARNONE, M. I. 2016. A pancreatic exocrine-like cell regulatory circuit operating in the upper stomach of the sea urchin Strongylocentrotus purpuratus larva. *BMC Evol Biol,* 16**,** 117.

PREGER-BEN NOON, E. & FRANKEL, N. 2015. Evolving Genital Structures: A Deep Look at Network Co-option. *Dev Cell,* 34**,** 485-6.

QI, F., QIAN, S., ZHANG, S. & ZHANG, Z. 2020. Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochem Biophys Res Commun,* 526**,** 135-140.

RAFIQ, K., CHEERS, M. S. & ETTENSOHN, C. A. 2012. The genomic regulatory control of skeletal morphogenesis in the sea urchin. *Development,* 139**,** 579-90.

RAST, J. P., SMITH, L. C., LOZA-COLL, M., HIBINO, T. & LITMAN, G. W. 2006. Genomic insights into the immune system of the sea urchin. *Science,* 314**,** 952-6.

RIZZO, F., FERNANDEZ-SERRA, M., SQUARZONI, P., ARCHIMANDRITIS, A. & ARNONE, M. I. 2006. Identification and developmental expression of the ets gene family in the sea urchin (Strongylocentrotus purpuratus). *Dev Biol,* 300**,** 35-48.

RODRIGUEZ-DIAZ, R., DANDO, R., JACQUES-SILVA, M. C., FACHADO, A., MOLINA, J., ABDULREDA, M. H., RICORDI, C., ROPER, S. D., BERGGREN, P. O. & CAICEDO, A. 2011. Alpha cells secrete acetylcholine as a non-neuronal paracrine signal priming beta cell function in humans. *Nat Med,* 17**,** 888-92.

SEA URCHIN GENOME SEQUENCING, C., SODERGREN, E., WEINSTOCK, G. M., DAVIDSON, E. H., CAMERON, R. A., GIBBS, R. A., ANGERER, R. C., ANGERER, L. M., ARNONE, M. I., BURGESS, D. R., BURKE, R. D., COFFMAN, J. A., DEAN, M., ELPHICK, M. R., ETTENSOHN, C. A., FOLTZ, K. R., HAMDOUN, A., HYNES, R. O., KLEIN, W. H., MARZLUFF, W., MCCLAY, D. R., MORRIS, R. L., MUSHEGIAN, A., RAST, J. P., SMITH, L. C., THORNDYKE, M. C., VACQUIER, V. D., WESSEL, G. M., WRAY, G., ZHANG, L., ELSIK, C. G., ERMOLAEVA, O., HLAVINA, W., HOFMANN, G., KITTS, P., LANDRUM, M. J., MACKEY, A. J., MAGLOTT, D., PANOPOULOU, G., POUSTKA, A. J., PRUITT, K., SAPOJNIKOV, V., SONG, X., SOUVOROV, A., SOLOVYEV, V., WEI, Z., WHITTAKER, C. A., WORLEY, K., DURBIN, K. J., SHEN, Y., FEDRIGO, O., GARFIELD, D., HAYGOOD, R., PRIMUS, A., SATIJA, R., SEVERSON, T., GONZALEZ-GARAY, M. L., JACKSON, A. R., MILOSAVLJEVIC, A., TONG, M., KILLIAN, C. E., LIVINGSTON, B. T., WILT, F. H., ADAMS, N., BELLE, R., CARBONNEAU, S., CHEUNG, R., CORMIER, P., COSSON, B., CROCE, J., FERNANDEZ-GUERRA, A., GENEVIERE, A. M., GOEL, M., KELKAR, H., MORALES, J., MULNER-LORILLON, O., ROBERTSON, A. J., GOLDSTONE, J. V., COLE, B., EPEL, D., GOLD, B., HAHN, M. E., HOWARD-ASHBY, M., SCALLY, M., STEGEMAN, J. J., ALLGOOD, E. L., COOL, J., JUDKINS, K. M., MCCAFFERTY, S. S., MUSANTE, A. M., OBAR, R. A., RAWSON, A. P., ROSSETTI, B. J., GIBBONS, I. R., HOFFMAN, M. P., LEONE, A., ISTRAIL, S., MATERNA, S. C., SAMANTA, M. P., STOLC, V., et al. 2006. The genome of the sea urchin Strongylocentrotus purpuratus. *Science,* 314**,** 941-52.

SEBE-PEDROS, A., SAUDEMONT, B., CHOMSKY, E., PLESSIER, F., MAILHE, M. P., RENNO, J., LOE-MIE, Y., LIFSHITZ, A., MUKAMEL, Z., SCHMUTZ, S., NOVAULT, S., STEINMETZ, P. R. H., SPITZ, F., TANAY, A. & MARLOW, H. 2018. Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. *Cell,* 173**,** 1520-1534 e20.

SEVERO, M. S., LANDRY, J. J. M., LINDQUIST, R. L., GOOSMANN, C., BRINKMANN, V., COLLIER, P., HAUSER, A. E., BENES, V., HENRIKSSON, J., TEICHMANN, S. A. & LEVASHINA, E. A. 2018. Unbiased classification of mosquito blood cells by single-cell genomics and high-content imaging. *Proc Natl Acad Sci U S A,* 115**,** E7568-E7577.

SHARMA, S., WANG, W. & STOLFI, A. 2019. Single-cell transcriptome profiling of the Ciona larval brain. *Dev Biol,* 448**,** 226-236.

SHEKHAR, K. & MENON, V. 2019. Identification of Cell Types from Single-Cell Transcriptomic Data. *Methods Mol Biol,* 1935**,** 45-77.

SLADITSCHEK, H. L., FIUZA, U. M., PAVLINIC, D., BENES, V., HUFNAGEL, L. & NEVEU, P. A. 2020. MorphoSeq: Full Single-Cell Transcriptome Dynamics Up to Gastrulation in a Chordate. *Cell,* 181**,** 922-935 e21.

SLOTA, L. A. & MCCLAY, D. R. 2018. Identification of neural transcription factors required for the differentiation of three neuronal subtypes in the sea urchin embryo. *Dev Biol,* 435**,** 138-149.

SLOTA, L. A., MIRANDA, E., PESKIN, B. & MCCLAY, D. R. 2020. Developmental origin of peripheral ciliary band neurons in the sea urchin embryo. *Dev Biol,* 459**,** 72-78.

SMITH, M. M., CRUZ SMITH, L., CAMERON, R. A. & URRY, L. A. 2008. The larval stages of the sea urchin, Strongylocentrotus purpuratus. *J Morphol,* 269**,** 713-33.

SOLIMAN, S. 1983. Pharmacological control of ciliary activity in the young sea urchin larva. Effects of monoaminergic agents. *Comp Biochem Physiol C Comp Pharmacol Toxicol,* 76**,** 181-91.

STRATHMANN, M. F. 1987. *Reproduction and development of marine invertebrates of the northern Pacific coast : data and methods for the study of eggs, embryos, and larvae,* Seattle, University of Washington Press.

STUART, T., BUTLER, A., HOFFMAN, P., HAFEMEISTER, C., PAPALEXI, E., MAUCK, W. M., 3RD, HAO, Y., STOECKIUS, M., SMIBERT, P. & SATIJA, R. 2019. Comprehensive Integration of Single-Cell Data. *Cell,* 177**,** 1888-1902 e21.

SUN, Z. & ETTENSOHN, C. A. 2014. Signal-dependent regulation of the sea urchin skeletogenic gene regulatory network. *Gene Expr Patterns,* 16**,** 93-103.

SUN, Z. & ETTENSOHN, C. A. 2017. TGF-beta sensu stricto signaling regulates skeletal morphogenesis in the sea urchin embryo. *Dev Biol,* 421**,** 149-160.

TANG, F., BARBACIORU, C., WANG, Y., NORDMAN, E., LEE, C., XU, N., WANG, X., BODEAU, J., TUCH, B. B., SIDDIQUI, A., LAO, K. & SURANI, M. A. 2009. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods,* 6**,** 377-82.

TEITELMAN, G., ALPERT, S., POLAK, J. M., MARTINEZ, A. & HANAHAN, D. 1993. Precursor cells of mouse endocrine pancreas coexpress insulin, glucagon and the neuronal proteins tyrosine hydroxylase and neuropeptide Y, but not pancreatic polypeptide. *Development,* 118**,** 1031-9.

TRAPNELL, C., WILLIAMS, B. A., PERTEA, G., MORTAZAVI, A., KWAN, G., VAN BAREN, M. J., SALZBERG, S. L., WOLD, B. J. & PACHTER, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol,* 28**,** 511-5.

TRITSCHLER, S., THEIS, F. J., LICKERT, H. & BOTTCHER, A. 2017. Systematic single-cell analysis provides new insights into heterogeneity and plasticity of the pancreas. *Mol Metab,* 6**,** 974-990.

TU, Q., BROWN, C. T., DAVIDSON, E. H. & OLIVERI, P. 2006. Sea urchin Forkhead gene family: phylogeny and embryonic expression. *Dev Biol,* 300**,** 49-62.

TU, Q., CAMERON, R. A. & DAVIDSON, E. H. 2014. Quantitative developmental transcriptomes of the sea urchin Strongylocentrotus purpuratus. *Dev Biol,* 385**,** 160-7.

WAGNER, D. E., WEINREB, C., COLLINS, Z. M., BRIGGS, J. A., MEGASON, S. G. & KLEIN, A. M. 2018. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science,* 360**,** 981-987.

WEI, Z., ANGERER, L. M. & ANGERER, R. C. 2016. Neurogenic gene regulatory pathways in the sea urchin embryo. *Development,* 143**,** 298-305.

WEI, Z., ANGERER, R. C. & ANGERER, L. M. 2011. Direct development of neurons within foregut endoderm of sea urchin embryos. *Proc Natl Acad Sci U S A,* 108**,** 9143-7.

WEI, Z., RANGE, R., ANGERER, R. & ANGERER, L. 2012. Axial patterning interactions in the sea urchin embryo: suppression of nodal by Wnt1 signaling. *Development,* 139**,** 1662-9.

WEI, Z., YAGUCHI, J., YAGUCHI, S., ANGERER, R. C. & ANGERER, L. M. 2009. The sea urchin animal pole domain is a Six3-dependent neurogenic patterning center. *Development,* 136**,** 1179-89.

WESSEL, G. M. & MCCLAY, D. R. 1987. Gastrulation in the sea urchin embryo requires the deposition of crosslinked collagen within the extracellular matrix. *Dev Biol,* 121**,** 149-65.

WOOD, N.J., T. M., WARD, E.T.M., ROWE, M.L., PERILLO, M., ARNONE, M.I. , ELPHICK, M.R. & OLIVERI, P. 2018. Neuropeptidergic systems in pluteus larvae of the sea urchin Strongylocentrotus purpuratus: neurochemical complexity in a “simple” nervous system. *Fronties in Endocrinology*.

XIMERAKIS, M., LIPNICK, S. L., INNES, B. T., SIMMONS, S. K., ADICONIS, X., DIONNE, D., MAYWEATHER, B. A., NGUYEN, L., NIZIOLEK, Z., OZEK, C., BUTTY, V. L., ISSERLIN, R., BUCHANAN, S. M., LEVINE, S. S., REGEV, A., BADER, G. D., LEVIN, J. Z. & RUBIN, L. L. 2019. Single-cell transcriptomic profiling of the aging mouse brain. *Nat Neurosci,* 22**,** 1696-1708.

YAGUCHI, J., TAKEDA, N., INABA, K. & YAGUCHI, S. 2016. Cooperative Wnt-Nodal Signals Regulate the Patterning of Anterior Neuroectoderm. *PLoS Genet,* 12**,** e1006001.

YAGUCHI, S., YAGUCHI, J. & TANAKA, H. 2017. Troponin-I is present as an essential component of muscles in echinoderm larvae. *Sci Rep,* 7**,** 43563.

YU, Z., LIAO, J., CHEN, Y., ZOU, C., ZHANG, H., CHENG, J., LIU, D., LI, T., ZHANG, Q., LI, J., YANG, X., YE, Y., HUANG, Z., LONG, X., YANG, R. & MO, Z. 2019. Single-Cell Transcriptomic Map of the Human and Mouse Bladders. *J Am Soc Nephrol,* 30**,** 2159-2176.

YUH, C. H., LI, X., DAVIDSON, E. H. & KLEIN, W. H. 2001. Correct Expression of spec2a in the sea urchin embryo requires both Otx and other cis-regulatory elements. *Dev Biol,* 232**,** 424-38.

ZARET, K. S. & GROMPE, M. 2008. Generation and regeneration of cells of the liver and pancreas. *Science,* 322**,** 1490-4.

ZHAO, J., ZHANG, S., LIU, Y., HE, X., QU, M., XU, G., WANG, H., HUANG, M., PAN, J., LIU, Z., LI, Z., LIU, L. & ZHANG, Z. 2020. Single-cell RNA sequencing reveals the heterogeneity of liver-resident immune cells in human. *Cell Discov,* 6**,** 22.