

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 Paganos¹, Danila Voronov¹, Jacob M Musser², Detlev Arendt², Maria Ina Arnone^{1*}

¹Stazione Zoologica Anton Dohrn, Department of Biology and Evolution of Marine Organisms, Naples, Italy; ²European Molecular Biology Laboratory, Developmental Biology Unit, Heidelberg, Germany

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 identify 21 distinct cell clusters, representing cells of the digestive, skeletal, immune, and nervous systems. Further subclustering of these reveal a highly detailed portrait of cell diversity across the larva, including the identification of neuronal cell types. We then validate important gene regulatory networks driving sea urchin development and reveal new domains of activity within the larval body. Focusing on neurons that co-express *Pdx-1* and *Brn1/2/4*, we identify an unprecedented number of genes shared by this population of neurons in sea urchin and vertebrate endocrine pancreatic cells. Using differential expression results from *Pdx-1* knock-down experiments, we show that *Pdx1* is necessary for the acquisition of the neuronal identity of these cells. We hypothesize that a network similar to the one orchestrated by *Pdx1* in the sea urchin neurons was active in an ancestral cell type and then inherited by neuronal and pancreatic developmental lineages in sea urchins and vertebrates.

*For correspondence:
miarnone@szn.it

Competing interest: The authors declare that no competing interests exist.

Funding: See page 22

Preprinted: 16 March 2021

Received: 17 May 2021

Accepted: 24 November 2021

Published: 25 November 2021

Reviewing Editor: Kristin Tessmar-Raible, University of Vienna, Austria

© Copyright Paganos et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editor's evaluation

This work provides a comprehensive analysis of cell state specification of a whole deuterostome organism, the sea urchin *Strongylocentrotus purpuratus*. It is also a vigorous example for the use of single-cell sequencing to identify cell type homologies across evolution. The paper is thus of significant interest to scientists within the broad fields of developmental biology and evolution, as well as to the more specific communities of researchers that use the sea urchin as a model system or those interested in employing the single-cell mRNA-sequencing technology for "non-conventional" (and marine) molecular model systems.

Introduction

Multicellular organisms consist of numerous cell types, specialized in performing different tasks that guide growth and survival. During embryonic development, cells go through rounds of proliferation, specification and differentiation into cell types with distinct functions. 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, development and evolution (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 description of distinct morphological features, that is linked to the functionality of a given cell type, 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-throughput 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 type families by clustering cells with similar transcriptional profiles. Further analysis of such cell clusters can lead to the identification of distinct cell types.

Within deuterostomes, echinoderms are a member of the phylogenetic sister group to chordates, making them an ideal model for understanding the origin and diversification of deuterostome and chordate cell types (Arnone et al., 2015). 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 lies 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 Etensohn, 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 type families. Our findings suggest that the larva consists of 21 genetically distinct cell clusters, representing distinct cell type families (Shekhar and Menon, 2019), which we validate using fluorescent in situ hybridization (FISH) and immunohistochemistry (IHC). Based on previous studies tracing developmental lineage, we assign cell type families to the specific embryonic germ layers they derived from. Comparing the transcription factor content of the grouped cell clusters reveals that most transcription factors are expressed pleiotropically independently of their germ layer origins, yet tend to be cell type family-specific within a germ layer. In addition, we illustrate how single-cell data complement and validate 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. Among these, we recover a unique neurosecretory type controlled by *Sp-Pdx1* and *Sp-Brn1/2/4* 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 three 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 10 x 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 Materials and methods). This revealed 21 genetically distinct cell clusters (Figure 1B, Figure 1—figure supplement 1A,B), each representing an individual cell type or a 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, benefiting from the unique and rich knowledge on sea urchin developmental lineages: 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; Annunziata et al., 2019), 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 the cell clusters according to their embryonic germ layer origin (Figure 1C) using knowledge from previous lineage tracing experiments (Angerer and Davidson, 1984; Cameron et al., 1987).

To validate these identities, we identified all genes expressed in each cell type, totaling in 15,578 WHL genes (transcriptome models) and 12,924 genes (SPU gene models) (Supplementary file 1), 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 of the single-cell dataset. Importantly, this approach identified various new markers for each cell type family, 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 as judged by the low number and expression level of total and marker genes combined with the lack of specific localization, likely represents not fully differentiated cells (Figure 1—figure supplement 1D,F) and we refer to this cluster as undefined.

Our scRNA-seq analysis and in situ hybridization protocol unraveled 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). 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). Using in situ hybridization and immunohistochemical detection methods we confirmed the predictions and we found *Sp-SoxC* expressed in *Sp-Fgf9/16/20* positive PMCs and *Sp-Hbn* expressed in skeletal cells located in the vertex (Figure 2 2D1-4 & 2D5-9). Moreover, we found *Sp-Fgf9/16/20* transcripts localized in distinct gut domains corresponding to the cardiac sphincter, intestinal, and anal regions (Figure 2).

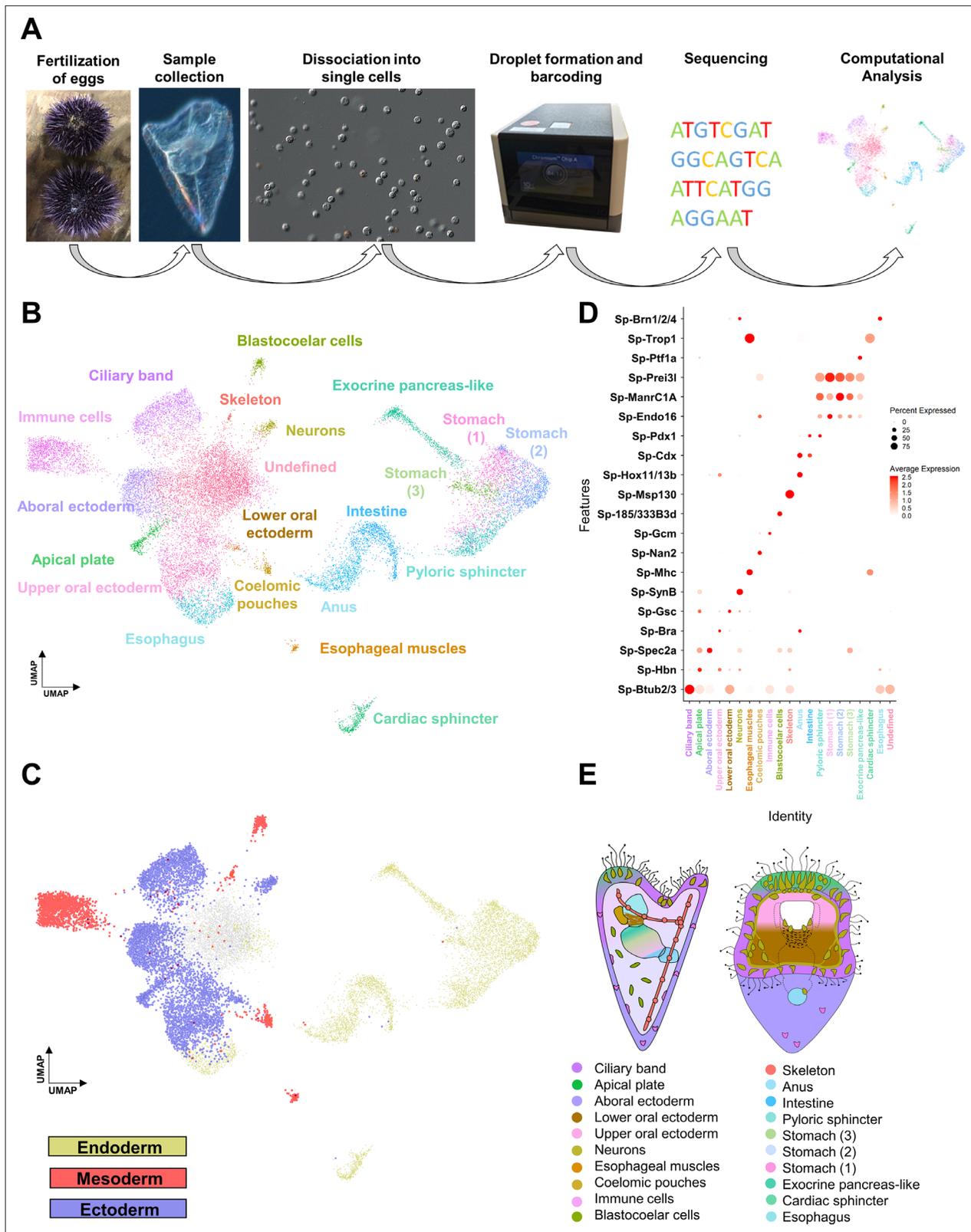


Figure 1. Cell type family atlas of the three dpf *S. purpuratus* larva. **(A)** Single-cell RNA sequencing pipeline from gamete fertilization to computational analysis. **(B)** UMAP showing three dpf larval cells colored by their assignment to the initial set of 21 distinct cell clusters. **(C)** UMAP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

Figure 1 continued on next page

Figure 1 continued

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Overlap of the different replicates and characterization of the undefined cluster.

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 pouches are derived from the mesoderm and the left coelomic pouch contributes to the formation of the rudiment and juvenile sea urchin after metamorphosis (Strathmann, 1987; Smith et al., 2008). The formation of the coelomic pouches 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). Previous attempts to characterize this population had involved screening of genes active in germline determination and maintenance in other species and revealed that, while some 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 co-authors' gene list, alongside previously described coelomic pouch specific genes (Luo and Su, 2012; Martik and McClay, 2015), they were all found in our analysis to be expressed in the same cell cluster (Figure 2—figure supplement 3). This suggests a higher sensitivity of single-cell RNA sequencing compared to the in situ hybridization, adding crucial missing information on the molecular fingerprint of a complex cell type.

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 populations (Figure 2—figure supplement 4A). Validating this, we observed S-phase cells in ciliary band, apical plate, oral ectoderm, endodermal and skeletal cells using Edu pulse labeling (Figure 2—figure supplement 4B). In contrast, we did not observe Edu fluorescence in cell populations that lacked expression of proliferation markers, such as the aboral ectoderm, suggesting that indeed they may not be proliferating.

Shared lineage information of the larval cell type families

We next compared transcription factor profiles of early larval cell type families. In general, cell populations derived from the same germ layer shared more factors with each other than with cell populations from other germ layers (Figure 3, Figure 3—figure supplement 1), consistent with the finding that cell type expression programs often retain information about their developmental lineage (Sladitschek et al., 2020). Further analysis, however, revealed very few transcription factors specific to cells from a single germ layer, contrasting with many regulators expressed in derivatives of more than one germ layer (Figure 3A), and nearly one third ($n = 187$) shared by cell type families from all three layers. Unexpectedly, mesodermal cell populations share expression of more transcription factors with ectodermal than with endodermal cell populations, even though they are more closely linked to endodermal lineage in development. We also noted that cell type families derived from the same germ layer share up to one third of transcription factors, while the majority are cell type-specific (Figure 3B, D and E). In general, neighboring cell populations and those with common developmental origins share a larger number of TFs (Figure 3D), compared to cell populations with different developmental histories (Figure 3C).

To further characterize the regulatory profile of larval cell type families we set out to identify the expression profiles of members of major transcription factor families (Figure 4, Figure 4—figure supplement 1). In *S. purpuratus*, most homeobox transcription factors were previously found expressed at the gastrula stage (two dpf), 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 type families. In the early pluteus larva, most homeobox class transcription factors are enriched in ectodermally derived cell populations, such as the apical plate and neurons. In contrast, ANTP Class and HNF class transcription factors are enriched in endodermal derivatives (Figure 4).

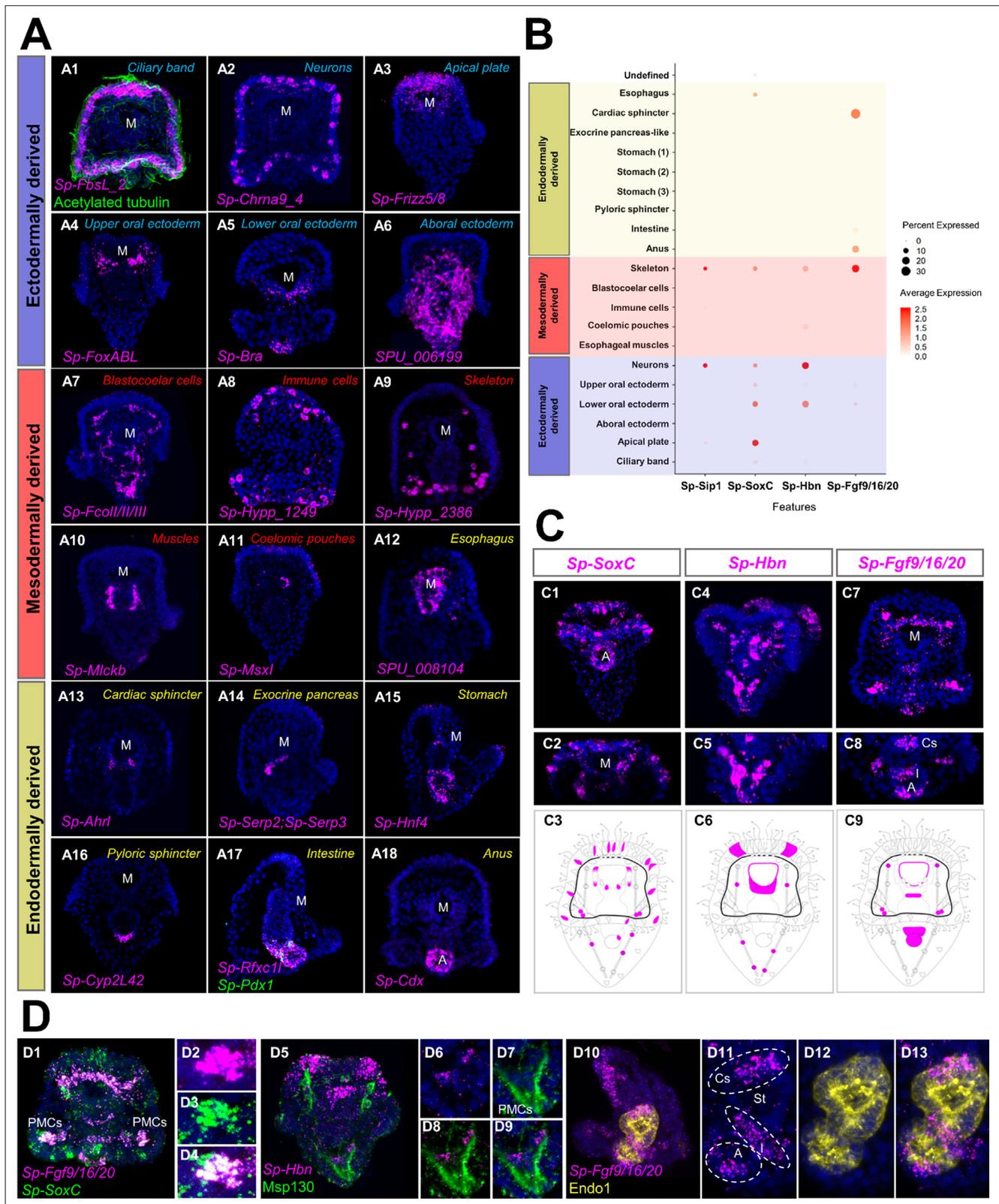


Figure 2. Validation of scRNA-seq predictions and novel expression domains. (A) FISH of *S. purpuratus* three dpf larvae with antisense probes for *Sp-FbsL_2* (A1), *Sp-Chrma9_4* (A2), *Sp-Frizz5/8* (A3), *Sp-FoxABL* (A4), *Sp-Bra* (A5), *SPU_006199* (A6), *Sp-Fcoll/II/III* (A7), *Sp-Hypp_1249* (A8), *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-Rfxc11* (A17), *Sp-Pdx1* (A17) and *Sp-Cdx* (A18). Color-code indicates germ layer embryonic origin: endoderm (yellow), mesoderm (red), ectoderm

Figure 2 continued on next page

Figure 2 continued

(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* three dpf larvae with antisense probes for *Sp-SoxC* (C1–C2), *Sp-Hbn* (C4–C5), and *Sp-Fgf9/16/20* (C7–C8). Illustrations depicting all the expression domains of *Sp-SoxC* (C3), *Sp-Hbn* (C6), and *Sp-Fgf9/16/20* (C9). Illustrations depicting all the expression domains of *Sp-SoxC* (C3), *Sp-Hbn* (C6), and *Sp-Fgf9/16/20* (C9). Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. (D) FISH of *S. purpuratus* three dpf larvae with antisense probes for *Sp-SoxC* and *Sp-Fgf9/16/20* (D1–D4), for *Sp-Hbn* combined with immunohistochemical detection for the skeletal cells marker Msp130 (D5–D9) and *Sp-Fgf9/16/20* with immunohistochemical detection for the midgut and hindgut protein Endo1 (D10–D13). A, Anus; Cs, Cardiac sphincter; I, Intestine; M, Mouth; PMCs, Primary mesenchyme cells; St, Stomach.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Expression pattern of genes used to annotate the clusters.

Figure supplement 2. scRNA-seq predicted expression patterns of genes used to annotate the clusters.

Figure supplement 3. scRNA-seq is able to detect expression patterns, previously undetectable by in situ hybridization.

Figure supplement 4. Proliferation status and dynamics of the larval cell type families.

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 populations. Forkhead and zinc-finger transcription factors are highly expressed in specific cell type families 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 family 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 is broadly consistent with previous studies, yet also identifies new domains and cell populations in which these regulatory networks may be active. For instance, we plotted all transcription factors known to be active in specifying coelomic pouch cells (Martik and McClay, 2015). Our data confirmed their co-expression in coelomic pouch, but also revealed their co-expression in the apical plate (Figure 5A). Similarly, 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 (Figure 5C) revealed most were still active in the pluteus larva and specific to skeletal cells (Oliveri et al., 2008; Shashikant and Etensohn, 2019; Khor et al., 2019; Etensohn, 2020). These findings illustrate the immediate benefit of our dataset to drastically expand our knowledge of larval regulatory networks. Finally, our scRNA-seq recreates a nearly identical three 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 level and the percentage of cells expressing each marker (Figure 5—figure supplement 1).

Unravelling the neuronal diversity and molecular signature of the nervous system

The assessment 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 interconnected ganglia (Burke et al., 2006a) that allows the animal to respond to environmental stimuli and coordinate its swimming behavior (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; Burke et al., 2014; 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 family 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 6, Figure 6—figure supplements 1 and 2). Two of the immune subclusters

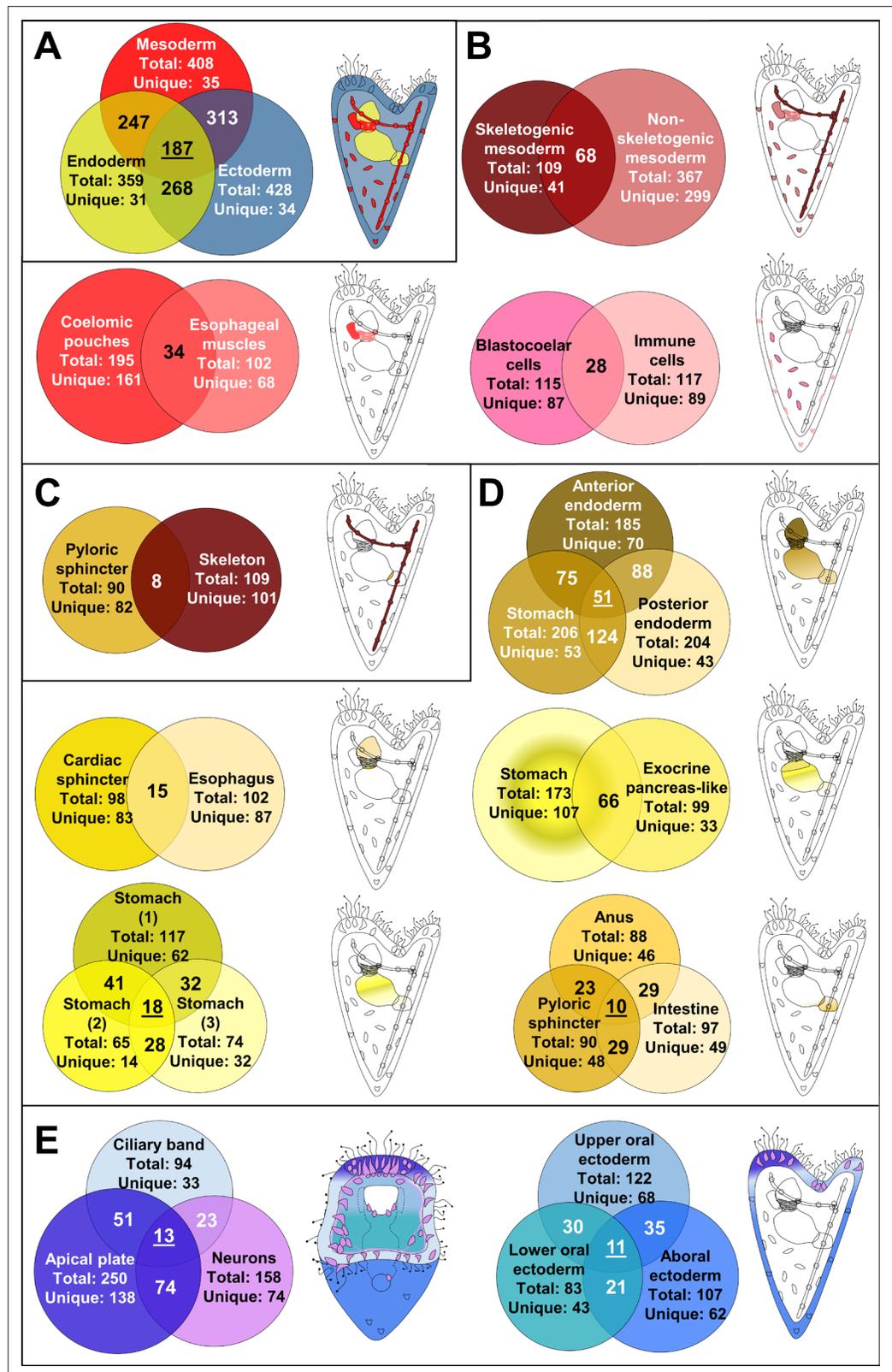


Figure 3. Regulatory states of the three dpf *S. purpuratus* larva. (A) Comparison of the transcription factor content per germ layer. Venn diagram showing the shared and unique transcription factors per germ layer. Ectodermally derived cell type families are shown in blue, mesodermally derived in red, and endodermally derived in yellow. (B) Comparison of the transcription factor content across mesodermal lineages and cell type families. Venn diagram

Figure 3 continued on next page

Figure 3 continued

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 type families. Venn diagram showing the shared and unique transcription factors per comparison. (E) TF signature comparison of ectodermally derived cell type families. Venn diagram showing the shared and unique transcription factors per comparison. Cartoons indicated the relative position of each cell type family/lineage. Mesodermal cell type families/lineages are shown in shades of red, endodermal ones in shades of yellow and endodermal ones in shades of blue.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Cell type family trees of the three dpf pluteus larva.

expressed polyketide synthase 1 (*Sp-Pks1*), suggesting these represent sea urchin pigment cell 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—figure supplement 2). Notably, our finding of 5 PMC subclusters corroborates previous reports showing five distinct groups of PMC cells along the syncytium (Sun and Etnensohn, 2014; Figure 6—figure supplement 1).

To annotate 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 neuronal markers, we resolved unique molecular signatures for each subcluster and assigned each a putative identity and location in the larva (Figure 6B–D). To validate this, we then 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*), as well as members of neurotransmitter (*Ddc*, *Nacha6*, *Chrna9_4*, *Tph*), and neuro-peptidergic signaling pathways (*An*, *Salmfap*, *Trh*).

The sea urchin larva neuronal differentiation proceeds stepwise, 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 the 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, and in particular the co-presence of *Sp-Fgfr1* and *Sp-Fgf9/16/20* (Figure 6—figure

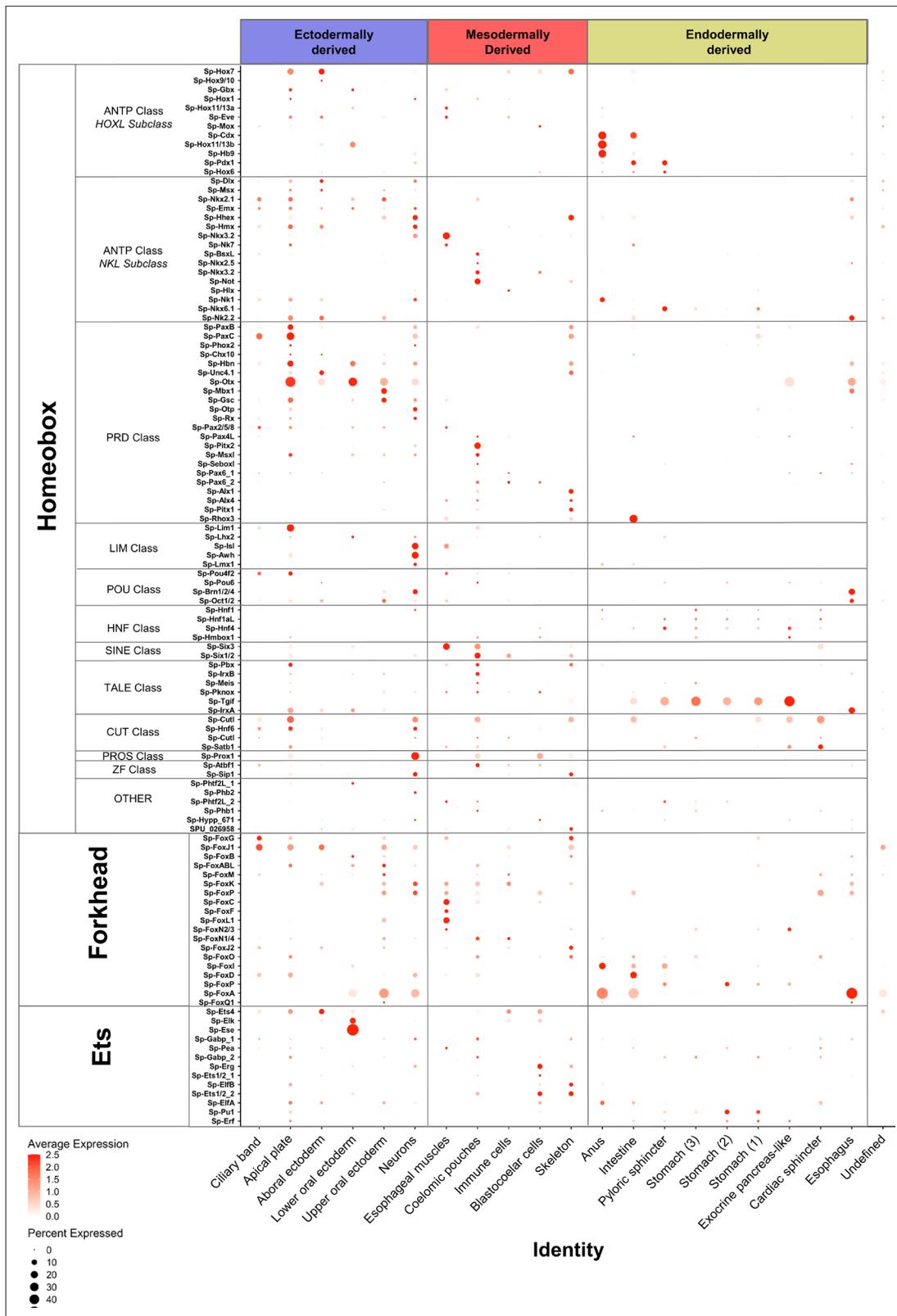


Figure 4. Localization of major transcription factor family members. Dotplot showing the average scaled expression of members of the Homeobox, Forkhead and Ets transcription factor families. The developmental origins of each cell type family are shown in blue for ectodermally derived, red for mesodermally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

Figure 4 continued on next page

Figure 4 continued

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Distribution of the Zinc finger TF family members across the cell type families.

Figure supplement 2. Differentially expressed TFs across cell type families.

supplement 3), could correspond to neurons associated with the cardiac sphincter within the esophageal endoderm (subcluster 12). Overall, our subclustering analysis shows the different neuronal subtypes present at this developmental stage at an increased resolution, 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 toward both the apical plate and ciliary band (Figure 7A1-A3). Double FISH of *Sp-An* and *Sp-FbsL_2*, a ciliary band marker revealed by this study, as well as double IHC of anti-*Sp-An* and anti-acetylated tubulin (labeling cilia), further indicate their distribution relative to the ciliary band (Figure 7A4-A6). Moreover, we found that the post-oral *Sp-An* neurons are found in close proximity to cells of both ciliary band subclusters (Figure 7A7-A9), and project axons to the cell bodies of the *Sp-Prox1* positive neurons (Figure 7A9). 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 7A10), and with immune globular cells (Figure 7A11). 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 immunohistochemical detection of *Sp-An*, reveal co-localization of these three genes in the post-oral and lateral neuronal population (Figure 7A13-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 7A12,A15).

The vertebrate orthologue of *Sp-Pdx1* is essential for pancreas development, β -cell differentiation, and maintaining mature β -cell function (Kaneto et al., 2008). 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 genes, 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

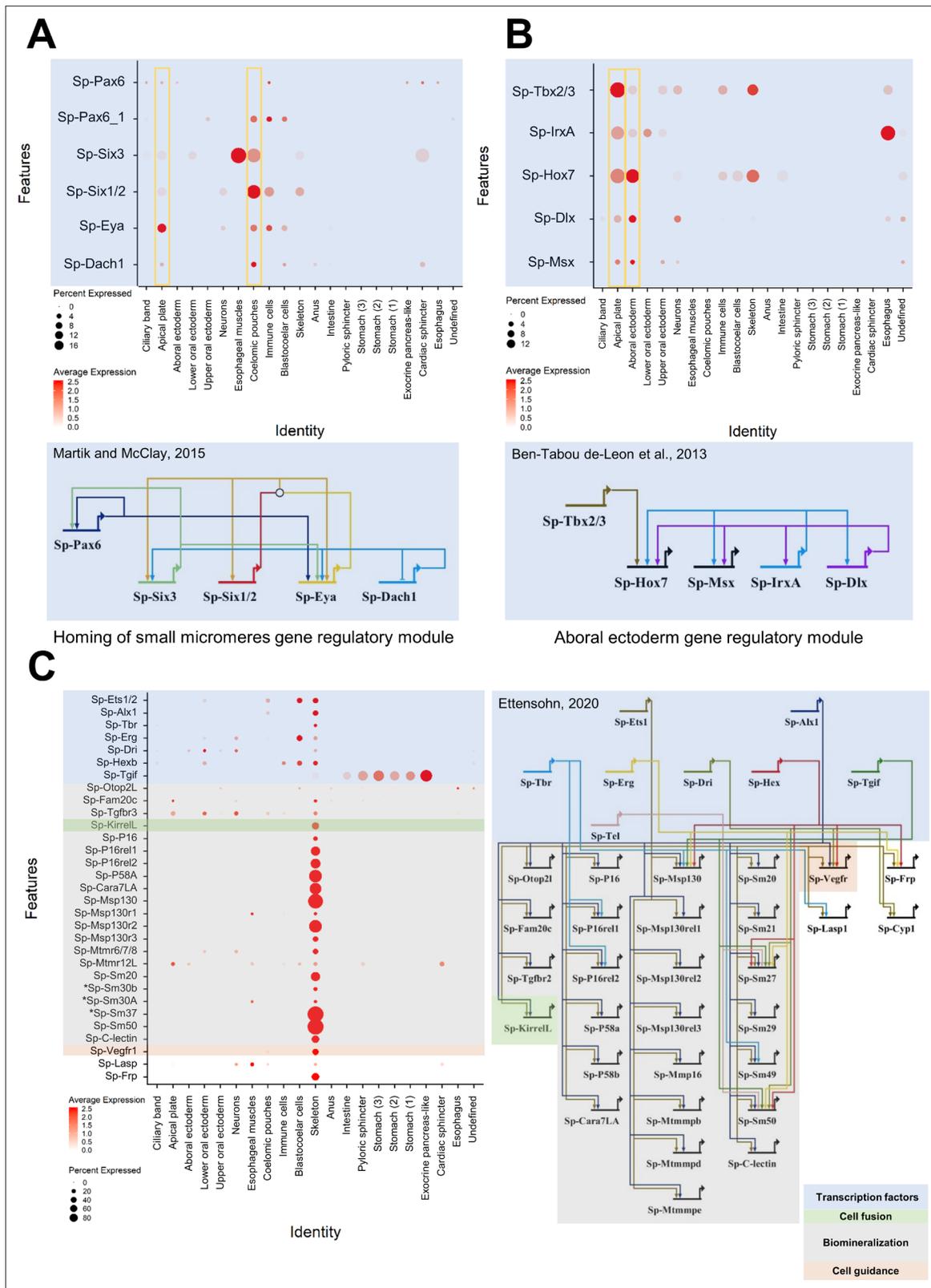


Figure 5. Validation of preexisting GRNs and putative novel function of specific gene regulatory modules. **(A)** Dotplot showing the mRNA localization of genes involved in the homing of small micromeres to the coelomic pouch and novel apical plate domain. **(B)** Dotplot of aboral ectoderm regulatory 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.

Figure 5 continued on next page

Figure 5 continued

The online version of this article includes the following figure supplement(s) for figure 5:

Figure supplement 1. Reconstruction of the molecular signature of the gut at a single-cell resolution.

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-Ngffap*, *Sp-Th*, and *Sp-Chat* (**Figure 7C**).

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 as 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 from insects (Davie et al., 2018; Severo et al., 2018; Cho et al., 2020), to cnidarians (Sebé-Pedrós et al., 2018; Siebert et al., 2019), planarians (Swapna et al., 2018; Zeng 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 cell types of the early 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 1A**), that consist of derivatives from all three germ layers. Similar ectodermal, mesodermal and endodermal representatives have also been reported at a single cell resolution in previous studies, carried out at either earlier developmental stages or different sea urchin species, highlighting the power of this technique to capture fine and dynamic developmental processes during sea urchin development (Foster et al., 2020; Massri et al., 2021). The cells comprising the 21 cell type families are expressing a total of 15,578 WHL genes (transcriptome models) and 12,924 genes (SPU gene models). 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 due to their calcite and syncytial nature are the last to dissociate and hence cells can get trapped within the debris and thus be underestimated in our datasets. Nonetheless, the cell type families and 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 15,579, 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 type families within the early pluteus larva. In particular, our study reveals unprecedented transcriptional diversity among cells of the pluteus larva that relates to feeding and digestion, as seen by the large representation of such clusters in our dataset. This includes two distinct oral ectoderm cell clusters, as well as distinct cell expression programs for the esophagus, cardiac and pyloric sphincters, exocrine pancreas-like cells, three distinct stomach cell clusters, intestine and anus. We also identified the mesodermally derived esophageal 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.

Cell type family tree reconstruction reveals two well supported groups (**Figure 3—figure supplement 1**). One group represents the endoderm. It comprises a core of closely related cell type families (stomach cells 1,2,3, pyloric and cardiac sphincter, and exocrine pancreas-like cells) and the more distant subgroup of intestinal and anal cells. The coherence of this group is also apparent from the transcription factor heatmap (**Figure 4—figure supplement 2**), which reveals numerous factors shared between subsets or the entire endodermal group such as *Klf15* that is known to control lipid and

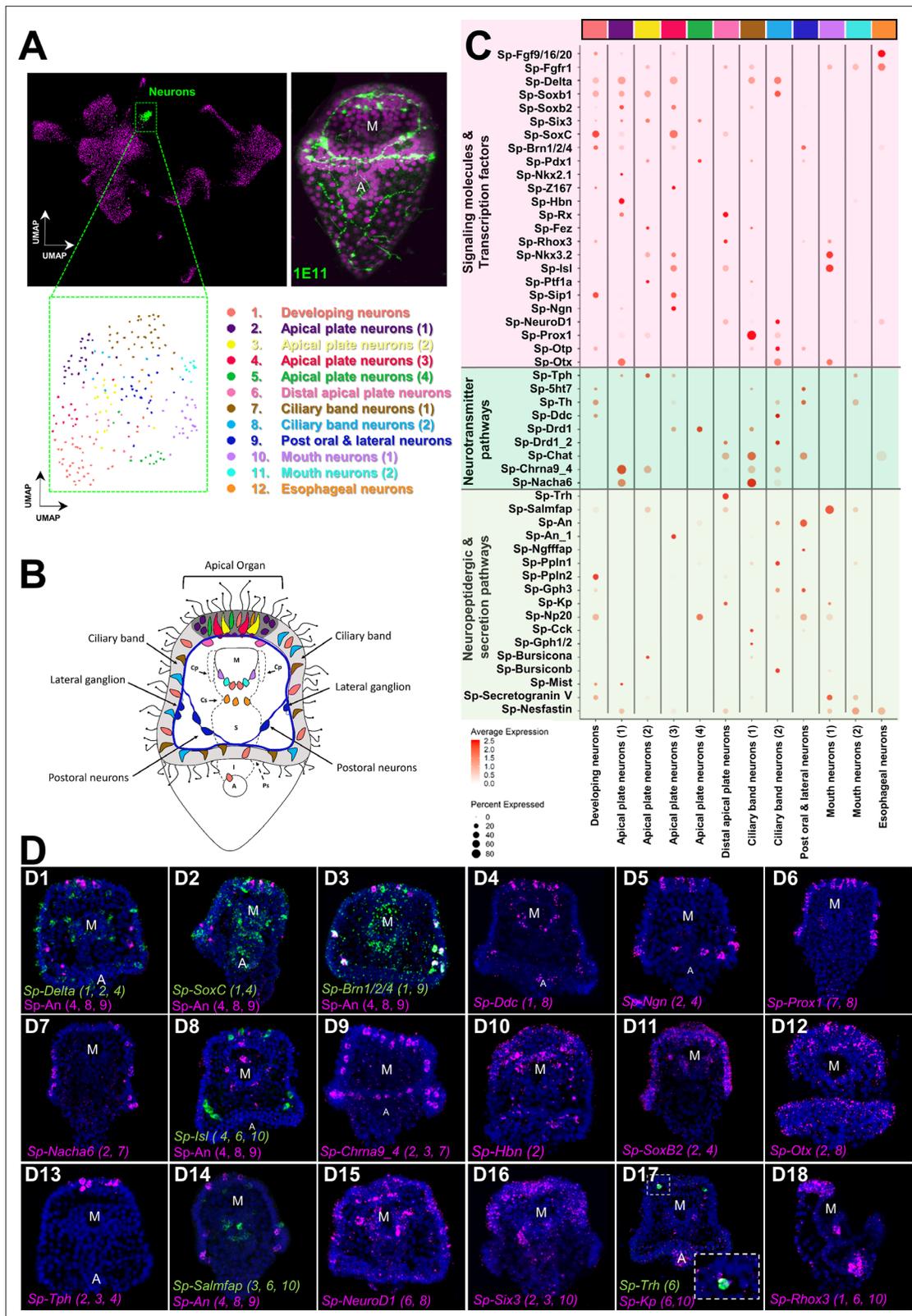


Figure 6. Neuronal complexity of the three dpf *S. purpuratus* larva. **(A)** (From left to right and top to bottom) UMAP highlighting the neurons cluster (green), immunohistochemical detection for the paneuronal sea urchin marker 1E11 (green), UMAP showing the 12 distinct neuronal subclusters. **(B)** Schematic representation of the three 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.*

Figure 6 continued on next page

Figure 6 continued

purpuratus three 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-Chrma9_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 labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. A, anus; M, mouth.

The online version of this article includes the following figure supplement(s) for figure 6:

Figure supplement 1. Subclustering of the skeletal cell type family.

Figure supplement 2. Subclustering of the immune cells cluster.

Figure supplement 3. Co-localization of *Sp-Fgf9/16/20* and *Sp-FgfR1* in the cardiac sphincter region.

glucose metabolism in the liver (**Oishi and Manabe, 2018**). The other group represents ectoderm. It contains apical plate, ciliary band, the aboral and upper oral ectoderm cells, and neurons. Again, this group manifests in shared transcription factor expression including *Klf2/4*, *Tbx2/3*, and *SoxB1* and *SoxB2*. Unexpectedly, this group firmly comprises some cell type families that are usually not considered ectodermal, such as the immune cells and the esophageal cells (**Figure 3—figure supplement 1**). Outside these two groups, the remaining cell type families do not assemble into a well-supported group. Instead, they exhibit molecular affinities to either the endodermal or ectodermal clades.

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 (mouth neurons and esophageal neurons, see **Figure 6A**). 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 populations, which we were able to trace back on the larva. By doing so, we identified (i) one cluster consisting of neurons that apparently are undergoing differentiation as judged by the combinatorial expression of *Delta*, *Soxc*, *Brn1/2/4* and terminal differentiation gene markers, (ii) four associated with the apical domain, one of which is expressing the same combination of TFs mentioned above, suggesting it might also be undergoing differentiation, (iii) one matching the distal apical plate neurons, (iv) two corresponding to ciliary band neurons, (v) two located in the rim of the mouth, (vi) one associated with esophageal structures, and (vii) 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 cell cluster (undefined) that could not be traced back to a specific domain. Based on the numbers and distributions of the features and transcript numbers that are comparable to the ones of real cell type families, such as the esophageal cell type family, we speculate that this

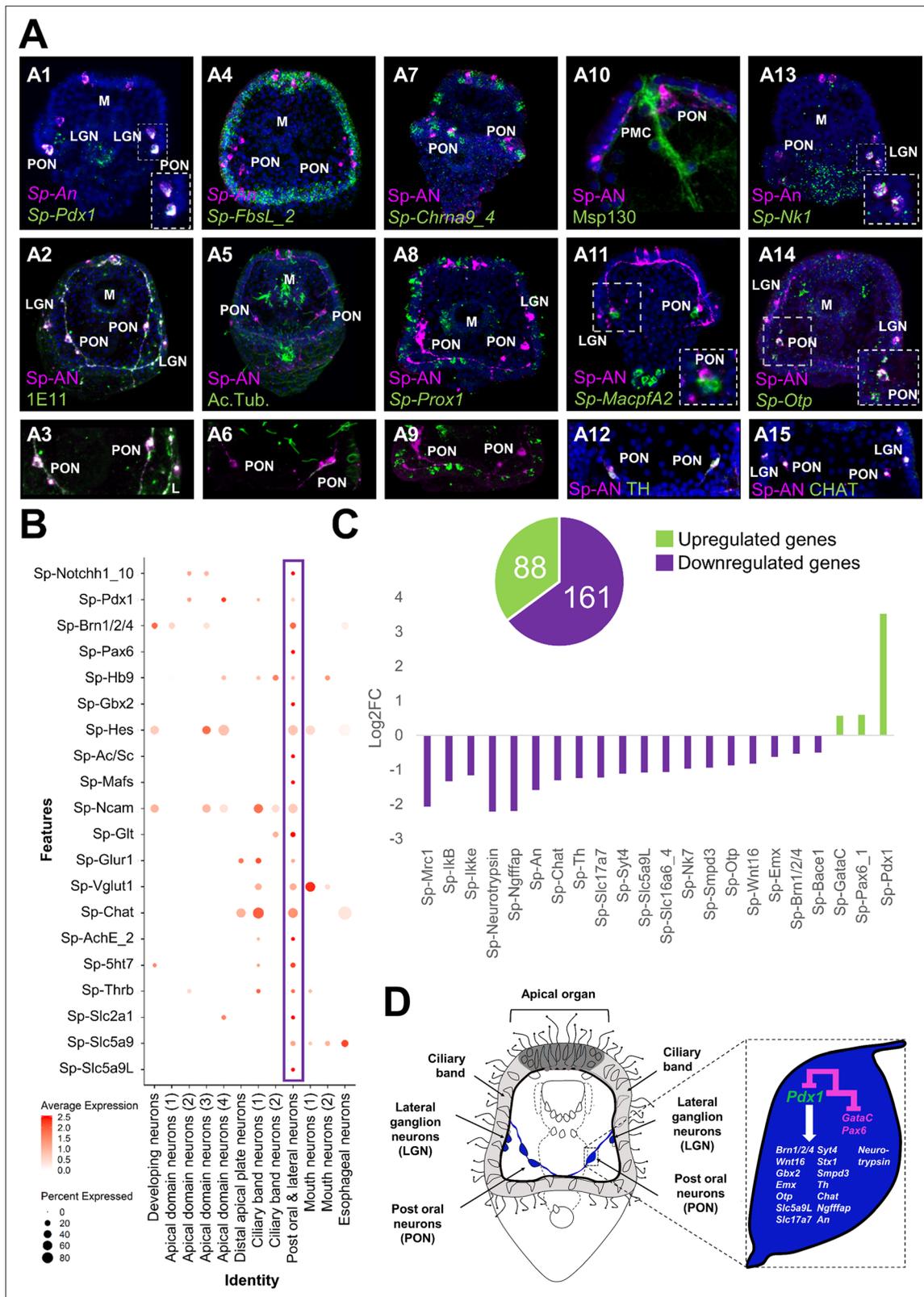


Figure 7. ScRNA-seq reveals a Pdx-1-dependent neuroendocrine cell type. (A) Molecular characterization of a *Sp-Pdx1*/*Sp-Brn1/2/4* double positive neuronal population. (A1) Double FISH of *S. purpuratus* 3 dpf larvae with specific antisense probes for *Sp-Pdx1* and *Sp-An*. (A2) Double immunohistochemical 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* three dpf larvae with specific antisense probes for *Sp-FbsL_2* and *Sp-An*. Double immunohistochemical detection of

Figure 7 continued on next page

Figure 7 continued

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* three dpf larvae with a specific antisense probe for *Sp-Chrma9_4* paired with immunodetection of Sp-An. (A8) FISH of *S. purpuratus* three 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* three 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* three dpf larvae with specific antisense probes for *Sp-Nk1* and *Sp-An*. (A14) FISH of *S. purpuratus* three 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 labeled 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) Schematic representation of the Sp-Pdx1/Sp-Brn1/2/4 neurons genetic wiring as revealed by the combination of scRNA-seq and differential RNA-seq analysis after *Sp-Pdx1* knockdown, highlighting an important role of *Sp-Pdx1* in the differentiation of this cell type.

The online version of this article includes the following figure supplement(s) for figure 7:

Figure supplement 1. Pancreatic genes expression significance testing in the ‘post oral and lateral neurons’ subcluster.

represents a real cell type family and not empty droplets. (Figure 1—figure supplement 1E). Overall, this cluster exhibited greatest transcriptional similarity to cell types derived from the ectoderm (Figure 1—figure supplement 1F), suggesting it may be of ectodermal origin. Consistent with this, a recent study 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 be 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. 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). 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 (Hes, *Ascl1*, *Pax6*, *Hb9*, *Mafs*, *Gbx2*), and signaling molecules (Notch) have been characterized as essential for this process (Ishibashi et al., 1995; de la Pompa et al., 1997; Li et al., 1999; Iso et al., 2003; Mizusawa et al., 2004; Zaret and Grompe, 2008; Arkhipova et al., 2012; Hart et al., 2013; Tritschler et al., 2017; Mitchell et al., 2017; Buckle et al., 2018; de la Pompa et al., 1997; Iso et al., 2003).

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.

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 ortholog of 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., 1987). Moreover, these neurons also express genes encoding members of the glutamate signaling pathway

(*Sp-Glt*: glutamate synthase; *Sp-Vglut1*: glutamate transporter; *Sp-Glut1*: glutamate receptor), which in mammals are involved in glucose-responsive insulin secretion (Gonoï *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 (Rodríguez-Díaz *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 that 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 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.

Evolutionary considerations

Taken together, our data show that the sea urchin *Sp-Pdx1/Sp-Brn1/2/4* neurons uniquely express a combination of key genes that are necessary for the endocrine pancreas differentiation and function in the vertebrates (Figure 7, Figure 7—figure supplement 1). These shared molecular features of *Sp-Pdx1/Sp-Brn1/2/4* neurons and pancreatic cells suggest these represent features of a cell type present in the deuterostome ancestor. What was the nature of these cells and what happened in the evolution of the descendant lineages so that the molecular characteristics of these cells are found in extant cell types as divergent as ectodermal neurons and pancreatic cells?

One possibility is that the ancestral deuterostomes already possessed a *Sp-Pdx1/Sp-Brn1/2/4* neuronal cell type similar to the one found in today's sea urchin, and that the regulatory program of this cell type was then co-opted into the endodermal lineage in the vertebrates. The co-option and incorporation of gene regulatory elements to different developmental or morphogenetic context has been postulated before and has the potential 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). In particular, it has been 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) based on the many physiological, morphological and molecular features endocrine pancreatic cells share with neurons (Alpert *et al.*, 1988; Eberhard *et al.*, 2013). However, given the long list of genes shared between neurons and pancreatic cells, this would have necessarily represented a massive co-option event profoundly altering the identity and function of the receiving cells. Notably, a flexibility toward the co-option of a neuroendocrine phenotype has also been reported in the vertebrate lineage in the case of the adrenal medulla derived PC12 cell line, that upon stimulation with NGF, adopts a neurosecretory fate (Fujita *et al.*, 1989).

A second possibility is that both ectodermal neurons and endodermal pancreatic cells are direct evolutionary descendants of the same precursor cell, representing sister cell types (Arendt, 2008). This explanation would avoid co-option; instead, it postulates a direct evolutionary link between cell types derived from different germ layers - ectoderm versus endoderm. Surprising at first, a similar link between digestive and neuronal-type expression profiles has been reported repeatedly. For example, it appears to hold true for choanocytes in the early-branching sponges (Musser *et al.*, 2021), indicating that the family of sponge choanocytes might be related to the hypothesized digestive-neuronal precursors. In the sea anemone *Nematostella vectensis*, pharyngeal cells give rise to both digestive cells and neurons (Steinmetz *et al.*, 2017), and neuronal cell types and secretory gland cells exhibit related molecular profiles (Sebé-Pedrós *et al.*, 2018). In fact, based on these findings a close

evolutionary link between endoderm and ectoderm has been postulated (*Steinmetz et al., 2017*). Finally, motor neurons in the vertebrate ventral neural tube and pancreatic islet cells share a specific combination of transcription factors including *pax6*, *nkx6*, and *hb9* (reviewed in *Arendt, 2021*). Overall, this second possibility is also in line with the hypothesis that at least subsets of neuron types in animal nervous systems may have originated from a mucociliary sole that initially harbored a neuro-secretory network of cells with digestive and communicative functions (*Arendt et al., 2015; Arendt, 2021*). Via division of labor, these precursor cell types may have then given rise to both digestive as well as neuronal cell types in different animal phyla.

Materials and methods

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 deionized water.

Larvae dissociation

Dissociation of the three dpf *Strongylocentrotus purpuratus* plutei into single cells was performed according to an 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 Ca²⁺ Mg²⁺-free artificial sea water. Larvae were spun down at 500 g for 5 min and resuspended in dissociation buffer containing 1 M glycine and 0.02 M EDTA in Ca²⁺ Mg²⁺-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 Ca²⁺ Mg²⁺-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 (10 x Genomics). Throughout this procedure samples were kept at 4 °C.

Single-cell RNA sequencing and data analysis

Single cell RNA sequencing was performed using the 10 x Genomics single-cell capturing system. Specimens from four independent biological replicates, ranging from 6000 to 20,000 cells, were loaded on the 10 X 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 88 M 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 Consortium 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. Of the 29,130 cells estimated by Cell Ranger, 19,699 cells 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 family of interest and performing similar analysis as described above. All resulting tables containing the genes transcribed within different cell type families 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).

Cell type tree reconstruction was performed in R (version 4.0.3) using the neighbor-joining tree function in package 'ape' (version 5.4–1). Two final trees were constructed, one that derived pairwise expression distances between clusters using all expressed genes, and a second in the analysis was restricted to only expressed transcription factors. A gene was defined as 'expressed' when it had an average expression of greater than three transcripts per million in at least one of the 21 major clusters. For each tree, pairwise distances were calculated as $1 - \text{spearman correlation}$ between each of the 20 major clusters, excluding the undefined cells, using the `spearman.dist` function in the `bioDist` package (version 1.60.0). To assess the support of each bifurcation in the tree, we performed 10,000 bootstrap replicates using the `boot.phylo` function in 'ape'. Although not reported in this manuscript, we also tested the effect of using different distance metrics and normalization methods on tree topology. Specifically, we reconstructed trees using distances derived from Pearson correlations of either log or square root normalized average expression values, and Euclidean distances of scaled log-normalized expression values. In general, similar topologies were found in all analyses, recovering clades of endodermal and ectodermal derivatives. Mesodermal derivatives often, but not always, grouped together, and the position of the neuron cluster moved in and out of the clade of ectodermal derivatives. This latter result may possibly arise because of conflicting signals resulting from cell type heterogeneity (i.e. different neuron types) within the cluster.

In order to assess the significance of identifying the homologs of the vertebrate genes involved in the development of the pancreas in the 'post oral and lateral neurons' sub-population of neuronal cells, the normalized counts from all the transcripts present in the neuronal subcluster were extracted in the form of a matrix using `GetAssayData` command from Seurat v3.0.2. The counts within this matrix were then randomized with 100,000 iterations using the `randomizeMatrix` function from the `picante` v1.8.2 package (Kembel et al., 2010) with the independent swap null model. A new Seurat object was created, which is the same as the neuronal subclustering object, however the actual transcript normalized counts were replaced with the randomized ones. Then the average expression of the 20 pancreatic genes was extracted from both from actual neuronal subclustering Seurat object and the one with randomized counts using `AverageExpression` function to visualize the expression of these pancreatic genes in the 'post oral and lateral neurons' subcluster in actual data and randomized data. This process was looped 500 times for randomized data to then calculate the mean number of pancreatic genes expressed in the 'post oral and lateral neurons' subcluster to obtain an estimate of the expected number of genes expressed, rounded to the closest integer, if their expression was found in this subcluster by random. The number of genes expressed in the actual data and the randomized data was then used to generate proportions of the number of expressed and not expressed pancreatic genes in this neuronal subcluster. Then, we have performed a chi-squared test within R (Hope, 1968) to show that the proportion from the actual data proportion is significantly different to the randomized data proportion with p-value less than 0.002.

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 from 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 file 2**. Specimens were imaged using a Zeiss LSM 700 confocal microscope.

Immunohistochemistry (IHC)

Immunohistochemical staining or IHC paired with FISH was performed as described in *Perillo et al., 2021*. Briefly three 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 1 hr. Primary antibodies were added in the appropriate dilution and incubated for 1 hr 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 (1:20), 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 (1:100). Specimens were washed multiple times with PBST and incubated for 1 hr 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.

EdU labeling 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 647 (Thermo Fisher Scientific). Larvae were treated with EdU at a final concentration of 10 μ M in FSW and let to grow for 2 hr. 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 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.

Gene regulatory network draft

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

Acknowledgements

The authors 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 thank Davide Caramiello for taking care of the adult sea urchins, Dr. Giovanna Benvenuto (SZN) for microscopy assistance and the Arnone lab members Maria Cocurullo and Inés Fournon Berodia for their help in gene cloning and preparation of several 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.

Additional information

Funding

Funder	Grant reference number	Author
H2020 Marie Skłodowska-Curie Actions	766053	Periklis Paganos Detlev Arendt Maria Ina Arnone
H2020 European Research Council	788921	Jacob M Musser Detlev Arendt

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Periklis Paganos, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review and editing; Danila Voronov, Data curation, Formal analysis, Methodology, Writing - review and editing; Jacob M Musser, Data curation, Formal analysis, Funding acquisition, Methodology, Writing - review and editing; Detlev Arendt, Funding acquisition, Writing - review and editing; Maria Ina Arnone, Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing - review and editing

Author ORCIDs

Periklis Paganos  <http://orcid.org/0000-0001-9525-4625>
 Danila Voronov  <http://orcid.org/0000-0002-2972-6484>
 Detlev Arendt  <http://orcid.org/0000-0001-7833-050X>
 Maria Ina Arnone  <http://orcid.org/0000-0002-9012-7624>

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.70416.sa1>

Author response <https://doi.org/10.7554/eLife.70416.sa2>

Additional files

Supplementary files

- Transparent reporting form
- Supplementary file 1. Differentially expressed genes per cell type family and putative target genes in the PDX1 positive neurons.
- Supplementary file 2. Primers used to generate specific antisense RNA probes.

Data availability

Sequencing data (mapped reads) have been deposited in Dryad under the unique identifier <https://doi.org/10.5061/dryad.n5tb2rbvz>.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Paganos P, Voronov D, Musser J, Arendt D, Arnone MI	2021	Data from: Sp 3 dpf scRNA-Seq	http://dx.doi.org/10.5061/dryad.n5tb2rbvz	Dryad Digital Repository, 10.5061/dryad.n5tb2rbvz

Continued on next page

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Lowe EK	2016	Data from: paraHox_analysis	https://github.com/elijahlowe/paraHox_analysis	Github, paraHox_analysis
Lowe EK	2015	S. purpurarus reads/ Sp-Lox 48 and 72	https://osf.io/cbsxr/files/	Open Science Framework, cbsxr

References

- Adomako-Ankomah A**, Etensohn CA. 2014. Growth factors and early mesoderm morphogenesis: insights from the sea urchin embryo. *Genesis* **52**:158–172. DOI: <https://doi.org/10.1002/dvg.22746>, PMID: 24515750
- Aiello V**, Moreno-Asso A, Servitja JM, Martín M. 2014. Thyroid hormones promote endocrine differentiation at expenses of exocrine tissue. *Experimental Cell Research* **322**:236–248. DOI: <https://doi.org/10.1016/j.yexcr.2014.01.030>, PMID: 24503054
- Alpert S**, Hanahan D, Teitelman G. 1988. Hybrid insulin genes reveal a developmental lineage for pancreatic endocrine cells and imply a relationship with neurons. *Cell* **53**:295–308. DOI: [https://doi.org/10.1016/0092-8674\(88\)90391-1](https://doi.org/10.1016/0092-8674(88)90391-1), PMID: 3282675
- Amisten S**, Neville M, Hawkes R, Persaud SJ, Karpe F, Salehi A. 2015. An atlas of G-protein coupled receptor expression and function in human subcutaneous adipose tissue. *Pharmacology & Therapeutics* **146**:61–93. DOI: <https://doi.org/10.1016/j.pharmthera.2014.09.007>, PMID: 25242198
- Andrikou C**, Iovene E, Rizzo F, Oliveri P, Arnone MI. 2013. Myogenesis in the sea urchin embryo: the molecular fingerprint of the myoblast precursors. *EvoDevo* **4**:33. DOI: <https://doi.org/10.1186/2041-9139-4-33>, PMID: 24295205
- Andrikou C**, Pai CY, Su YH, Arnone MI. 2015. Logics and properties of a genetic regulatory program that drives embryonic muscle development in an echinoderm. *eLife* **4**:e07343. DOI: <https://doi.org/10.7554/eLife.07343>, PMID: 26218224
- Angerer RC**, Davidson EH. 1984. Molecular indices of cell lineage specification in sea urchin embryos. *Science* **226**:1153–1160. DOI: <https://doi.org/10.1126/science.6594757>, PMID: 6594757
- Arishchenko E**, Arnone MI, D'Aniello S. 2018. SoxB2 in sea urchin development: implications in neurogenesis, ciliogenesis and skeletal patterning. *EvoDevo* **9**:5. DOI: <https://doi.org/10.1186/s13227-018-0094-1>, PMID: 29479411
- Annunziata R**, Arnone MI. 2014. A dynamic regulatory network explains ParaHox gene control of gut patterning in the sea urchin. *Development* **141**:2462–2472. DOI: <https://doi.org/10.1242/dev.105775>, PMID: 24850857
- Annunziata R**, Perillo M, Andrikou C, Cole AG, Martinez P, Arnone MI. 2014. Pattern and process during sea urchin gut morphogenesis: the regulatory landscape. *Genesis* **52**:251–268. DOI: <https://doi.org/10.1002/dvg.22738>, PMID: 24376127
- Annunziata R**, Andrikou C, Perillo M, Cuomo C, Arnone MI. 2019. Development and evolution of gut structures: from molecules to function. *Cell and Tissue Research* **377**:445–458. DOI: <https://doi.org/10.1007/s00441-019-03093-9>, PMID: 31446445
- Arendt D**. 2008. The evolution of cell types in animals: emerging principles from molecular studies. *Nature Reviews. Genetics* **9**:868–882. DOI: <https://doi.org/10.1038/nrg2416>, PMID: 18927580
- Arendt D**, Benito-Gutierrez E, Brunet T, Marlow H. 2015. Gastric pouches and the mucociliary sole: setting the stage for nervous system evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **370**:20150286. DOI: <https://doi.org/10.1098/rstb.2015.0286>, PMID: 26554050
- Arendt D**. 2021. Elementary nervous systems. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **376**:20200347. DOI: <https://doi.org/10.1098/rstb.2020.0347>, PMID: 33550948
- 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. *Developmental Biology* **365**:290–302. DOI: <https://doi.org/10.1016/j.ydbio.2012.03.001>, PMID: 22426004
- Arnone MI**, Byrne M, Martínez P. 2015. Echinodermata. Wanninger A (Ed). *Evolutionary Developmental Biology of Invertebrates* 6. Vienna: Springer. p. 1–214. DOI: <https://doi.org/10.1007/978-3-7091-1856-6>, PMID: 25674161
- Arnone MI**, Andrikou C, Annunziata R, Andrikou C. 2016. Echinoderm systems for gene regulatory studies in evolution and development. *Current Opinion in Genetics & Development* **39**:129–137. DOI: <https://doi.org/10.1016/j.gde.2016.05.027>, PMID: 27389072
- Arntfield ME**, van der Kooy D. 2011. β -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: News and Reviews in Molecular, Cellular and Developmental Biology* **33**:582–587. DOI: <https://doi.org/10.1002/bies.201100015>, PMID: 21681773

- Ben-Tabou de-Leon S**, Su YH, Lin KT, Li E, Davidson EH. 2013. Gene regulatory control in the sea urchin aboral ectoderm: spatial initiation, signaling inputs, and cell fate lockdown. *Developmental Biology* **374**:245–254. DOI: <https://doi.org/10.1016/j.ydbio.2012.11.013>, PMID: 23211652
- Berger C**, Zdzienicka D. 2020. Glucose transporters in pancreatic islets. *Pflugers Archiv* **472**:1249–1272. DOI: <https://doi.org/10.1007/s00424-020-02383-4>, PMID: 32394191
- Bisgrove B**, Burke R. 1987. Development of the nervous system of the pluteus larva of *Strongylocentrotus droebachiensis*. *Cell and Tissue Research* **248**:335–343. DOI: <https://doi.org/10.1007/BF00218200>
- Buckle A**, Nozawa RS, Kleinjan DA, Gilbert N. 2018. Functional characteristics of novel pancreatic Pax6 regulatory elements. *Human Molecular Genetics* **27**:3434–3448. DOI: <https://doi.org/10.1093/hmg/ddy255>, PMID: 30007277
- Burke RD**, Angerer LM, Elphick MR, Humphrey GW, Yaguchi S, Kiyama T, Liang S, Mu X, Agca C, Klein WH, Brandhorst BP, Rowe M, Wilson K, Churcher AM, Taylor JS, Chen N, Murray G, Wang D, Mellott D, Olinski R, et al. 2006a. A genomic view of the sea urchin nervous system. *Developmental Biology* **300**:434–460. DOI: <https://doi.org/10.1016/j.ydbio.2006.08.007>, PMID: 16965768
- Burke RD**, Osborne L, Wang D, Murabe N, Yaguchi S, Nakajima Y. 2006b. Neuron-specific expression of a synaptotagmin gene in the sea urchin *Strongylocentrotus purpuratus*. *The Journal of Comparative Neurology* **496**:244–251. DOI: <https://doi.org/10.1002/cne.20939>, PMID: 16538680
- Burke RD**, Moller DJ, Krupke OA, Taylor VJ. 2014. Sea urchin neural development and the metazoan paradigm of neurogenesis. *Genesis* **52**:208–221. DOI: <https://doi.org/10.1002/dvg.22750>, PMID: 25368883
- Butler A**, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology* **36**:411–420. DOI: <https://doi.org/10.1038/nbt.4096>, PMID: 29608179
- Calestani C**, Rogers DJ. 2010. Cis-regulatory analysis of the sea urchin pigment cell gene polyketide synthase. *Developmental Biology* **340**:249–255. DOI: <https://doi.org/10.1016/j.ydbio.2010.01.026>, PMID: 20122918
- Cameron RA**, Hough-Evans BR, Britten RJ, Davidson EH. 1987. Lineage and fate of each blastomere of the eight-cell sea urchin embryo. *Genes & Development* **1**:75–85. DOI: <https://doi.org/10.1101/gad.1.1.75>, PMID: 2448185
- Cameron RA**, Davidson EH. 1991. Cell type specification during sea urchin development. *Trends in Genetics* **7**:212–218. DOI: [https://doi.org/10.1016/0168-9525\(91\)90367-y](https://doi.org/10.1016/0168-9525(91)90367-y), PMID: 1887502
- Cao C**, Lemaire LA, Wang W, Yoon PH, Choi YA, Parsons LR, Matese JC, Wang W, Levine M, Chen K. 2019. Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* **571**:349–354. DOI: <https://doi.org/10.1038/s41586-019-1385-y>, PMID: 31292549
- Cary GA**, McCauley BS, Zueva O, Pattinato J, Longabaugh W, Hinman VF. 2020. Systematic comparison of sea urchin and sea star developmental gene regulatory networks explains how novelty is incorporated in early development. *Nature Communications* **11**:6235. DOI: <https://doi.org/10.1038/s41467-020-20023-4>, PMID: 33277483
- Chestnut B**, Casie Chetty S, Koenig AL, Sumanas S. 2020. Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle. *Nature Communications* **11**:2796. DOI: <https://doi.org/10.1038/s41467-020-16515-y>, PMID: 32493965
- Cho B**, Yoon SH, Lee D, Koranteng F, Tattikota SG, Cha N, Shin M, Do H, Hu Y, Oh SY, Lee D, Vipin Menon A, Moon SJ, Perrimon N, Nam JW, Shim J. 2020. Single-cell transcriptome maps of myeloid blood cell lineages in *Drosophila*. *Nature Communications* **11**:4483. DOI: <https://doi.org/10.1038/s41467-020-18135-y>, PMID: 32900993
- Cole AG**, Arnone MI. 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 Expression Patterns* **9**:324–328. DOI: <https://doi.org/10.1016/j.gexp.2009.02.005>, PMID: 19250980
- Cole AG**, Rizzo F, Martinez P, Fernandez-Serra M, Arnone MI. 2009. Two ParaHox genes, SpLox and SpCdx, interact to partition the posterior endoderm in the formation of a functional gut. *Development* **136**:541–549. DOI: <https://doi.org/10.1242/dev.029959>, PMID: 19144720
- Davidson EH**, Cameron RA, Ransick A. 1998. Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. *Development* **125**:3269–3290. DOI: <https://doi.org/10.1242/dev.125.17.3269>, PMID: 9693132
- Davidson EH**, McClay DR, Hood L. 2003. Regulatory gene networks and the properties of the developmental process. *PNAS* **100**:1475–1480. DOI: <https://doi.org/10.1073/pnas.0437746100>, PMID: 12578984
- Davidson EH**, Erwin DH. 2006. Gene regulatory networks and the evolution of animal body plans. *Science* **311**:796–800. DOI: <https://doi.org/10.1126/science.1113832>, PMID: 16469913
- Davie K**, Janssens J, Koldere D, De Waegeneer M, Pech U, Kreft L, Aibar S, Makhzami S, Christiaens V, Bravo González-Blas C, Poovathingal S, Hulselmans G, Spanier KI, Moerman T, Vanspauwen B, Geurs S, Voet T, Lammertyn J, Thienpont B, Liu S, et al. 2018. A Single-Cell Transcriptome Atlas of the Aging *Drosophila* Brain. *Cell* **174**:982–998. DOI: <https://doi.org/10.1016/j.cell.2018.05.057>, PMID: 29909982
- de la Pompa JL**, Wakeham A, Correia KM, Samper E, Brown S, Aguilera RJ, Nakano T, Honjo T, Mak TW, Rossant J, Conlon RA. 1997. Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development* **124**:1139–1148. DOI: <https://doi.org/10.1242/dev.124.6.1139>, PMID: 9102301
- 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**:2293–2302. DOI: <https://doi.org/10.1242/dev.005108>, PMID: 17507391

- Eberhard CE**, Fu A, Reeks C, Sreaton RA. 2013. CRT2 is required for β -cell function and proliferation. *Endocrinology* **154**:2308–2317. DOI: <https://doi.org/10.1210/en.2012-2088>, PMID: 23677932
- Esaulova E**, Cantoni C, Shchukina I, Zaitsev K, Bucelli RC, Wu GF, Artyomov MN, Cross AH, Edelson BT. 2020. Single-cell RNA-seq analysis of human CSF microglia and myeloid cells in neuroinflammation. *Neurology(R) Neuroimmunology & Neuroinflammation* **7**:e732. DOI: <https://doi.org/10.1212/NXI.0000000000000732>, PMID: 32371549
- Ettensohn CA**. 2020. The gene regulatory control of sea urchin gastrulation. *Mechanisms of Development* **162**:103599. DOI: <https://doi.org/10.1016/j.mod.2020.103599>, PMID: 32119908
- Finn RD**, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. 2014. Pfam: the protein families database. *Nucleic Acids Research* **42**:D222–D230. DOI: <https://doi.org/10.1093/nar/gkt1223>, PMID: 24288371
- Foster S**, Oulhen N, Wessel G. 2020. A single cell RNA sequencing resource for early sea urchin development. *Development* **147**:dev191528. DOI: <https://doi.org/10.1242/dev.191528>, PMID: 32816969
- Fu X**, He F, Li Y, Shahveranov A, Hutchins AP. 2017. Genomic and molecular control of cell type and cell type conversions. *Cell Regeneration* **6**:1–7. DOI: <https://doi.org/10.1016/j.cr.2017.09.001>, PMID: 29348912
- Fujita K**, Lazarovici P, Guroff G. 1989. Regulation of the differentiation of PC12 pheochromocytoma cells. *Environmental Health Perspectives* **80**:127–142. DOI: <https://doi.org/10.1289/ehp.8980127>, PMID: 2647474
- Garner S**, Zysk I, Byrne G, Kramer M, Moller D, Taylor V, Burke RD. 2016. Neurogenesis in sea urchin embryos and the diversity of deuterostome neurogenic mechanisms. *Development* **143**:286–297. DOI: <https://doi.org/10.1242/dev.124503>, PMID: 26511925
- 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. *The Journal of Biological Chemistry* **269**:16989–16992. DOI: [https://doi.org/10.1016/S0021-9258\(17\)32507-3](https://doi.org/10.1016/S0021-9258(17)32507-3), PMID: 8006003
- Harkey MA**, Whiteley HR, Whiteley AH. 1992. Differential expression of the msp130 gene among skeletal lineage cells in the sea urchin embryo: a three dimensional in situ hybridization analysis. *Mechanisms of Development* **37**:173–184. DOI: [https://doi.org/10.1016/0925-4773\(92\)90079-y](https://doi.org/10.1016/0925-4773(92)90079-y), PMID: 1498042
- Harlow P**, Nemer M. 1987. Coordinate and selective beta-tubulin gene expression associated with cilium formation in sea urchin embryos. *Genes & Development* **1**:1293–1304. DOI: <https://doi.org/10.1101/gad.1.10.1293>, PMID: 3428593
- Hart AW**, Mella S, Mendrychowski J, van Heyningen V, Kleinjan DA. 2013. The developmental regulator Pax6 is essential for maintenance of islet cell function in the adult mouse pancreas. *PLOS ONE* **8**:e54173. DOI: <https://doi.org/10.1371/journal.pone.0054173>, PMID: 23326594
- Ho EC**, Buckley KM, Schrankel CS, Schuh NW, Hibino T, Solek CM, Bae K, Wang G, Rast JP. 2017. Perturbation of gut bacteria induces a coordinated cellular immune response in the purple sea urchin larva. *Immunology and Cell Biology* **95**:647. DOI: <https://doi.org/10.1038/icb.2017.40>, PMID: 28785056
- Hope ACA**. 1968. A Simplified Monte Carlo Significance Test Procedure. *Journal of the Royal Statistical Society* **30**:582–598. DOI: <https://doi.org/10.1111/j.2517-6161.1968.tb00759.x>
- Howard-Ashby M**, Materna SC, Brown CT, Chen L, Cameron RA, Davidson EH. 2006. Identification and characterization of homeobox transcription factor genes in *Strongylocentrotus purpuratus*, and their expression in embryonic development. *Developmental Biology* **300**:74–89. DOI: <https://doi.org/10.1016/j.ydbio.2006.08.039>, PMID: 17055477
- Hu Y**, Schmitt-Engel C, Schwirz J, Stroehlein N, Richter T, Majumdar U, Bucher G. 2018. A morphological novelty evolved by co-option of a reduced gene regulatory network and gene recruitment in a beetle. *Proceedings. Biological Sciences* **285**:20181373. DOI: <https://doi.org/10.1098/rspb.2018.1373>, PMID: 30135167
- Hui H**, Perfetti R. 2002. Pancreas duodenum homeobox-1 regulates pancreas development during embryogenesis and islet cell function in adulthood. *European Journal of Endocrinology* **146**:129–141. DOI: <https://doi.org/10.1530/eje.0.1460129>, PMID: 11834421
- Hussain MA**, Miller CP, Habener JF. 2002. Brn-4 transcription factor expression targeted to the early developing mouse pancreas induces ectopic glucagon gene expression in insulin-producing beta cells. *The Journal of Biological Chemistry* **277**:16028–16032. DOI: <https://doi.org/10.1074/jbc.M107124200>, PMID: 11834725
- Ishibashi M**, Ang SL, Shiota K, Nakanishi S, Kageyama R, Guillemot F. 1995. Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. *Genes & Development* **9**:3136–3148. DOI: <https://doi.org/10.1101/gad.9.24.3136>, PMID: 8543157
- Iso T**, Kedes L, Hamamori Y. 2003. HES and HERP families: multiple effectors of the Notch signaling pathway. *Journal of Cellular Physiology* **194**:237–255. DOI: <https://doi.org/10.1002/jcp.10208>, PMID: 12548545
- Iturriza FC**, Thibault J. 1993. Immunohistochemical investigation of tyrosine-hydroxylase in the islets of Langerhans of adult mice, rats and guinea pigs. *Neuroendocrinology* **57**:476–480. DOI: <https://doi.org/10.1159/000126394>, PMID: 8100618
- Juliano CE**, Voronina E, Stack C, Aldrich M, Cameron AR, Wessel GM. 2006. Germ line determinants are not localized early in sea urchin development, but do accumulate in the small micromere lineage. *Developmental Biology* **300**:406–415. DOI: <https://doi.org/10.1016/j.ydbio.2006.07.035>, PMID: 16970939
- Juliano CE**, Yajima M, Wessel GM. 2010. Nanos functions to maintain the fate of the small micromere lineage in the sea urchin embryo. *Developmental Biology* **337**:220–232. DOI: <https://doi.org/10.1016/j.ydbio.2009.10.030>, PMID: 19878662
- Juliano C**, Swartz SZ, Wessel G. 2014. Isolating specific embryonic cells of the sea urchin by FACS. *Methods in Molecular Biology* **1128**:187–196. DOI: https://doi.org/10.1007/978-1-62703-974-1_12, PMID: 24567215

- Jung M**, Wells D, Rusch J, Ahmad S, Marchini J, Myers SR, Conrad DF. 2019. Unified single-cell analysis of testis gene regulation and pathology in five mouse strains. *eLife* **8**:e43966. DOI: <https://doi.org/10.7554/eLife.43966>, PMID: 31237565
- Kaneto H**, Miyatsuka T, Kawamori D, Yamamoto K, Kato K, Shiraiwa T, Katakami N, Yamasaki Y, Matsuhisa M, Matsuoka TA. 2008. PDX-1 and MafA play a crucial role in pancreatic beta-cell differentiation and maintenance of mature beta-cell function. *Endocrine Journal* **55**:235–252. DOI: <https://doi.org/10.1507/endocrj.k07e-041>, PMID: 17938503
- Katow H**, Suyemitsu T, Ooka S, Yaguchi J, Jin-Nai T, Kuwahara I, Katow T, Yaguchi S, Abe H. 2010. Development of a dopaminergic system in sea urchin embryos and larvae. *The Journal of Experimental Biology* **213**:2808–2819. DOI: <https://doi.org/10.1242/jeb.042150>, PMID: 20675551
- Kemmel SW**, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**:1463–1464. DOI: <https://doi.org/10.1093/bioinformatics/btq166>, PMID: 20395285
- Khor JM**, Guerrero-Santoro J, Etensohn CA. 2019. Genome-wide identification of binding sites and gene targets of Alx1, a pivotal regulator of echinoderm skeletogenesis. *Development* **146**:dev180653. DOI: <https://doi.org/10.1242/dev.180653>, PMID: 31331943
- Klein AM**, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW. 2015. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**:1187–1201. DOI: <https://doi.org/10.1016/j.cell.2015.04.044>, PMID: 26000487
- Krouk G**, Lingeman J, Colon AM, Coruzzi G, Shasha D. 2013. Gene regulatory networks in plants: learning causality from time and perturbation. *Genome Biology* **14**:123. DOI: <https://doi.org/10.1186/gb-2013-14-6-123>, PMID: 23805876
- Kudtarkar P**, Cameron RA. 2017. Echinobase: an expanding resource for echinoderm genomic information. *Database* **2017**:bax074. DOI: <https://doi.org/10.1093/database/bax074>, PMID: 29220460
- Langley OK**, Aletsee MC, Gratzl M. 1987. Endocrine cells share expression of N-CAM with neurones. *FEBS Letters* **220**:108–112. DOI: [https://doi.org/10.1016/0014-5793\(87\)80885-2](https://doi.org/10.1016/0014-5793(87)80885-2), PMID: 2440721
- Li H**, Arber S, Jessell TM, Edlund H. 1999. Selective agenesis of the dorsal pancreas in mice lacking homeobox gene Hlx9. *Nature Genetics* **23**:67–70. DOI: <https://doi.org/10.1038/12669>, PMID: 10471501
- Longabaugh WJR**. 2012. BioTapestry: a tool to visualize the dynamic properties of gene regulatory networks. *Methods in Molecular Biology* **786**:359–394. DOI: https://doi.org/10.1007/978-1-61779-292-2_21, PMID: 21938637
- Luo YJ**, Su YH. 2012. Opposing nodal and BMP signals regulate left-right asymmetry in the sea urchin larva. *PLOS Biology* **10**:e1001402. DOI: <https://doi.org/10.1371/journal.pbio.1001402>, PMID: 23055827
- Lyons DC**, Kaltenbach SL, McClay DR. 2012. Morphogenesis in sea urchin embryos: linking cellular events to gene regulatory network states. *Wiley Interdisciplinary Reviews. Developmental Biology* **1**:231–252. DOI: <https://doi.org/10.1002/wdev.18>, PMID: 23801438
- Maechler P**, Wollheim CB. 1999. Mitochondrial glutamate acts as a messenger in glucose-induced insulin exocytosis. *Nature* **402**:685–689. DOI: <https://doi.org/10.1038/45280>, PMID: 10604477
- Martik ML**, McClay DR. 2015. Deployment of a retinal determination gene network drives directed cell migration in the sea urchin embryo. *eLife* **4**:e08827. DOI: <https://doi.org/10.7554/eLife.08827>, PMID: 26402456
- Massri AJ**, Greenstreet L, Afanassiev A, Berrio A, Wray GA, Schiebinger G, McClay DR. 2021. Developmental single-cell transcriptomics in the *Lytechinus variegatus* sea urchin embryo. *Development* **148**:dev198614. DOI: <https://doi.org/10.1242/dev.198614>, PMID: 34463740
- Materna SC**, Howard-Ashby M, Gray RF, Davidson EH. 2006. The C2H2 zinc finger genes of *Strongylocentrotus purpuratus* and their expression in embryonic development. *Developmental Biology* **300**:108–120. DOI: <https://doi.org/10.1016/j.ydbio.2006.08.032>, PMID: 16997293
- Materna SC**, Ransick A, Li E, Davidson EH. 2013. Diversification of oral and aboral mesodermal regulatory states in pregastrular sea urchin embryos. *Developmental Biology* **375**:92–104. DOI: <https://doi.org/10.1016/j.ydbio.2012.11.033>, PMID: 23261933
- McClay DR**. 1986. Embryo dissociation, cell isolation, and cell reassociation. *Methods in Cell Biology* **27**:309–323. DOI: [https://doi.org/10.1016/s0091-679x\(08\)60355-6](https://doi.org/10.1016/s0091-679x(08)60355-6), PMID: 2422529
- McClay DR**. 2004. Methods for embryo dissociation and analysis of cell adhesion. *Methods in Cell Biology* **74**:311–329. DOI: [https://doi.org/10.1016/s0091-679x\(04\)74014-5](https://doi.org/10.1016/s0091-679x(04)74014-5), PMID: 15575613
- McClay DR**. 2011. Evolutionary crossroads in developmental biology: sea urchins. *Development* **138**:2639–2648. DOI: <https://doi.org/10.1242/dev.048967>, PMID: 21652646
- McClay DR**, Miranda E, Feinberg SL. 2018. Neurogenesis in the sea urchin embryo is initiated uniquely in three domains. *Development* **145**:dev167742. DOI: <https://doi.org/10.1242/dev.167742>, PMID: 30413529
- McClay DR**, Warner J, Martik M, Miranda E, Slota L. 2020. Gastrulation in the sea urchin. *Current Topics in Developmental Biology* **136**:195–218. DOI: <https://doi.org/10.1016/bs.ctdb.2019.08.004>, PMID: 31959288
- McQueen E**, Rebeiz M. 2020. On the specificity of gene regulatory networks: How does network co-option affect subsequent evolution? *Current Topics in Developmental Biology* **139**:375–405. DOI: <https://doi.org/10.1016/bs.ctdb.2020.03.002>, PMID: 32450967
- Mitchell RK**, Nguyen-Tu MS, Chabosseau P, Callingham RM, Pullen TJ, Cheung R, Leclerc I, Hodson DJ, Rutter GA. 2017. The transcription factor Pax6 is required for pancreatic β cell identity, glucose-regulated ATP synthesis, and Ca²⁺ dynamics in adult mice. *The Journal of Biological Chemistry* **292**:8892–8906. DOI: <https://doi.org/10.1074/jbc.M117.784629>, PMID: 28377501

- 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. DOI: <https://doi.org/10.1016/j.gene.2004.01.016>, PMID: 15094191
- 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 : News and Reviews in Molecular, Cellular and Developmental Biology* **34**:181–186. DOI: <https://doi.org/10.1002/bies.201100160>, PMID: 22223407
- Morgulis M**, Gildor T, Roopin M, Sher N, Malik A, Lalar 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. *PNAS* **116**:12353–12362. DOI: <https://doi.org/10.1073/pnas.1902126116>, PMID: 31152134
- Musser JM**, Schippers KJ, Nickel M, Mizzon G, Kohn AB, Pape C, Ronchi P, Papadopoulos N, Tarashansky AJ, Hammel JU, Wolf F, Liang C, Hernández-Plaza A, Cantalapiedra CP, Achim K, Schieber NL, Pan L, Ruperti F, Francis WR, Vargas S, et al. 2021. Profiling Cellular Diversity in Sponges Informs Animal Cell Type and Nervous System Evolution. *bioRxiv*. DOI: <https://doi.org/10.1101/758276>
- Navale AM**, Paranjape AN. 2016. Glucose transporters: physiological and pathological roles. *Biophysical Reviews* **8**:5–9. DOI: <https://doi.org/10.1007/s12551-015-0186-2>, PMID: 28510148
- Nestorowa S**, Hamey FK, Pijuan Sala B, Diamanti E, Shepherd M, Laurenti E, Wilson NK, Kent DG, Göttgens B. 2016. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood* **128**:e20–e31. DOI: <https://doi.org/10.1182/blood-2016-05-716480>, PMID: 27365425
- Oishi Y**, Manabe I. 2018. Krüppel-Like Factors in Metabolic Homeostasis and Cardiometabolic Disease. *Frontiers in Cardiovascular Medicine* **5**:69. DOI: <https://doi.org/10.3389/fcvm.2018.00069>, PMID: 29942807
- Okazaki K**. 1965. Skeleton formation of sea urchin larvae. V. Continuous observation of the process of matrix formation. *Experimental Cell Research* **40**:585–596. DOI: [https://doi.org/10.1016/0014-4827\(65\)90236-3](https://doi.org/10.1016/0014-4827(65)90236-3), PMID: 5866069
- Oliveri P**, Tu Q, Davidson EH. 2008. Global regulatory logic for specification of an embryonic cell lineage. *PNAS* **105**:5955–5962. DOI: <https://doi.org/10.1073/pnas.0711220105>, PMID: 18413610
- Pehrson JR**, Cohen LH. 1986. The fate of the small micromeres in sea urchin development. *Developmental Biology* **113**:522–526. DOI: [https://doi.org/10.1016/0012-1606\(86\)90188-0](https://doi.org/10.1016/0012-1606(86)90188-0), PMID: 3512335
- Perillo M**, Arnone MI. 2014. Characterization of insulin-like peptides (ILPs) in the sea urchin *Strongylocentrotus purpuratus*: insights on the evolution of the insulin family. *General and Comparative Endocrinology* **205**:68–79. DOI: <https://doi.org/10.1016/j.ygcen.2014.06.014>, PMID: 24971803
- Perillo M**, Wang YJ, Leach SD, Arnone MI. 2016. A pancreatic exocrine-like cell regulatory circuit operating in the upper stomach of the sea urchin *Strongylocentrotus purpuratus* larva. *BMC Evolutionary Biology* **16**:117. DOI: <https://doi.org/10.1186/s12862-016-0686-0>, PMID: 27230062
- Perillo M**, Paganos P, Mattiello T, Cocurullo M, Oliveri P, Arnone MI. 2018. New Neuronal Subtypes With a “Pre-Pancreatic” Signature in the Sea Urchin *Strongylocentrotus purpuratus*. *Frontiers in Endocrinology* **9**:650. DOI: <https://doi.org/10.3389/fendo.2018.00650>, PMID: 30450080
- 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**:e60388. DOI: <https://doi.org/10.7554/eLife.60388>, PMID: 32812865
- Perillo M**, Paganos P, Spurrell M, Arnone MI, Wessel GM. 2021. Methodology for Whole Mount and Fluorescent RNA In Situ Hybridization in Echinoderms: Single, Double, and Beyond. *Methods in Molecular Biology* **2219**:195–216. DOI: https://doi.org/10.1007/978-1-0716-0974-3_12, PMID: 33074542
- Preger-Ben Noon E**, Frankel N. 2015. Evolving Genital Structures: A Deep Look at Network Co-option. *Developmental Cell* **34**:485–486. DOI: <https://doi.org/10.1016/j.devcel.2015.08.022>, PMID: 26374760
- 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. *Biochemical and Biophysical Research Communications* **526**:135–140. DOI: <https://doi.org/10.1016/j.bbrc.2020.03.044>, PMID: 32199615
- Rafiq K**, Cheers MS, Ettensohn CA. 2012. The genomic regulatory control of skeletal morphogenesis in the sea urchin. *Development* **139**:579–590. DOI: <https://doi.org/10.1242/dev.073049>, PMID: 22190640
- Rast JP**, Smith LC, Loza-Coll M, Hibino T, Litman GW. 2006. Genomic insights into the immune system of the sea urchin. *Science* **314**:952–956. DOI: <https://doi.org/10.1126/science.1134301>, PMID: 17095692
- Rizzo F**, Fernandez-Serra M, Squarzon P, Archimandritis A, Arnone MI. 2006. Identification and developmental expression of the ets gene family in the sea urchin (*Strongylocentrotus purpuratus*). *Developmental Biology* **300**:35–48. DOI: <https://doi.org/10.1016/j.ydbio.2006.08.012>, PMID: 16997294
- Rodriguez-Diaz R**, Dando R, Jacques-Silva MC, Fachado A, Molina J, Abdulreda MH, Ricordi C, Roper SD, Berggren PO, Caicedo A. 2011. Alpha cells secrete acetylcholine as a non-neuronal paracrine signal priming beta cell function in humans. *Nature Medicine* **17**:888–892. DOI: <https://doi.org/10.1038/nm.2371>, PMID: 21685896
- Sea Urchin Genome Sequencing Consortium**, Sodergren E, Weinstock GM, Davidson EH, Cameron RA, Gibbs RA, Angerer RC, Angerer LM, Arnone MI, Burgess DR, Burke RD, Coffman JA, Dean M, Elphick MR, Ettensohn CA, Foltz KR, Hamdoun A, Hynes RO, Klein WH, Marzluff W, et al. 2006. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* **314**:941–952. DOI: <https://doi.org/10.1126/science.1133609>, PMID: 17095691
- Sebé-Pedrós A**, Saudemont B, Chomsky E, Plessier F, Mailhé MP, Renno J, Loe-Mie Y, Lifshitz A, Mukamel Z, Schmutz S, Novault S, Steinmetz PRH, 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. DOI: <https://doi.org/10.1016/j.cell.2018.05.019>, PMID: 29856957
- Severo MS**, Landry JJM, Lindquist RL, Goosmann C, Brinkmann V, Collier P, Hauser AE, Benes V, Henriksson J, Teichmann SA, Levashina EA. 2018. Unbiased classification of mosquito blood cells by single-cell genomics and high-content imaging. *PNAS* **115**:E7568–E7577. DOI: <https://doi.org/10.1073/pnas.1803062115>, PMID: 30038005
- Sharma S**, Wang W, Stolfi A. 2019. Single-cell transcriptome profiling of the *Ciona* larval brain. *Developmental Biology* **448**:226–236. DOI: <https://doi.org/10.1016/j.ydbio.2018.09.023>, PMID: 30392840
- Shashikant T**, Etensohn CA. 2019. Genome-wide analysis of chromatin accessibility using ATAC-seq. *Methods in Cell Biology* **151**:219–235. DOI: <https://doi.org/10.1016/bs.mcb.2018.11.002>, PMID: 30948010
- Shekhar K**, Menon V. 2019. Identification of Cell Types from Single-Cell Transcriptomic Data. *Methods in Molecular Biology* **1935**:45–77. DOI: https://doi.org/10.1007/978-1-4939-9057-3_4, PMID: 30758819
- Siebert S**, Farrell JA, Cazet JF, Abeykoon Y, Primack AS, Schnitzler CE, Juliano CE. 2019. Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. *Science* **365**:922–935. DOI: <https://doi.org/10.1126/science.aav9314>, PMID: 31346039
- Sladitschek HL**, Fiuza UM, Pavlinic D, Benes V, Hufnagel L, Neveu PA. 2020. MorphoSeq: Full Single-Cell Transcriptome Dynamics Up to Gastrulation in a Chordate. *Cell* **181**:922–935. DOI: <https://doi.org/10.1016/j.cell.2020.03.055>, PMID: 32315617
- Slota LA**, McClay DR. 2018. Identification of neural transcription factors required for the differentiation of three neuronal subtypes in the sea urchin embryo. *Developmental Biology* **435**:138–149. DOI: <https://doi.org/10.1016/j.ydbio.2017.12.015>, PMID: 29331498
- Slota LA**, Miranda E, Peskin B, McClay DR. 2020. Developmental origin of peripheral ciliary band neurons in the sea urchin embryo. *Developmental Biology* **459**:72–78. DOI: <https://doi.org/10.1016/j.ydbio.2019.12.011>, PMID: 31881199
- Smith MM**, Cruz Smith L, Cameron RA, Urry LA. 2008. The larval stages of the sea urchin, *Strongylocentrotus purpuratus*. *Journal of Morphology* **269**:713–733. DOI: <https://doi.org/10.1002/jmor.10618>, PMID: 18438782
- Soliman S**. 1983. Pharmacological control of ciliary activity in the young sea urchin larva: Effects of monoaminergic agents. *Comparative Biochemistry and Physiology. C, Comparative Pharmacology and Toxicology* **76**:181–191. DOI: [https://doi.org/10.1016/0742-8413\(83\)90061-0](https://doi.org/10.1016/0742-8413(83)90061-0), PMID: 6139247
- Steinmetz PRH**, Aman A, Kraus JEM, Technau U. 2017. Gut-like ectodermal tissue in a sea anemone challenges germ layer homology. *Nature Ecology & Evolution* **1**:1535–1542. DOI: <https://doi.org/10.1038/s41559-017-0285-5>, PMID: 29185520
- Strathmann MF**. 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. DOI: <https://doi.org/10.1086/416265>
- Stuart T**, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive Integration of Single-Cell Data. *Cell* **177**:1888–1902. DOI: <https://doi.org/10.1016/j.cell.2019.05.031>, PMID: 31178118
- Sun Z**, Etensohn CA. 2014. Signal-dependent regulation of the sea urchin skeletogenic gene regulatory network. *Gene Expression Patterns* **16**:93–103. DOI: <https://doi.org/10.1016/j.gep.2014.10.002>, PMID: 25460514
- Sun Z**, Etensohn CA. 2017. TGF- β sensu stricto signaling regulates skeletal morphogenesis in the sea urchin embryo. *Developmental Biology* **421**:149–160. DOI: <https://doi.org/10.1016/j.ydbio.2016.12.007>, PMID: 27955944
- Swapna LS**, Molinaro AM, Lindsay-Mosher N, Pearson BJ, Parkinson J. 2018. Comparative transcriptomic analyses and single-cell RNA sequencing of the freshwater planarian *Schmidtea mediterranea* identify major cell types and pathway conservation. *Genome Biology* **19**:124. DOI: <https://doi.org/10.1186/s13059-018-1498-x>, PMID: 30143032
- Tang F**, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, Lao K, Surani MA. 2009. mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods* **6**:377–382. DOI: <https://doi.org/10.1038/nmeth.1315>, PMID: 19349980
- Teitelman G**, Alpert S, Polak JM, 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**:1031–1039. DOI: <https://doi.org/10.1242/dev.118.4.1031>, PMID: 7903631
- Trapnell C**, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28**:511–515. DOI: <https://doi.org/10.1038/nbt.1621>, PMID: 20436464
- Tritschler S**, Theis FJ, Lickert H, Böttcher A. 2017. Systematic single-cell analysis provides new insights into heterogeneity and plasticity of the pancreas. *Molecular Metabolism* **6**:974–990. DOI: <https://doi.org/10.1016/j.molmet.2017.06.021>, PMID: 28951822
- Tu Q**, Brown CT, Davidson EH, Oliveri P. 2006. Sea urchin Forkhead gene family: phylogeny and embryonic expression. *Developmental Biology* **300**:49–62. DOI: <https://doi.org/10.1016/j.ydbio.2006.09.031>, PMID: 17081512

- Tu Q**, Cameron RA, Davidson EH. 2014. Quantitative developmental transcriptomes of the sea urchin *Strongylocentrotus purpuratus*. *Developmental Biology* **385**:160–167. DOI: <https://doi.org/10.1016/j.ydbio.2013.11.019>, PMID: 24291147
- Wagner DE**, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. 2018. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**:981–987. DOI: <https://doi.org/10.1126/science.aar4362>, PMID: 29700229
- Wei Z**, Yaguchi J, Yaguchi S, Angerer RC, Angerer LM. 2009. The sea urchin animal pole domain is a Six3-dependent neurogenic patterning center. *Development* **136**:1179–1189. DOI: <https://doi.org/10.1242/dev.032300>, PMID: 19270175
- Wei Z**, Angerer RC, Angerer LM. 2011. Direct development of neurons within foregut endoderm of sea urchin embryos. *PNAS* **108**:9143–9147. DOI: <https://doi.org/10.1073/pnas.1018513108>, PMID: 21576476
- 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**:1662–1669. DOI: <https://doi.org/10.1242/dev.075051>, PMID: 22438568
- Wei Z**, Angerer LM, Angerer RC. 2016. Neurogenic gene regulatory pathways in the sea urchin embryo. *Development* **143**:298–305. DOI: <https://doi.org/10.1242/dev.125989>, PMID: 26657764
- Wood NJ**, Mattiello T, Rowe ML, Ward L, Perillo M, Arnone MI, Elphick MR, Oliveri P. 2018. Neuropeptidergic Systems in Pluteus Larvae of the Sea Urchin *Strongylocentrotus purpuratus*: Neurochemical Complexity in a “Simple” Nervous System. *Frontiers in Endocrinology* **9**:628. DOI: <https://doi.org/10.3389/fendo.2018.00628>, PMID: 30410468
- Ximerakis M**, Lipnick SL, Innes BT, Simmons SK, Adiconis X, Dionne D, Mayweather BA, Nguyen L, Niziolek Z, Ozek C, Butty VL, Isserlin R, Buchanan SM, Levine SS, Regev A, Bader GD, Levin JZ, Rubin LL. 2019. Single-cell transcriptomic profiling of the aging mouse brain. *Nature Neuroscience* **22**:1696–1708. DOI: <https://doi.org/10.1038/s41593-019-0491-3>, PMID: 31551601
- Yaguchi J**, Takeda N, Inaba K, Yaguchi S. 2016. Cooperative Wnt-Nodal Signals Regulate the Patterning of Anterior Neuroectoderm. *PLOS Genetics* **12**:e1006001. DOI: <https://doi.org/10.1371/journal.pgen.1006001>, PMID: 27101101
- Yaguchi S**, Yaguchi J, Tanaka H. 2017. Troponin-I is present as an essential component of muscles in echinoderm larvae. *Scientific Reports* **7**:43563. DOI: <https://doi.org/10.1038/srep43563>, PMID: 28272398
- 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. *Journal of the American Society of Nephrology* **30**:2159–2176. DOI: <https://doi.org/10.1681/ASN.2019040335>, PMID: 31462402
- Yuh CH**, Li X, Davidson EH, Klein WH. 2001. Correct Expression of *spec2a* in the sea urchin embryo requires both *Otx* and other cis-regulatory elements. *Developmental Biology* **232**:424–438. DOI: <https://doi.org/10.1006/dbio.2001.0207>, PMID: 11401403
- Zaret KS**, Grompe M. 2008. Generation and regeneration of cells of the liver and pancreas. *Science* **322**:1490–1494. DOI: <https://doi.org/10.1126/science.1161431>, PMID: 19056973
- Zeng A**, Li H, Guo L, Gao X, McKinney S, Wang Y, Yu Z, Park J, Semerad C, Ross E, Cheng LC, Davies E, Lei K, Wang W, Perera A, Hall K, Peak A, Box A, Sánchez Alvarado A. 2018. Prospectively Isolated Tetraspanin+ Neoblasts Are Adult Pluripotent Stem Cells Underlying Planaria Regeneration. *Cell* **173**:1593–1608. DOI: <https://doi.org/10.1016/j.cell.2018.05.006>, PMID: 29906446
- 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 Discovery* **6**:22. DOI: <https://doi.org/10.1038/s41421-020-0157-z>, PMID: 32351704