Developmental Biology
Evolutionary Biology

Reviewed Preprint
v1 • June 5, 2024
Not revised

Single Cell RNA Sequencing Provides Clues for the Developmental Genetic Basis of Syngnathidae's Evolutionary Adaptations

Hope M. Healey, **Hayden B. Penn**, **Clayton M. Small**, **Susan Bassham**, **Vithika Goyal**, **Micah A. Woods**, **William A. Cresko**

Institute of Ecology and Evolution, University of Oregon • School of Computer and Data Science, University of Oregon • Knight Campus for Accelerating Scientific Impact, University of Oregon

Abstract

Seahorses, pipefishes, and seadragons are fishes from the family Syngnathidae that have evolved extraordinary traits including male pregnancy, elongated snouts, loss of teeth, and dermal bony armor. The developmental genetic and cellular changes that led to the evolution of these traits are largely unknown. Recent syngnathid genomes revealed suggestive gene content differences and provide the opportunity for detailed genetic analyses. We created a single cell RNA sequencing atlas of Gulf pipefish embryos to understand the developmental basis of four traits: derived head shape, toothlessness, dermal armor, and male pregnancy. We completed marker gene analyses, built genetic networks, and examined spatial expression of select genes. We identified osteochondrogenic mesenchymal cells in the elongating face that express regulatory genes *bmp4*, *sfrp1a*, and *prdm16*. We found no evidence for tooth primordia cells, and we observed re-deployment of osteoblast genetic networks in developing dermal armor.

Finally, we found that epidermal cells expressed nutrient processing and environmental sensing genes, potentially relevant for the brooding environment. The examined pipefish evolutionary innovations are composed of recognizable cell types, suggesting derived features originate from changes within existing gene networks. Future work addressing syngnathid gene networks across multiple stages and species is essential for understanding how their novelties evolved.

Impact Statement

The production of a single cell atlas for developing syngnathid fish permits study of their unique traits.

eLife assessment

This study provides a single-cell atlas for syngnathid fishes (seahorses, pipefishes, and seadragons), a **valuable** new resource to investigate the molecular basis of the many unique characters that define the pipefish embryo. The findings are generally supported by **solid** arguments, but whereas the single-cell RNA-sequencing analysis appears to be of good quality, the spatiotemporal expression data only **incompletely** support the authors' arguments. Additional computational analyses on cell identity and developmental trajectories would allow a deeper examination of the current data from these unconventional model organisms, to provide new insights into understanding the extraordinary adaptations of the Syngnathidae family. If appropriately improved, the work could be of broad interest for evolutionary developmental biology, particularly for fishes.

https://doi.org/10.7554/eLife.97764.1.sa3

Introduction

Seahorses, pipefishes, and seadragons are extraordinary fishes in the family Syngnathidae with diverse body plans, coloration, and elaborate structures for paternal brooding. The syngnathid clade comprises over 300 diverse species that vary in conservation status, distribution, ecology, and morphology (Leysen et al., 2011; Manning et al., 2019; Schneider, Woltering, et al., 2023; Stiller et al., 2022). Syngnathids have numerous highly altered traits, trait losses, and evolutionary novelties. They have elongated snouts bearing small, toothless jaws (Leysen et al., 2011) specialized for capture of small zooplankton (Van Wassenbergh et al., 2009). Additionally, syngnathids are distinctly protected by a bony dermal armor rather than scales (Jungerson, 1910). Other skeletal differences include a lack of ribs and pelvic fins, and an expansion in the number of vertebrae (Schneider, Woltering, et al., 2023). Finally, syngnathids exhibit male pregnancy, which has involved the evolution of specialized brooding tissues and structures (Whittington & Friesen, 2020). Paternal investment varies among lineages; for example, seadragons tether embryos externally to their tails while seahorses and some pipefishes have enclosed brood pouches proposed to support embryos through nutrient transfer and osmotic regulation (Carcupino, 2002; Melamed et al., 2005; Ripley & Foran, 2006).

Despite advances in understanding the ecology and evolution of syngnathid novelties, the developmental genetic basis for these traits is largely unknown. The recent production of high quality syngnathid genome assemblies (Qu et al., 2021; Ramesh et al., 2023; Small et al., 2022; Wolf et al., 2024) provides initial clues for the developmental genetic basis of some evolutionary changes. Studies have found that syngnathids lack several genes with deeply conserved roles in vertebrate development, including pharyngeal arch development (*fgf3*), tooth development (*fgf3*, *fgf4*, and *eve1*, and most scpp genes), fin development (*tbx4*), and immune function (MHC pathway components) (Lin et al., 2016; Qu et al., 2021; Small et al., 2016; Small et al., 2022; Zhang et al., 2020). Though these gene losses are highly suggestive of leading to unique changes, exploration of the actual developmental consequences of their losses is needed.

To fill this gap in knowledge, we used single cell RNA sequencing (scRNAseq) to investigate how these striking genomic changes have affected the developmental genetic and cellular basis of syngnathids' evolutionary derived traits. The Gulf pipefish (*Syngnathus scovelli*) is an attractive model for this study (**Figure 1A**). This species has a high-quality reference genome annotated by NCBI and is amenable to laboratory culture (Anderson & Jones, 2019; Ramesh et al., 2023).

Furthermore, species from the *Syngnathus* genus are used worldwide to address questions about syngnathid evolution in microbial, developmental, histological, transcriptomic, ecotoxicological, and genomic studies (Berglund et al., 1986; Carcupino, 2002; Fuiten & Cresko, 2021; Harada et al., 2022; Harlin-Cognato et al., 2006; Partridge et al., 2007; Ripley & Foran, 2006a; Rose et al., 2023; Roth et al., 2012; Small et al., 2016; Small et al., 2013). In this paper, we focus on a subset of unique traits, including their elongated head, toothlessness, dermal armor, and development of embryos inside the brood pouch. These traits represent the diversity of evolutionary changes observed in the syngnathid clade (highly altered, lost, and novel traits) and hypotheses from studies of model organisms suggest developmental pathways involved in their evolution (Lin et al., 2016b; Roth et al., 2020; Small et al., 2016; Small et al., 2022).

scRNAseq atlases are a powerful complement to genomic analyses (Shema et al., 2019; Ton et al., 2020). They can provide crucial insights into types of cells present, genes that distinguish cell types (marker genes), active gene networks, and a means to identify expression of genes of interest within predicted cell types (Farnsworth et al., 2020; Farrell et al., 2018; Williams et al., 2019). Specifically, scRNAseq captures RNA expression profiles from individual cells allowing cell types to be inferred *post-hoc*. scRNAseq has successfully been applied to syngnathid adult kidneys (Parker et al., 2022), but there are no published syngnathid developmental atlases.

Here we report the first developmental scRNAseq atlas for syngnathids from late embryogenesis staged Gulf pipefish. We delineate the overall structure of this atlas, which describes 38 cell clusters composed of 35,785 cells, and use these data to make inferences about the morphological evolution of several syngnathid innovations. In addition to inferring present cell types and their underlying genetic networks, we detail *in situ* spatial expression patterns of select marker and other candidate genes in pipefish embryos and juveniles. We found conserved signaling pathways expressed during craniofacial development but did not detect evidence of tooth primordia. The embryonic dermis and epidermis, respectively, expressed genes for the dermal armor development (bone development pathways) and genes potentially involved in interaction with the male brood pouch (e.g., nutrient acquisition genes). Overall, this atlas provides a deeper understanding of the development of Gulf pipefish and identifies gene candidates for understanding the development of syngnathid evolutionary innovations. In addition to these discoveries, this atlas provides a significant resource for researchers studying syngnathid evolution and development.

Results

Valuable scRNAseq atlas for studying syngnathid development

We produced the first developmental scRNAseq atlas for a syngnathid from two samples comprising similarly staged embryos from pregnant, wild-caught Gulf pipefish (*Syngnathus scovelli*) males. The samples represent a late organogenesis developmental stage (**Figure 1B,C**,D,E,F,G). These embryos had a primarily cartilaginous skeleton with minimal mineralization, including jaw cartilages that were at the onset of mineralization and ethmoid elongation. The embryos also possessed cartilaginous dorsal fin pterygiophores but had no signs of dermal armor mineralization.

The atlas included 35,785 cells (19,892 and 15,893 cells from each sample), which formed 38 cell clusters (**Figure 2A**, Supplementary File 2). We classified cells into 4 different broad tissue types – epithelial, connective, neural, and muscle – using Seurat identified marker genes and published model organism resources. We next used Seurat identified marker genes to pinpoint single marker genes that were most unique to each cluster (**Figure S1**, Supplementary File 3). We completed *in situ* hybridization using Gulf and bay pipefish embryos for cell clusters for which examining gene expression would help hone and validate cluster annotations (**Figure S2**, **S3**).

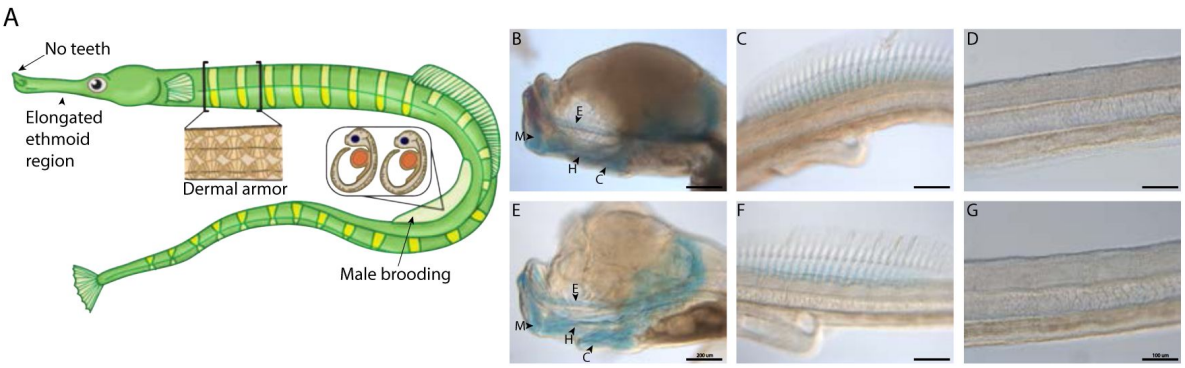

Figure 1

Gulf pipefish exemplify syngnathid derived traits.

Gulf pipefish have elongate snouts, have lost teeth on their oral and pharyngeal jaws, possess dermal armor, and have brood pouches in males (panel A). Cartilage (alcian) and bone (alizarin) stained clutch siblings of embryos from the two samples are shown in panels B-G. Embryos have cartilaginous craniofacial skeletons (panels B, E; E marks the Mes(ethmoid) cartilage, C indicates the Ceratohyal, H shows the Hyosymplectic cartilage, and M marks the Meckel's cartilage) with the onset of ossification in the jaw. They have cartilaginous fin radials in the dorsal fin (C and F). The embryos do not have signs of ossification in the trunk where the exoskeleton will form later (panels D,G). Panels B, C, E, and F, scale bar 200 μM; D, G scale bar 100 μM.

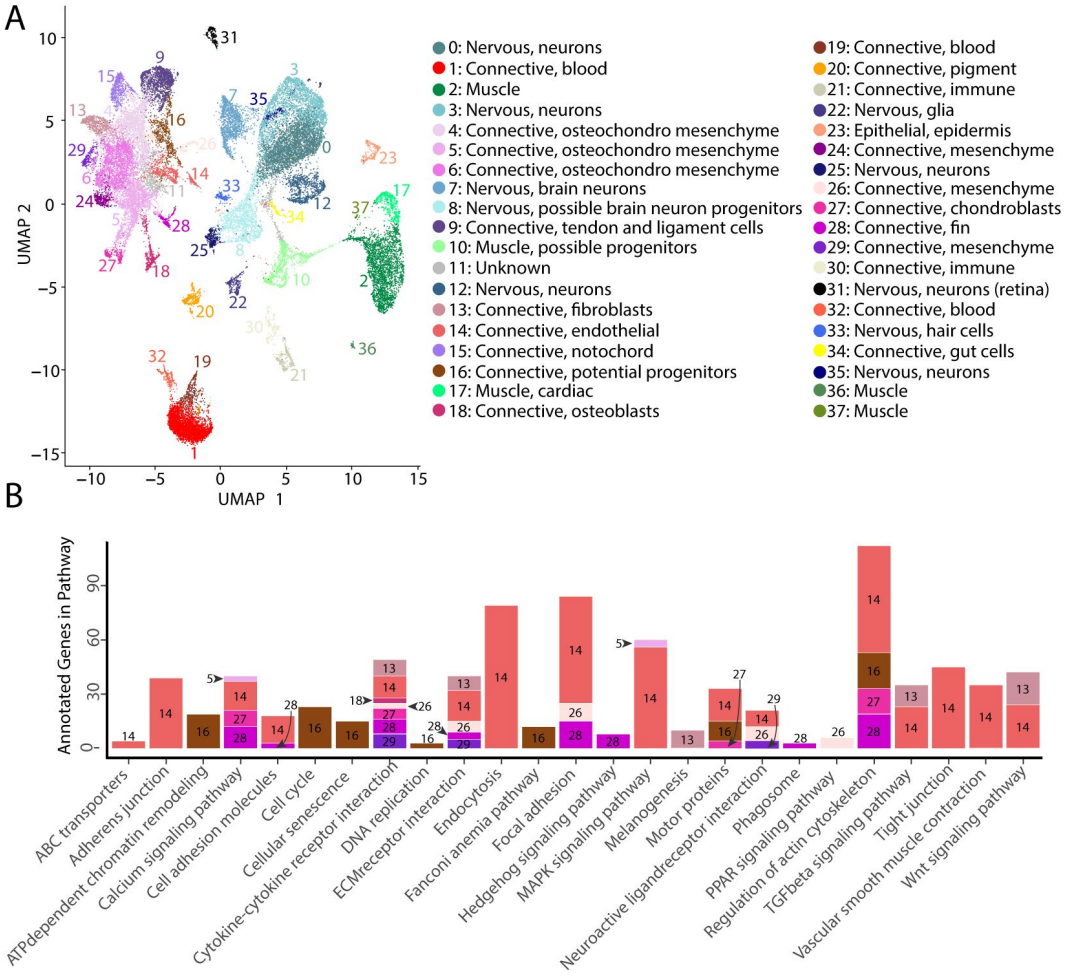

Figure 2

Gulf pipefish single cell atlas contains cells from the entire embryo and identifies genetic pathways active in different cell types.

The UMAP plot (panel A) shows all of the cell clusters and their identities reduced to the first two UMAP dimensions. The graph in panel B displays results of the KEGG pathway analysis in cell clusters identified as connective tissue (excluding blood, pigment, digestive, and immune cells). The number of Seurat identified marker genes for each cluster that was a part of each pathway is displayed on the y axis. Bars are colored and labeled by cell cluster.

In total, our atlas contained 13,027 connective tissue cells (excluding cells from blood, immune, and the digestive system) from 14 clusters, 10,112 nervous system cells from ten clusters, 4,363 muscle cells from five clusters, 4,133 blood cells from three clusters, 650 immune cells from two clusters, 432 pigment cells from one cluster, 370 epidermal cells from one cluster, and 137 gut cells from one cluster. Within the connective tissue cell types, we also identified cartilage (302 cells), developing bone (442 cells), fins (253 cells), and notochord (693 cells).

Discovery of cell cluster function and state using KEGG analysis

To affirm identities and discover potential properties of each cluster, we completed a KEGG pathway analysis for each cluster using Seurat's marker genes (**Figure S4**, **Figure 2B**). For eight of the clusters (1, 4, 6, 9, 11, 15, 19, and 24), we did not find any significantly enriched pathways, possibly due to similar gene expression profiles across cell types that reduced the number of identified markers. However, we found one or more significantly enriched pathways for the other 29 cell clusters. We observed enriched pathway terms that supported cluster annotations. For example, phototransduction in the retina cluster, melanogenesis in the pigment cluster, cardiac muscle contraction in muscle clusters, and neuroactive ligand receptor interaction in neuronal clusters.

The inferred KEGG pathways demonstrated some commonalities across the different tissue types, including in signaling pathways and cell states. Notably, our identified KEGG terms delineated progenitor and differentiated cell clusters. Based on their KEGG terms, we classified clusters 8, 10, and 16 as possible neural, muscle, and connective tissue progenitor cells, respectively. We also detected expression of *pax3a* and *pax3b*, muscle primordia markers, in cluster 10, supporting this annotation. These clusters had enriched KEGG terms associated with cell division ('cell cycle', 'DNA replication', 'nucleotide excision repair', and 'homologous recombination'), and lacked enrichment for KEGG pathways present with differentiated cell types of their lineage. Specifically, cluster 8 lacked the neural KEGG term 'neuroactive ligand receptor interaction', cluster 10 lacked muscle KEGG terms 'adrenergic signaling in cardiomyocytes', 'calcium signaling pathways', and 'cardiac muscle contraction', and cluster 16 lacked connective tissue term 'ECM receptor interaction'. Thus, it is likely that clusters 8, 10, and 16 represented undifferentiated cells within the major lineages of neural, muscle, and connective cells.

Commonalities of cell clusters, unique networks, and elusive cell types identified in network analysis

We built gene networks/modules from 3,000 variable genes using Weighted Gene Network Correlation Analysis (WGCNA; Langfelder & Horvath, 2008). This produced 43 gene modules in total (Supplementary File 5,6), assessed for each cluster-module pair for their strength of association (**Figure 3A**, Supplementary File 7,8) and every module's dependence on each cluster for their network connectivity (**Figure S5**, Supplementary File 9,10). Using the genes from each network, we completed a KEGG pathway analysis to identify whether gene modules indicated specific cellular pathways or states (**Figure 3B**). We initially explored whether these network-cluster associations could reveal commonalities between cell clusters or identify whether particular clusters contained multiple cell identities.

First, we asked whether gene modules that associate with three or more cell clusters signify commonalities between clusters that have similar cell types. We identified seven gene modules (6, 7, 21, 14, 17, 41, and 42) that each associate with three or more cell clusters. These gene modules do connect clusters of similar identities or cell states. For example, Modules 6 and 7 associate with connective tissue cells. Module 6, the larger of the two modules, contains numerous other KEGG pathways found in most connective cell clusters, such as 'cytokine-cytokine receptor interaction' (**Figure 3C**). Interestingly, Module 21, associated with clusters 8, 10, 16, and 25, contains genes from KEGG pathways related to the 'cell cycle' and 'cellular senescence', supporting the results

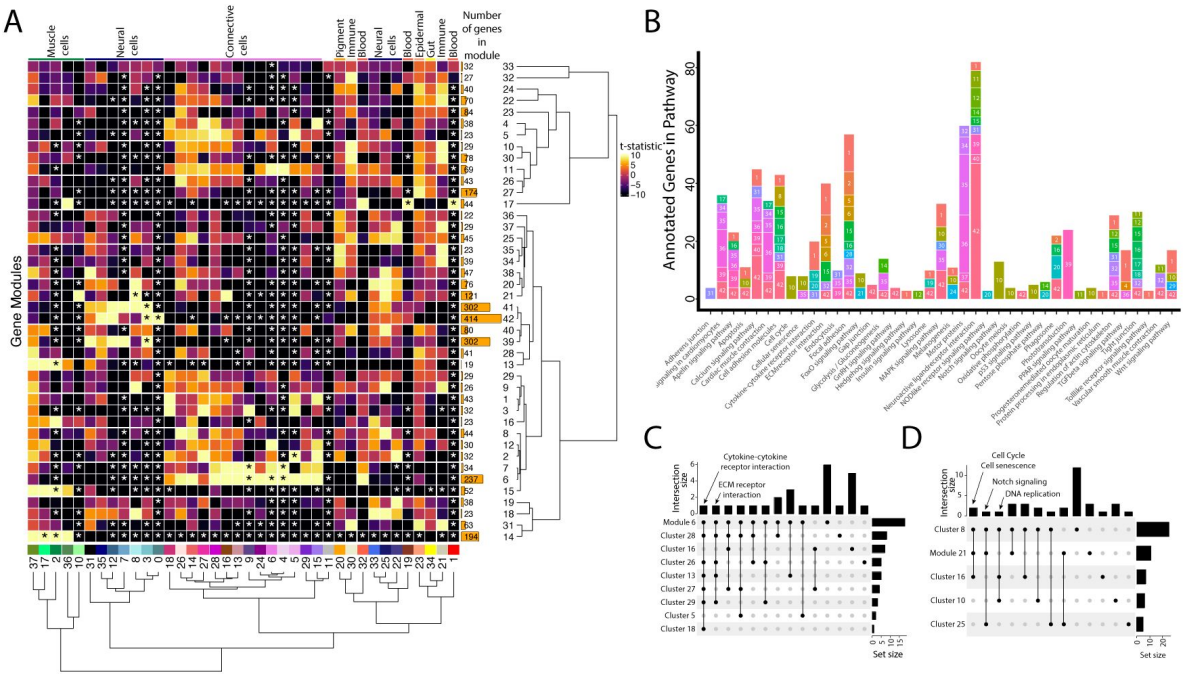

Figure 3

Weighted Gene Network Analysis (WGCNA) identifies gene modules that define and unite cell clusters.

A) The strength of association between the gene modules and cell clusters is shown in panel A with dendrogram clustering illustrating the distance between modules and cell clusters. Gene modules are represented by rows and cell clusters by columns. The modules and clusters are clustered using the Pearson distance method. The number of genes in each gene module are shown in the right hand bar plots. Cell clusters are colored based on their identity. The asterisks indicate the module-cluster relationships that have a p-value less than .05 from a two-sided permutation test after correction for multiple tests (FDR). The heatmap is colored by t-statistics in a range of -10 to 10, with highly positive values in yellow and highly negative values in black. B) The identified gene modules possess genes from KEGG pathways. The bars are labeled with the gene module and the size of each bar corresponds to the number of genes from the KEGG pathway in the module. Since WGCNA modules do not have p-values, only KEGG pathways with more than two genes included in the gene module are shown on the plot. C) Identified gene modules contain similar KEGG pathways as the cell clusters that correlated with them. These relationships are shown in Upset plots where each row is a cell cluster or gene module, each column represents KEGG pathways shared by the modules and clusters (shared condition is shown filled in black dots connected by lines), the interaction size is the number of pathways in common between the set of modules and clusters, and the set size is the number of pathways that are enriched in each cluster and module. Panel C1 highlights that 'cytokine-cytokine receptor interaction' and 'ECM receptor interaction' are present in module 6 as well as 6 and 4 connective cell clusters respectively. Panel C2 shows that 'cell cycle' and 'senesence' are present in module 21 as well as clusters 8 and 16, 'Notch signaling' genes are present in clusters 8 and 25 as well as module 21, and 'DNA replication' is present in clusters 8, 10, and 16.

found in our cluster based KEGG pathway analysis (**Figure 3D**). Module 21 also contained 'Notch signaling' genes, possibly due to similar correlations with cell cycle genes; however, these were only expressed in the neural cell clusters (8 and 25).

Where one cell cluster is associated with multiple gene networks, we wondered if multiple cell identities existed within the cell cluster. We explored this possibility by examining the pigment cell cluster and whether the five different correlated gene modules (#34, 35, 36, 37, and 38) are expressed in distinct subgroups of cells within the cluster. We found cases of more than one genetic network (modules 35 and 38) expressed in the same cells. Modules 35 and 38 contain conserved pigment genes, *pmela*, *mlana*, and *dct* in module 35 and *tryp1b* and *pmel* in module 38, that mark melanocytes (**Figure S6**; Du et al., 2003; Johnson et al., 2011; Lamason et al., 2005; B. Thisse & Thisse, 2004). However, we also found non-overlapping expression of networks, notably modules 36 and 37. We inferred that module 36 is associated with xanthophores and xanthoblasts due to the presence of *plin6* and *scarb1*, genes involved with lipid binding and activity in xanthophores (Ahi et al., 2020). On the other hand, module 37 likely represents iridophores and iridoblasts because it contains *pnp4a,* which is involved in purine-nucleoside phosphorylase activity in iridoblasts (Kimura et al., 2017).

Conserved signaling pathways are active during syngnathid craniofacial development

The specialized pipefish feeding apparatus is composed of an elongate, tubular snout, toothless mandible and pharyngeal jaws, large tendons, and associated muscles. Therefore, numerous cell types contribute to their distinct faces: cartilage, bone, tendon, muscle, and connective tissues as well as their progenitors. We sought to identify markers of these cell types and signaling pathways active in them.

We found marker genes uniquely expressed in the face, genes that mark cell types important to craniofacial development, and markers with potentially relevant functions for craniofacial development using *in situ* hybridizations of cell cluster marker genes. For instance, we observed the marker for osteochondro-mesenchymal cells (cluster #6), *elnb*, specifically expressed at the intersection between the ethmoid plate and palatoquadrate as well as on the Meckel's cartilage (**Figure 4C**). Although elnb is observed in the zebrafish cranial skeleton, it is primarily studied for its proposed role in teleost heart evolution (Miao et al., 2007; Moriyama et al., 2016).

Other genes identified here as cell markers were not uniquely craniofacial but provide insights into the cell types that build the face. For example, *tnmd* marked tendons and ligaments (cluster #9) throughout the face and body (**Figure 4E**; **Figure S3**). Our finding is consistent with *tnmd's* role in tenocyte development in model systems, namely zebrafish and mouse (J. W. Chen & Galloway, 2014; Docheva et al., 2005). Identifying tenocyte cells is particularly relevant in syngnathid fishes where tendons are enlarged and store elastic energy necessary for their specialized feeding (Van Wassenbergh et al., 2007).

The last category of markers contains genes with regulatory roles important for craniofacial development such as *prdm16,* the marker for osteochondro-mesenchymal cells (cluster #5). *prdm16* mediates the methylation of histones and regulates gene expression, key for promoting craniofacial chondrocyte differentiation (Ding et al., 2013; Kaneda-Nakashima et al., 2022; Shull et al., 2020), We found *prdm16* expressed in mesenchymal cells directly above the ethmoid plate and in fins (**Figure 4B**; **Figure S2**). While *prdm16* is also observed in developing limb/fins in mice/zebrafish, and in the palate of mice, *prdm16* had additional expression domains in those models: tongue, teeth, and Meckel's cartilage in mice, and in hindbrain and all pharyngeal arches in zebrafish (Bjork et al., 2010; Ding et al., 2013). Our observations are consistent with the literature, except *prdm16* has not been observed above the ethmoid in zebrafish.

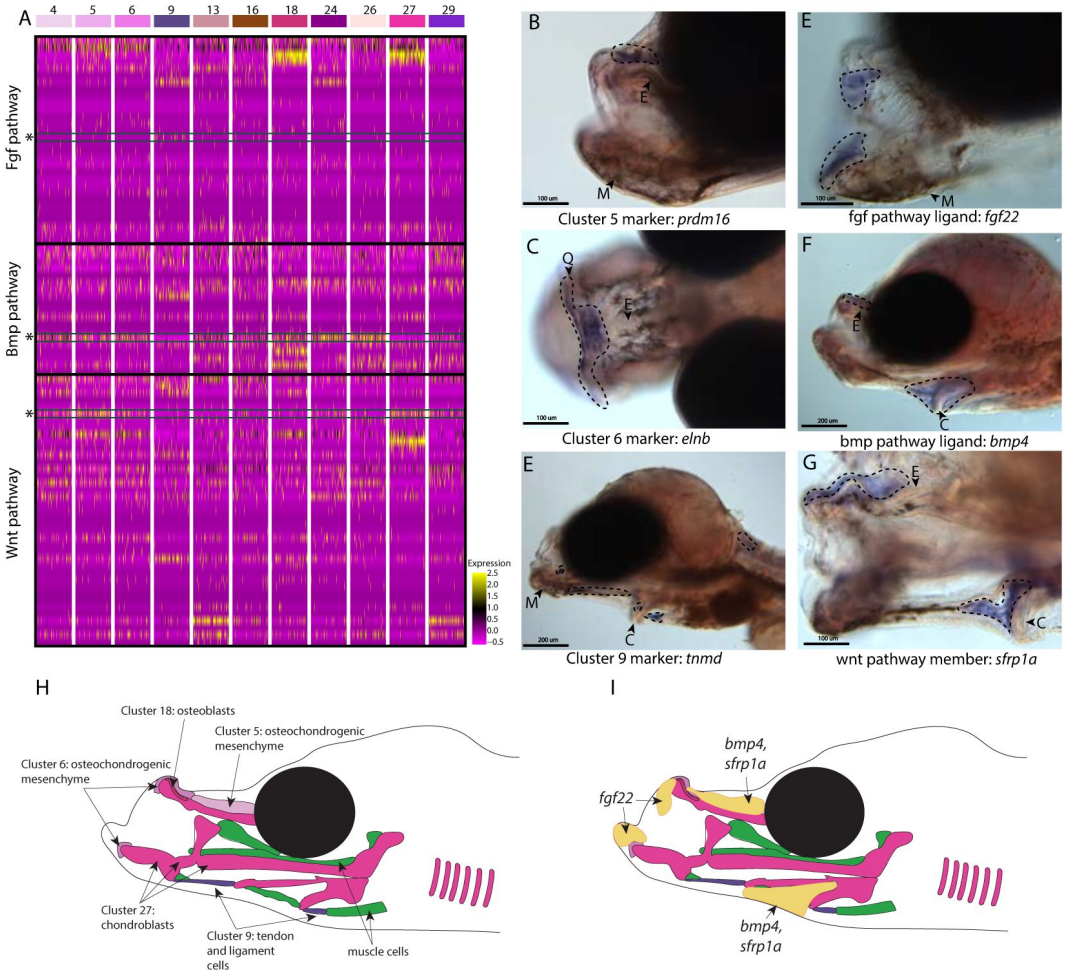

Figure 4

Conserved cell types and gene pathways build unique faces of syngnathids.

Three main conserved signaling pathways are enriched in connective cell types, MAPK signaling (including Fgf signaling), TGF-beta signaling (including BMP signaling), and Wnt signaling. Receptors and ligands expression patterns are shown in panel A heatmap from all cell types with cells present in the head. This heatmap features 100 cells downsampled from each cluster and illustrates that many genes from these families are expressed in these cells. Yellow lines indicate high expression of a gene, while hot pink lines indicate no expression. The pathways are boxed in black. Rows representing *fgf22*, *bmp4*, and *sfrp1a* expression are marked with an asterisk and green box for each respective section of signaling (Fgf, BMP, and Wnt). Panels B, C, and D are *in situ* hybridizations of three marker genes, *prdm16*, *elnb*, and *tnmd*. *prdm16* and *elnb* mark osteochondrogenic mesenchyme and *tnmd* marks tendons and ligaments. Panels E, F, and G, show expression patterns of three pathway representatives (*fgf22*, *bmp4*, and *sfrp1a*). All three genes are expressed in the face: *fgf22* at the tip of the mandible and *bmp4* and *sfrp1a* above the ethmoid and near the ceratohtyal. Staining is circled with dashed lines. The Meckel's cartilage (M), mes(Ethmoid) cartilage (E), Quadrate (Q), and Ceratohyal (C) are labeled. Panel C is a dorsal view. Panels B, C, E, F, G, H, I are in lateral view. *In situs* of *fgf22* were completed using 10dpf Gulf pipefish. *bmp4*, *sfrp1a*, *tnmd*, *elnb*, and *prdm16 in situs* were completed using wild caught bay pipefish at the onset craniofacial elongation. Panels H and I are summary illustrations of our findings, panel H shows where cells from various clusters were present in the developing head and panel I illustrates where *bmp4*, *sfrp1a*, and *fgf22* were expressed.

We next examined signaling pathways active in craniofacial development. Our KEGG pathway analysis revealed that MAPK, Wnt, and TGF-beta signaling pathways were significantly enriched in one or more craniofacial contributing cell clusters (module #6; **Figure 2B**, **Figure 3C**, **Figure 4D**). We chose three genes, one from each major pathway, for *in situ* hybridizations: *fgf22* (a fgf ligand present in the actively dividing cells module), *bmp4* (a bmp ligand present in the largest connective tissue module), and *sfrp1a* (a Wnt pathway enabler present in a cartilage gene module). We found *fgf22* is expressed at the tip of the palatoquadrate and Meckel's cartilage and in the pharyngeal arches (**figure 4E**). Our finding is surprising given that *fgf22* is primarily expressed in the nervous system and has never been observed in craniofacial skeletal development (Miyake & Itoh, 2013; Umemori et al., 2004). We observed *bmp4* and *sfrp1a* expressed above the ethmoid plate and along the ceratohyal (**figure 4F-G**). *Bmp4* has a conserved role in craniofacial development, particularly important at later stages for driving chondrocyte differentiation (Wang et al., 2024; Zhou et al., 2013). However, contrary to pipefish, *bmp4* in zebrafish is minimally expressed in the ethmoid region during cartilage formation (Schilling & Kimmel, 1997; Swartz et al., 2011; Wang et al., 2024). Interestingly, *sfrp1a* has not been observed in the palate of mice or ethmoid region of zebrafish, but *sfrp1a* craniofacial expression has been observed in other fishes with derived craniofacial structures (Ahi et al., 2014; Schneider, Gunter, et al., 2023).

Gulf pipefish retain tooth development genes but likely lack onset of tooth development

Previous papers have identified possible candidate genes for the loss of teeth in syngnathid fishes including those from genes that initiate tooth bud formation (*fgf4, eve1*), regulate tooth morphogenesis (*fgf3, fgf4*), and synthesize tooth minerals (*scpp4, scpp7, scpp9, odam,* and *scpp5*; Lin et al., 2016; Qu et al., 2021; Small et al., 2016; Small et al., 2022; Zhang et al., 2020). However, it is unknown whether syngnathid tooth development initiates then halts or whether it never begins. We searched for signs of early tooth primordia within our atlas to ask whether tooth development might initiate in syngnathids. Additionally, we examined whether genes present in mature teeth are still expressed in syngnathids and what types of cells express them.

Our cell cluster annotations did not identify a tooth primordium cluster. We therefore searched for tooth primordia by examining expression of specific fodontogenesis marker genes (*aldh1a2*, *bmp4*, *dlx2a*, *dlx3b*, *lef1*, *lhx6a*, *lhx8*, *msx1a*, *msx2*, and *wnt10a*; **Figure 5A**). We observed several primordium genes expressed in our atlas. However, there was no cluster with every marker gene expressed in over 10% of cells. Several markers (*dlx3b*, *lef1*, *msx1a*, *msx2*, and *wnt10a*) were expressed in cluster #28, a fin cluster distinguished by *hoxa13a* and *hoxa13b* expression. Since we previously noted that cluster #16 seems to be a primordial connective tissue cluster, we wondered if it could contain tooth primordial cells. In this cluster, we found the following percentage of expression of each gene in the cells: *bmp4* in 46%, *aldh1a2* in 29%, *lef1* in 24%, *dlx2a* in 18%, *dlx3b* in 16.4%, *eda* in 12.7%, *msx1a* in 10.6%, *pitx2* in 10%, *lhx6a* in 8%, *wnt10a* in 4.6%, *msx2* in 2.65%, *lhx8* in 1.8%, and *shha* in 1%.

We suggest that, given the low expression of most tooth marker genes, cluster #16 is unlikely to contain tooth primordial cells. To test this assertion, we examined spatial gene expression using *in situ* hybridization of *bmp4*, *pitx2*, *lhx6a*, *dlx2a*, and *dlx3b* in pipefish to ask whether the two definitive cell types are present, namely the dental epithelium (marked by *pitx2* and *bmp4*) and dental mesenchyme (distinguished by *dlx2a*, *dlx3b*, and *lhx6a*; **Figure 5B,C**,D,E,F; Gibert et al., 2019; A. Tucker & Sharpe, 2004). Tooth- specific expression of d*lx2a* is observed solely in the dental mesenchyme in zebrafish and mice, however, it is also expressed in the dental epithelium in medaka (Stock et al., 2006). For this study, we labelled it as a mesenchyme marker, though it could be expressed in both dental tissues in a syngnathid outgroup. We found expression of all genes except *dlx2a* in the developing jaws. However, *bmp4*, *pitx2*, *lhx6a*, and *dlx3b* were expressed throughout the jaw rather than in the punctate pattern observed in tooth primordia development.

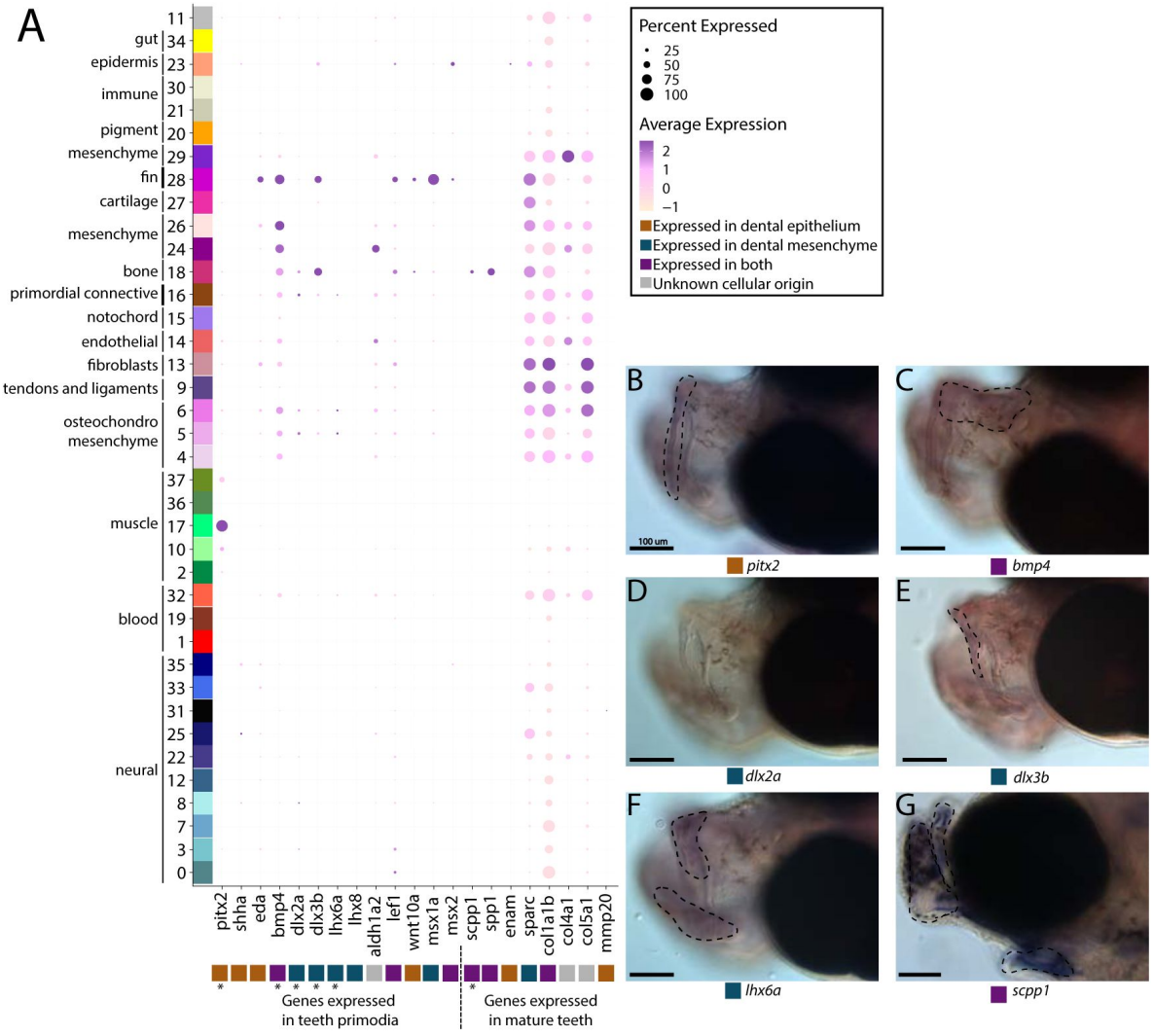

Figure 5

Pipefish do not possess identifiable tooth primordium cells, but continue to express tooth development genes in other contexts.

Panel A presents a dotplot of genes classified by the tissue layer in which they are reported to be expressed during tooth development in other vertebrate (Tucker and Sharpe 2004; Gilbert et al. 2019; Kawasaki 2009). The x-axis contains the assayed genes, with asterisks under the genes that were also examined with *in situ* hybridizations. The y-axis contains all of the cell clusters. The size of the dot is representative of the percentage of cells from the cluster that express the gene. The color of the dots is average expression of the gene in the cluster (darker purples represent higher expression). Panel B includes *in situ* hybridizations of selected tooth primordia markers (*bmp4*, *pitx2*, *lhx6a*, *dlx2a*, and *dlx3b*) and mature tooth marker (*scpp1*). *In situs* of *bmp4*, *pitx2*, *lhx6a*, *dlx2a*, and *dlx3b* were completed with wild caught bay pipefish that had begun craniofacial elongation. *scpp1 in situs* were completed using 9dpf Gulf pipefish. The scale bars for all images represents 100 μM.

We next investigated Scpp genes (*enam*, *scpp1*, and *spp1*), which are expressed in late tooth development and tooth maintenance in other vertebrates (**Figure 5A**). These genes also have some expression outside of teeth such as in dental bone (*scpp1*, *spp1;* Kawasaki, 2009) and fins (*enam;* Jain et al., 2007). We identified *spp1* and *scpp1* expression in 54.5% and 24% of bone cells, respectfully, and sparse *spp1* expression (<3%) in other connective tissue cell types. We found minimal expression of *enam* (in 11.4%) of epidermal cells and in less than 8% of muscle cells. In wholemount *in situ* hybridization, we found that *scpp1* is expressed in all developing pipefish bones, both endochondral and dermal (**Figure 5G**). Since Scpp gene losses observed in syngnathids have been hypothesized to be responsible for their tooth loss (Lin et al., 2016a; Qu et al., 2021; Zhang et al., 2020), we explored Scpp gene family content in the close toothed relative to syngnathids, the blue spotted cornetfish, and we found several gene losses (*scpp4*, *scpp7*, and *scpp9*) and a likely pseudogene (*scpp5*) (**Figure S7**; Hughes et al., 2018; Stiller et al., 2022).

Exploration of additional tooth maturation genes (*col1a1b, col4a1, col5a1*, *sparc*, and *mmp20b*) similarly found that these genes were expressed in non-tooth derivatives, including connective tissue, smooth muscle, and neural cells.

Tooth and skeletal genes are expressed during dermal armor development

Syngnathid dermal armor is mineralized dermal bone underneath the skin (**Figure 6D**, G, H). It is unknown when dermal armor primordia initiate and how they are patterned. Spatial expression analysis of *pitx2* and *dlx3b* in search of tooth primordia instead revealed expression of these genes in possible dermal armor primordia (**Figure 6A,B**, E,F). We found *pitx2* staining localized dorsally to striated muscle underneath developing dermal armor. *dlx3b* is expressed in a repeating pattern along the body in the epidermal and dermal tissues. Both staining patterns were in discrete regions of the muscle and epidermal layers rather than being continuously expressed across the tissues. We did not observe the expression of other tooth primordium genes (*bmp4, lhx6a*, and *dlx2a*) in this region.

Because the embryos from our atlas had not begun dermal armor mineralization, the atlas cannot be directly used for discovery of genes active in dermal armor. However, the atlas contains osteoblasts from craniofacial bones which we used to create osteoblast specific gene networks. We therefore asked whether these osteoblast genes were present in mineralizing dermal armor at later stages. *In situ* hybridization expression analysis revealed that *scpp1*, an osteoblast and tooth mineralization gene, and *ifitm5*, an osteoblast gene, were expressed in the dermal armor at the onset of mineralization (**Figure 6C**; **Figure S3**).

Epithelial expression of immune and nutrient processing genes may facilitate embryo-paternal interactions in the brood pouch

Within the brood pouch, embryos could interact with male placenta-like tissues, the male brood pouch epithelium, and/or the pouch microbiome. Once the thin chorion is shed, the embryos' epidermis is directly exposed to the pouch environment. We therefore asked if the embryonic epidermal cells expressed nutrient acquisition and/or immune genes that would indicate active transfer of nutrients and immune response.

Within our larger KEGG analysis, we asked whether nutrient absorption KEGG terms were among the enriched pathways for the epidermal cells. We identified 106 enriched genes in the endocytosis pathway (p-value=.036; **Figure 7A**). Four metabolism pathways (galactose, glutathione, sphingolipid, and starch and sucrose) are also enriched. For comparison, we investigated whether these KEGG terms are also enriched in epidermal cells of non-brooding fishes, we completed a KEGG pathway analysis on a comparably staged zebrafish single cell RNA sequencing atlas (3dpf;

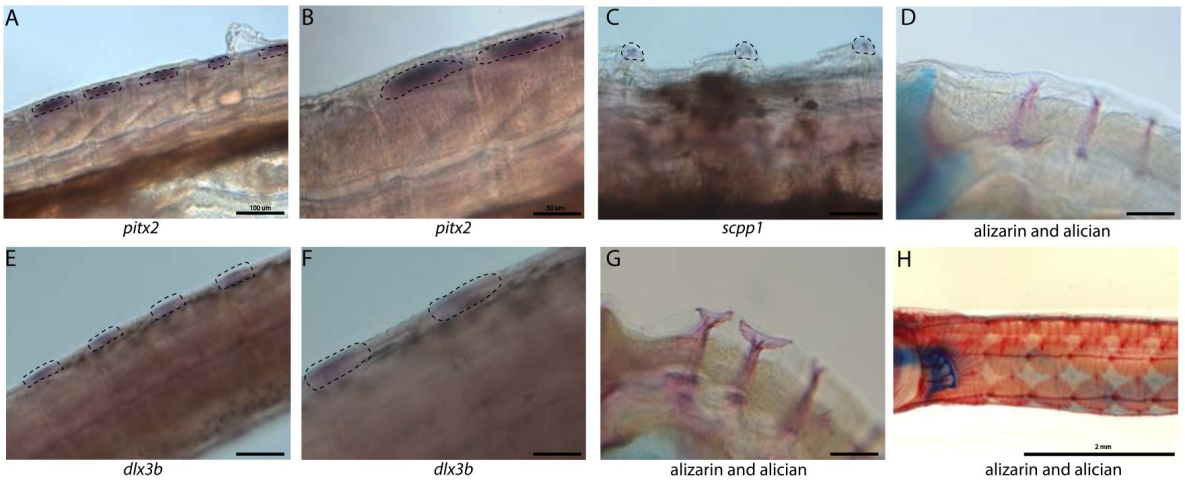

Figure 6

Tooth and bone development genes expressed during exoskeleton development.

We discovered *pitx2* (A1 and A2) and *dlx3b* (B1 and B2) expression during possible emergence of exoskeletal primordium in wild caught bay pipefish. The embryos used for these *in situs* were the same stage as those from **Figure 5**, at the beginning of craniofacial elongation. Figures A1 and B1 have 100 μM scale bars, figures A2 and B2 are taken from the same fish and have 50 μM scale bars. We further found *scpp1* is exppressed at the mineralization front of the exoskeleton in 12dpf Gulf pipefish (C) and has a 50 μM scale bar. Alizarin and alcian stained pipefish are shown in panels D1 (12dpf), D2 (1dpf), and D3 (adult Gulf pipefish) to illustrate how the exoskeleton forms. Panels D1 and D2 have 100 μM scale bars.

Lange et al., 2023). The zebrafish epidermal cells did not have a significant enrichment of the endocytosis pathway (23 genes, p-value=.99) or any metabolism pathway. However, 11 of the 23 endocytosis genes expressed in the zebrafish epidermis were also present in the pipefish epidermis.

We next examined epidermal gene networks. We found that the largest epidermal gene network (#16) contained a striking enrichment of C type lectin genes, carbohydrate binding proteins that possess antimicrobial properties (**Figure 7B**). This network contained 14 total lectin genes expressed in the epidermal cells: five galactose-specific lectin nattectin, 2 alpha-N-acetylgalactosamine-specific lectins, one L-rhamnose-binding lectins, four ladderlectin, one C-type lectin 37Dd-like, and one C-type lectin domain family 4 member G-like. Through examining lectin gene expression in the entire dataset, we found that these genes were specific to epidermal cells. Interestingly, previous literature has identified an upregulation of C-type lectins in brood pouch tissues throughout different stages of syngnathid pregnancy (Roth et al., 2020; Small et al., 2013; Whittington et al., 2015).

Unlike pipefish, we did not find any expression of C-type lectin genes in zebrafish epidermal cells (**Figure S8**). However, C-type lectin genes that were unannotated in zebrafish may have been missed. Interestingly, the 'C-type lectin receptor signaling pathway' was significantly enriched in zebrafish epidermal cells (13 genes, p-value=.04) but not in pipefish epidermal cells (34 genes, p-value=.4). Although these results seem paradoxical, the KEGG term 'C-type lectin signaling pathway' does not include any of the C- type lectin genes themselves. Additionally, zebrafish genes from this pathway include R- Ras2, p38 MAP kinase, and NF-kappa-B inhibitor alpha, all of which are present in 10 or more different KEGG pathways. Overall, these data suggest that the expression of C-type lectin genes in the pipefish embryonic epidermis is potentially unique and warrants further investigation.

Discussion

Our study examines the development of syngnathids, with a particular focus on novel and adaptive characters, using single cell RNA sequencing of Gulf pipefish embryos coupled with *in situ* analysis of gene expression. Our single cell atlas represents early craniofacial skeleton development in Gulf pipefish, at a stage when the cartilages of the head skeleton were formed but the face has not elongated. We used the atlas to explore craniofacial and dermal armor development and to investigate potential interactions between the embryos and the brood pouch environment. Our dataset is both an opportunity to explore the developmental genetic underpinnings of syngnathid innovations, and a resource for teleost researchers for future studies in this fascinating lineage.

Our atlas represents a novel resource for Evo-Devo research

Developmental single cell atlases have elucidated cell identities and genetic pathways active in zebrafish, mice, and chick (Farnsworth et al., 2020; Farrell et al., 2018; Feregrino et al., 2019; Morrison et al., 2017; Soldatov et al., 2019; D. E. Wagner et al., 2018; Williams et al., 2019). In less traditional models, the majority of scRNAseq atlases are produced from adult tissues, allowing investigations into cell types, population differences, and genetic networks (e.g. Chari et al., 2021; Fuess & Bolnick, 2023; Hain et al., 2022; Hong et al., 2023; Koiwai et al., 2023; Parker et al., 2022; Potts et al., 2022; Royan et al., 2021; Songco-Casey et al., 2022; Vonk et al., 2023; Woych et al., 2022). Creation of developmental scRNA-seq atlases in non-model organisms is just beginning to accelerate, but many emerging models still lack such a resource, limiting our understanding of their development (e.g. Healey et al., 2022; Salamanca-Díaz et al., 2022; Steger et al., 2022; Ton et al., 2023). Our single cell atlas is one of the first created to understand the development of derived traits in a non-model organism.

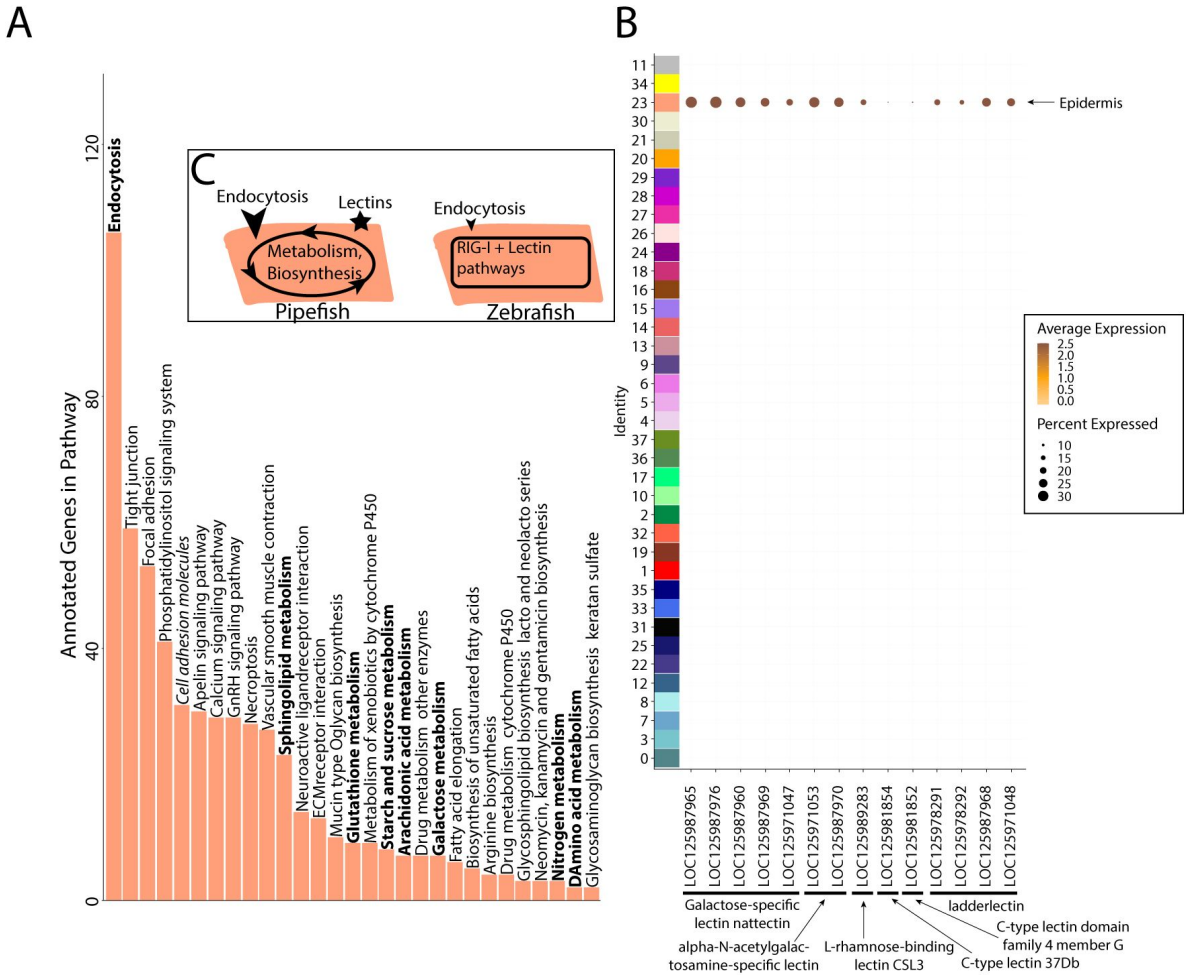

Figure 7

Gene expression signatures suggest embryonic interactions within brood pouch environment.

Epidermal cells (panel A), with pathways that suggest increased endocytosis and metabolism pathways are in bold text. Pathways upregulated in 3dpf zebrafish epidermal cells are in italics. Pipefish epidermal cells also express 15 lectin genes not found in other cell types (panel B). We suggest an epidermal cell model (panel C), in which we predict pipefish have an enrichment of nutrient processing genes and lectins in comparison with zebrafish cells.

For syngnathids specifically, this atlas represents an important step towards understanding the genetic nature of unique syngnathid traits. Numerous developmental genetic changes can lead to evolutionary innovations, including the evolution of novel genes, gene duplications, gene losses, gene family expansions or contractions, evolution of regulatory elements, co-option of gene regulatory networks, re-wiring of gene networks, assembly of novel gene networks, and/or the emergence of novel cells (Arendt et al., 2016; Cãestro et al., 2007; Teichmann & Babu, 2004; A. Wagner, 2011; G. P. Wagner & Lynch, 2010). Considering these possibilities, we examined select syngnathid traits and speculated on developmental genetic mechanisms influencing their evolution. Continuing to investigate these proposed mechanisms through expanded scRNAseq atlases and other studies will be critical for understanding syngnathid evolution.

Conserved pathways may contribute to derived syngnathid heads

Syngnathids have highly derived heads including an elongated ethmoid region, uniquely shaped hyoid, and altered muscles and tendons to support specialized “pivot feeding”. The developmental underpinnings of these derived traits have remained underexplored. In our atlas, we identified cell types present in the developing pipefish head and genetic pathways active in those cell types. We identified numerous cells that were present in the developing face: cartilage, bone, tendons, ligaments, osteochondrogenic mesenchyme, fibroblasts, and unclassified connective tissue cells. Overall, we did not find any unrecognizable cell types, suggesting that genetic modifications within conserved cell types may drive craniofacial modifications.

We next investigated signaling pathways expressed in these cells to determine whether and to what extent developmental genetic reorganization might have occurred. Specifically, we examined the expression of one gene each from three different highly conserved signaling pathways: Wnt (*sfrp1a*), TGF-beta (*bmp4*), and MAPK (*fgf22*). Using *in situ* hybridizations, we found *sfrp1a* and *bmp4* expressed dorsal of the elongating ethmoid plate and surrounding the ceratohyal, suggesting that Wnt and BMP signaling may be active in the lengthening structures. *sfrp1a* and *bmp4* are proposed to influence the development of elongated and broadened craniofacial morphologies in other species (Ahi et al., 2014; Schneider, Gunter, et al., 2023; A. S. Tucker et al., 2000). Pipefish *prdm16* is similarly expressed dorsal to the elongating ethmoid plate. Since *prdm16* regulates Wnt and TGF- beta signaling and these genes regulate chondrocyte differentiation (Bjork et al., 2010; Kaneda-Nakashima et al., 2022; Shull et al., 2022; Wang et al., 2024), their co-expression might suggest a prolonged period of chondrocyte differentiation along the pipefish ethmoid region.

We found *fgf22* expressed in the mandible and pharyngeal arches using *in situ* hybridizations, but not in the elongating regions of the head. Interestingly, *fgf22* has not been reported in craniofacial development of any other species (Miyake & Itoh, 2013). Fgf signaling, however, is a conserved and essential pathway for craniofacial development (Crump et al., 2004; Leerberg et al., 2019; McCarthy et al., 2016; Szabo-Rogers et al., 2008; Walshe & Mason, 2003; Woronowicz & Schneider, 2019), raising the possibility that *fgf22* has been co-opted into a role played by different Fgf genes in other species. Future work should investigate whether *fgf22* craniofacial expression is observed in other species and consider a relationship between *fgf22*’s novel craniofacial expression and the loss of Fgf ligands *fgf3* and *fgf4* in syngnathids. If *fgf22* is active in existing gene networks, particularly those where *fgf3* or *fgf4* is active in other species, then its novel expression may indicate evolved genetic compensation.

Our analysis suggests ways in which unique syngnathid craniofacial structure could have evolved through genetic network evolution. Unusual expression location (e.g., *fgf22* and *sfrp1a*) and timing (e.g., bmp4 and *prdm16*) in pipefish compared to zebrafish suggests that changes in signaling gene expression and/or content within craniofacial gene networks, particularly genes from Wnt, Fgf, or TGF-beta families could underly the exceptionally elongated syngnathid face.

No evidence of primordia suggests changes in early, not late, tooth development are at the root of evolutionary tooth loss

Tooth loss has occurred independently in numerous lineages and has often been studied to understand the developmental basis of character loss. For instance, research in birds and turtles found that tooth programs initiate but are subsequently truncated, explaining toothlessness in mature animals (Y. P. Chen et al., 2000; Tokita et al., 2013).

Additional studies in birds have found losses in tooth maturation genes (specifically scpp genes, Sire et al., 2008). Since numerous primordium and maturation genes are lost in syngnathids (Lin et al., 2016; Qu et al., 2021; Small et al., 2016; Small et al., 2022; Zhang et al., 2020), we asked if syngnathids begin tooth development at all.

We found that early tooth development genes were still expressed in pipefish, which is unsurprising given their pleiotropic roles, but found no convincing evidence either in the atlas or in spatial gene expression analysis tissues with tooth primordium identity. It is possible that syngnathid *fgf3* and *fgf4* losses resulted in insufficient Fgf signaling from the oral epithelium to the dental mesenchyme for tooth initiation (Small et al., 2022; Stock et al., 2006).

Despite having lost the earliest stages of tooth development, why then have syngnathids retained some members of the Scpp tooth maturation gene cluster? Since studies in birds propose Scpp gene losses can occur from relaxed selection (Sire et al., 2008), we speculated syngnathids lost Scpp genes with expression limited to teeth and retained genes with ancestrally pleiotropic expression patterns. In Gulf pipefish, we found the retained genes *scpp1*, *spp1*, and *enam,* are expressed in structures outside of tooth development, suggesting developmental pleiotropic constraint. Specifically, we found *spp1* and *scpp1* expressed in osteoblasts which is consistent with zebrafish (Bergen et al., 2022; Kawasaki, 2009; Liu et al., 2016) and *enam* expressed in the epidermis which has not been reported in zebrafish (Goldsmith et al., 2003; Jain et al., 2007; Liu et al., 2016). Through examining the conservation of the scpp genes in close syngnathid relatives, we found that most scpp gene losses (*scpp4*, *scpp7*, and *scpp9* and a functional *scpp5*) are shared with a tooth-bearing outgroup to the family, and likely occurred prior to the loss of teeth in syngnathids. Overall, our analysis favors the hypothesis that pleiotropic Scpp genes were retained in syngnathids while other Scpp genes were lost due to relaxed selection.

Redeployment of the bone gene network to build dermal armor

The syngnathid dermal armor is a type of evolutionary novelty, which can arise through either differentiation of serially repeated elements or *de novo* origination, derived from either the redeployment of existing gene networks, rewiring of existing gene networks, or the assemblage of new gene networks (Wagner & Lynch, 2010). Currently, there is no understanding of the developmental genetic underpinnings of the syngnathid dermal armor.

We identified *dlx3b* and *pitx2* expression in tissues where the dermal plates mineralize later in development. Using *in situ* hybridizations, we showed that epithelial and dermal layers expressed *dlx3b* and underlying muscle cells expressed *pitx2*. In some species, dermal bone, plate, or denticle development occurs from the re-deployment of tooth gene regulatory networks (Mori & Nakamura, 2022). However, this does not appear to be the case in syngnathids because the dermal armor lacks the characteristic epithelial– mesenchyme interactions distinguished by *pitx2* expression in the epithelia.

Instead, dermal armor might originate from the co-option of existing bone development gene regulatory networks. *Dlx3b* expression has been observed in epithelia and mesenchyme during dermal and perichondral bone development in zebrafish (Verreijdt et al., 2006). In addition to

dlx3b, we observed bone development genes *scpp1* and *ifitm5* expressed in the ossifying dermal armor. Future studies could test our hypothesis that the dermal armor evolved through re-deployed osteoblast networks by examining osteoblast gene network expression over time.

Signatures of embryonic interactions within the novel pouch environment

Syngnathid embryos are reared within the brood pouch, a novel structure and environment composed of male derived tissues (epithelium and placental-like tissues that include specialized cell types) that harbors a pouch microbiome (Stölting & Wilson, 2007). During pregnancy, the male brood pouch undergoes numerous changes including increased vascularization and altered expression of immune genes (Harada et al., 2022; Ripley et al., 2010; Roth et al., 2020; Small et al., 2013; Whittington et al., 2015). Researchers predict that these changes relate to nutrient and waste transfer and prevention of embryonic rejection and bacterial infection (Dudley et al., 2021; Whittington & Friesen, 2020).

However, there are few studies that examine whether and how embryos interact with the brood pouch environment (Kvarnemo et al., 2011; Ripley & Foran, 2006). To consider whether embryos have specializations for life in this brood pouch environment, we asked about cell type-specific expression of nutrient acquisition and/or immune genes.

Pipefish embryos uptake paternally derived carbohydrates, proteins, and lipids (Kvarnemo et al., 2011; Ripley & Foran, 2006). Our data suggests that this uptake could occur through the embryonic epidermis. Specifically, we noticed an enrichment of endocytosis and metabolism genes in epidermal cells. Epidermal absorption of maternally derived nutrients has been suggested in viviparous fishes (Tengfei et al., 2021; Wourms, 1981). Interestingly, microvilli, a type of cellular projection, have been observed on the anal fin of developing seahorses (Wetzel & Wourms, 2004). Possibly, these microvilli may be functionally equivalent to those on the small intestine, maximizing nutrient absorption from the environment.

During pregnancy, the male brood pouch increases expression of C-type lectin genes (Roth et al., 2020; Small et al., 2013; Whittington et al., 2015). These genes are transmembrane or secreted receptors that sense self or non-self and are primarily studied for their role in innate and adaptive immunity (Brown et al., 2018). We identified 14 C-type lectin genes expressed in the embryonic epidermis. Our work suggests that lectin genes are produced by both the father and the embryos, but their function is still unclear. Syngnathid research has primarily suggested that lectin genes are produced to prevent bacterial infection (Melamed et al., 2005), though they could be important for male-embryo recognition.

Overall, our findings suggest that pipefish embryos have evolved to be specialized for development within the brood pouch by expressing genes related to nutrient acquisition and immunity. Future studies could provide insights into when nutrient acquisition and lectin genes are expressed in development, their functional role, and how their expression varies across syngnathid lineages that have exposed versus enclosed embryos, for example, to examine how embryonic development has been impacted by the brood pouch.

Conclusions

Our study represents the first scRNAseq developmental atlas in syngnathids, and one of the first non-model developmental scRNAseq atlases, providing a major step forward for evo-devo research. We used our atlas to begin addressing questions on the evolution and development of syngnathid innovations including their unique craniofacial structure, loss of teeth, dermal armor, and development within the male brood pouch. By combining scRNAseq analysis with spatial

expression data from *in situ* hybridization, we made important discoveries in cell type identity and distribution as well as spatial expression of marker and signaling genes. We found that syngnathids express conserved signaling pathways during craniofacial development, suggesting that alterations within these pathways may be important for the evolution of their craniofacial skeletons. We did not find evidence for tooth primordia within syngnathids and propose that genetic changes early in tooth development could have led to their loss of teeth. We propose that the re- deployment of bone gene networks, but probably not tooth networks, could play a role in the dermal armor development. Finally, we observed an enrichment of endocytosis genes and many C-type lectin genes in epidermal cells, which suggests ways these cells might interact with the brood pouch environment. Our atlas pushes forward our understanding of syngnathid development and evolution, and provides resources for developmental genetic analysis in nascent evo-devo model species.

Methods

Single Cell RNA Sequencing Libraries Preparation

We created scRNAseq atlases from embryos of wild caught Gulf pipefish (*Syngnathus scovelli*, acquired from our collaborator Emily Rose), and all work was performed under an approved IACUC protocol. We harvested 20 embryos per pouch from two wild caught male pipefish. The embryos were at a stage before the tubular face was fully elongated, and while the head skeleton is cartilaginous with minimal signs of mineralization of superficial intramembranous bones. This corresponds to a stage termed “frontal jaws” in a recent description of pipefish development (Sommer et al., 2012).

We dissociated the embryos using 460ul of .25% trypsin in water and 40ul 100 mg/mL collagenase I (Sigma C0130-200mg) for 16 minutes. We filtered cells using a 40uM cell strainer (Thomas Scientific #1181X52). We quantified cell concentrations using the TC20 Automated Cell Counter (Biorad) then diluted the samples to 800 cells/ul in 4% BSA in PBS. The University of Oregon Genomics and Cell Characterization Core (GC3F; *https://gc3f.uoregon.edu*) prepared single cell libraries for each sample using 10X

Genomics Single Cell 3’ Genome Expression mRNAseq kit with NextGEM v3.1 chemistry. We sequenced these libraries on an S4 lane on the NovaSeq 6000 at the GC3F. To improve the 3’ UTR genome annotations, we also prepared scISOrSeq libraries from the first embryonic sample and from dissociated pouch cells from pregnant and nonpregnant males. These libraries were produced in accordance with (Healey et al., 2022). Embryonic, pregnant pouch and non-pregnant pouch libraries were sequenced separately on PacBio Sequel II - SMRT Cells 8M.

To turn the scISOrSeq reads into gene models, we followed the pipeline from Healey et al. 2022. We ran our custom script (scISOr_Seq_processing.py) to remove barcodes, identify cell barcodes, and demultiplex with the single cell flag and appropriate barcodes (5’ CCCATGTACTCTGCGTTGATACCACTGCT and 3’ CTACACGACGCTCTTCCGATCT). We aligned the reads to the 2022 Gulf pipefish genome GenBank: GCA_024217435.2) using minimap v2.9 (cite: Li 2018). We filtered the reads using cDNA cupcake to remove duplicate transcripts (Tseng, 2021). We used SQANTI3 to identify gene models and filter them (Tardaguila et al., 2018). We merged the SQANTI3 annotations with the Gulf pipefish genome (NCBI: ROL_Ssco_1.1) using TAMA merge (Kuo et al., 2017). Since the Gulf pipefish genome does not contain mitochondrial genes, we appended the genome files with the Gulf pipefish mitochondrial genome (NCBI RefSeq: NC_065499.1).

Single Cell Atlas Construction

We ran Cell Ranger (10X Genomics 3.0.2) using our scRNAseq reads, the Gulf pipefish genome assembly with mitochondrial genome, and the modified gene annotations. Cell Ranger estimated 20,733 cells for sample one, 23,682 genes expressed, and 21,039 mean reads per cell. For sample two, Cell Ranger predicted 17,626 cells, 23,740 genes expressed, and 29,804 mean reads per cell. We analyzed Cell Ranger's output using Seurat (v4.1.0) on R (v4.0.2; Butler et al., 2018; Hafemeister & Satija, 2019) on R (v4.0.2).

To remove extraneous RNA counts from the dataset, we used SoupX (v1.5.2; Young & Behjati, 2020). We identified doublet scores for our dataset using scrublet (v0.2.3). The doublet removal step reduced the first sample by 114 cells (from 20,733 cells to 20,619 cells) and second sample by 167 cells (from 17,626 cells to 17,459 cells). We finally removed cells with less than 500 features, greater than 9000 features, greater than 1E5 RNA counts, with a scrublet score greater than the detected threshold (.76 for sample 2 and .21 for sample 2), or greater than 10% mitochondrial reads. The second filtering step removed 727 cells from sample one (20,619 cells to 19,892 cells) and 1,566 cells from sample two (from 17,459 cells to 15,893 cells).

We normalized the datasets with SCTransform (v0.3.3). We used Seurat's integration tools, SelectIntegrationFeatures using 3,000 feature genes, FindIntegrationAnchors using SCT normalization, and IntegrateData using SCT normalization, to integrate the two datasets. After integration, our combined atlas had 35,785 cells (Supplementary File 1,2). We then used the integrated dataset to complete PCA analysis. We tested using a variety of principle components for further analysis and chose 30 PCs for our analysis based on the clear delineation of major cell types. We next clustered the cells using 30 PCs and plotted the data on a UMAP with Seurat.

Single Cell Atlas Cluster Identification

To identify cluster identities, we used the RNA assay of the scRNAseq data to find cluster markers with Seurat's FindAllMarkers command with the parameters only.pos = TRUE and logfc.threshold = 0.25, requiring markers to be upregulated in the cluster and have a log fold change of at least .25. We found a second set of cluster markers through our custom function which searched through all genes and identified genes uniquely expressed in greater than 60% of cells in the cluster and in less than 10% of cells in every other cluster using Seurat's DotPlots. We searched for our identified markers in available zebrafish datasets (Fabian et al., 2022; Farnsworth et al., 2020; Lange et al., 2023), ZFIN (Howe et al., 2013), NCBI, medlineplus.gov, and genecards to give the clusters initial annotations. For each cluster, we used multiple genes to propose the cluster identity.

Next, we identified one gene for each cluster which marked the cluster best (expressed in the most cells in the focal cluster and expressed in as few of the other clusters as possible) through consulting the two marker gene lists and examining markers with Dot Plots (Supplementary File3). Using these markers, we completed a set of *in situ* hybridizations to hone our cluster annotations. Due to challenges culturing Gulf pipefish, we used both embryos and larvae from *Syngnathus leptorhynchus*, a local pipefish from the same genus, and from cultured Gulf pipefish for the cluster annotation *in situ*s. We caught pregnant male *Syngnathus leptorhynchus* using a beach seine near Coos Bay, Oregon under Oregon Department of Fisheries and Wildlife permit number 26987.

Syngnathus scovelli used for *in situs* were purchased from Alyssa's Seahorse Saavy and Gulf Specimens Marine Lab then reared in our facility at 25 degrees C water and 25-28 PPT Salinity. Since *Syngnathus leptorhynchus* does not have a published genome, we designed probes using NCBI Primer Blast with the Gulf pipefish genome and produced these probes using Gulf pipefish

embryonic cDNA pools. To create the probes, we completed two rounds of PCR. The first round used a gene specific forward primer with a reverse gene specific primer that had 10 nucleotides of T7 promotor sequence attached.

PCR products were cleaned with Zymo clean and concentrator DNA kit and eluted in 15 ul of elution buffer. Round two of PCR used the same gene specific forward primer with a modified T7 promoter sequence (TGGACTAATACGACTCACTATAGGG) as the reverse primer, and finally the product was cleaned again with Zymo clean and concentrator DNA kit and eluted in 15 ul of elution buffer.

Round one PCR conditions are as followed: 95 degrees Celsius for 3:00 minutes, 40 cycles of denaturation (95 degrees for 30 seconds), annealing (30 seconds, annealing temperature varied by probe), and extension (72 degrees for 1:00 minute), and a final extension step of 3:00 minutes. Round two PCR conditions are as followed: 98 degrees Celsius for 30 seconds, 30 degrees for 10 seconds, 72 degrees for 50 seconds, then 35 rounds of denaturation (98 degrees for 10 seconds), annealing (50 degrees for 10 seconds), and extension (72 degrees for 50 seconds), and a final extension step of 10 minutes at 72 degrees. For round two PCRs with multiple bands (specifically, *ifitm5*), the band of the expected size was excised with a razorblade and DNA was extracted with Zymoclean gel DNA recovery kit.

The round two PCR product was sanger sequenced to confirm identity. All the marker genes primers and successful PCR conditions are in Supplementary File 4. The probes were transcribed with T7 polymerase for 2-6 hours then cleaned with Zymo RNA clean and concentrator and eluted into 30 ul of water. For the *in situ* hybridizations, we selected embryos and newly spawned larvae close to the developmental stage used in the atlas. We completed *in situ* hybridizations in keeping with Thisse & Thisse, 2007, leaving the embryos in stain until background was observed. For spawned larvae, we completed a bleaching step (1% H2O2 and .5% KOH for 8 minutes) prior to the proteinase K digest. After imaging, we used the levels tool in Adobe Photoshop (v23.4.2) to white balance the pictures.

Single Cell KEGG Analysis

To identify pathways upregulated in cell clusters, we completed a KEGG analysis. We downloaded gulf pipefish KEGG pathways (*https://www.kegg.jp*). For the KEGG analysis, we used the marker genes identified from our FindAllMarkers list as the input genes. We converted these gene ids using keggConv to KEGG ids. We used a Wilcoxon enrichment test to ask whether cluster marker genes were enriched for each KEGG pathway.

Single Cell Atlas Network Analysis

To identify genetic networks present in our atlas, we completed a weighted gene network correlation analysis using WGCNA (v1.72-1,Langfelder & Horvath, 2008). We selected 3,000 variable features from the integrated assay of the single cell dataset for the WGCNA. We created an adjacency matrix from the data using bicor with a maxPOutliers of .05. To decide on a beta value or the soft threshold power, we created an adjacency matrix plot using pickSoftThreshold and picked the threshold where the scale free topology model fit leveled off. We selected the value of two. Then, we raised the adjacency matrix to the power of two.

We calculated the dissimilarity matrix by calculating TOM similarity of the adjacency matrix and subtracting it from one. We then created a gene tree through hclust of the dissimilarity matrix with the average method. From the gene tree, we created modules using cutreeDynamic with deepSplit of two and the minClusterSize of 15 (setting the smallest cluster size to 15). We calculated module eigengenes using moduleEigengenes of the adjacency matrix then calculated the dissimilarity of the module eigengenes with cor of the module eigengenes subtracted from one and clustered the module eigengenes with hclust. We then chose a dissimilarity of 35% as the cut

off for merging modules; however, no modules had dissimilarity scores with each other below 40%. Since the module eigengenes are calculated for each cell in the dataset, we used these values to calculate the t-statistic for each cell cluster (considered each sample) in the module (all of the module eigengene scores were used as the population).

We next tested the hypothesis that certain cell clusters were strongly associated with specific gene modules through a two-way permutation test (1000 permutations) and corrected p-values to control the False Discovery Rate (FDR). Additionally, we tested whether specific cell clusters drove underlying gene network structure by measuring network connectivity. The network connectivity was first calculated using the entire dataset, then each cell cluster was progressively dropped from the dataset and the connectivity was remeasured. This resulted in change in connectivity scores for each module-cluster pair. To assess whether these changes were significant, we completed 1000 permutations whereby cells were randomly dropped (the number of cells dropped was equal to the cell cluster size of the focal cluster), connectivity measured, and the change in connectivity recorded. The p-value is the number of instances where the change in connectivity is greater in the permutations than in the focal cell cluster run. P-values were corrected to control the FDR.

Using the module gene lists, we identified the number of genes from each module that were found in each KEGG pathway. Since the KEGG modules do not have p-values associated with the genes, we could not complete a Wilcoxon enrichment test. We instead removed any pathways where there were less than three genes present for any pathway and note that there is no statistical test run on these KEGG results. We visualized the networks using Cytoscape (v3.10.0).

In situ Hybridization

For genes chosen for further follow up analysis, we completed *in situ* hybridizations in Gulf pipefish (*Syngnathus scovelli*) or bay pipefish (*Syngnathus leptorhynchus*). Primer sequences were designed using NCBI Primer Blast with the Gulf Pipefish genome and synthesized using Gulf Pipefish embryonic cDNA. The *sfrp1a* probe was synthesized using the PCR based probe preparation protocol described in the Single Cell Identification section. All other probes (*bmp4*, *dlx2a*, *dlx3b*, *fgf22*, *lhx6a*, *pitx2*, and *scpp1*) were prepared using TOPO cloning. Probe primer sequences as well as the species used for the *in situ*s are described in supplementary file In_Situ_Probe_and_Samples_Used.xlsx. Pregnant male bay pipefish were caught as described above. For Gulf pipefish embryos, we allowed the fish to mate in our facility then harvested embryos once they reached the appropriate stages. Fish were reared in 25 degrees C water with 25-28 ppt salinity. We completed *in situ* hybridizations in keeping with Thisse & Thisse, 2007, leaving the embryos in stain until background was observed. After imaging, we used the levels tool in Adobe Photoshop (v23.4.2) to white balance the pictures.

Bone and Cartilage Staining

We used alcian and alizarin stains to mark cartilage and bones. We followed the protocol from Walker and Kimmel (2007) with minor alterations. We stored samples in 50% glycerol/0.1% KOH at 4C and imaged them in 100% glycerol. After imaging, we white balanced the photographs using Photoshop (v23.4.2) levels tool.

Gene Cluster Analysis

To examine close syngnathid outgroups, we downloaded the 2023 mandarin dragonet genome from NCBI (GenBank assembly accession: GCA_027744825.1) and used an unpublished highly contiguous cornetfish genome. Since these genomes were unannotated, we manually identified scpp genes. We searched for scpp genes using BLASTN with medaka and additional fish sequences as the query. We also searched for these genes with mVISTA plots across conserved gene synteny

regions (LAGAN alignment using translated anchoring) with medaka as the focal species. To identify scpp genes in additional species, we gathered cluster information from NCBI and ensembl. Additionally, we used mVISTA plots to further search for unannotated scpp genes.

Acknowledgements

We are indebted to Emily Rose for her collection of the Gulf pipefish specimens for the single cell libraries. We are grateful for everyone in the Cresko Lab for their ideas and participation in early morning pipefish collection trips. We are additionally thankful for Emily Beck for her assistance collecting pipefish. In particular, we are immensely appreciative of Mark Currey and Tiffany Thornton's valiant efforts to culture Gulf pipefish. Additionally, we are grateful for Balan Ramesh and Adam Jones for collaborating with the Cresko Lab to produce new Syngnathiformes genomes. We thank Tina Arredondo from the UO GC3F for her preparation of our single cell RNA sequencing libraries. This work was funded by the National Science Foundation Grant OPP-2015301 (to WAC, SB, and CMS), University of Oregon Research Excellence funds (WAC) and National Institute of Health Fellowship 5F31DE032559-02 (to HMH). Additionally, HMH was supported by the Genetics Training Program (NIH T32GM007413). Finally, VG received funding and mentorship from the Hui Undergraduate Research Scholars and MW was supported by the Knight Campus Undergraduate Scholars program at the University of Oregon.

Author Contributions

Hope Healey: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Review, and Editing, Visualization, Supervision, Funding acquisition, **Hayden Penn:** Investigation, Writing – Review and Editing, **Clayton Small:** Methodology, Investigation, Writing – Review and Editing, **Susan Bassham:** Methodology, Writing – Review and Editing, **Vithika Goyal:** Investigation, Writing – Review and Editing, **Micah Woods:** Investigation, Writing – Review and Editing, Visualization, **William Cresko:** Conceptualization, Resources, Writing – Review and Editing, Supervision, Funding acquisition

Data Availability statement

All raw sequencing data associated with this study will be published upon publication via NCBIThe updated Gulf pipefish annotation will be on the Dryad repository. Code used for the analysis will be available on github (Repository: ssc_emb_atlas_scripts). The portion of the Cornetfish genome used for the Scpp analysis will be published on NCBI upon publication.

Supplementary file captions

SupplementaryFile1: 230816_scRNAseq_cellranger_info.csv, a table containing quality metrics for the single cell libraries.

SupplementaryFile2: 230818_Cell_Cluster_Sizes. csv, a table containing the number of cells in each cell cluster and cluster identities.

SupplementaryFile3: Selected_Markers_Information.xlsx, a file containing additional information on the marker gene identified for every cluster.

SupplementaryFile4: In_Situ_Probe_and_Samples_Used_SuccessfulProbes.xlsx, a list of the in situ probes used in this study, the conditions used to prepare the probes, and the staging/sample information for the embryos.

SupplementaryFile5: WGCNA_networks_with_descriptions.xlsx, a file containing all the genetic networks, the genes inside of them, and additional information for the genetic networks highlighted in this paper.

SupplementaryFile6: WGCNA_module_number_to_color_translation_table, genetic networks were initially labeled with colors, we converted these labels to numeric annotations for simplicity using this conversion table. The table also contains the number of genes in each network.

SupplementaryFile7: 240228_ssc_WGCNA_mod_association_wClus_tstat.csv, this file contains the t-statistics derived for each module-cell cluster pair. The cell clusters are in the rows and the gene modules are in the columns.

SupplementaryFile8: 240228_ssc_WGCNA_associat_mods_w_clus_tstat_pvalues_corrected.csv, p-values for the t-statistics of the strength of association between gene modules and cell clusters. P-values are corrected for multiple testing hypotheses using fdr. The cell clusters are in the rows and the gene modules are in the columns.

SupplementaryFile9: 240227_WGCNA_connectivity_change.csv, the change in connectivity for gene modules when individual cell clusters are removed. The cell clusters are in the columns and the gene networks are in the rows.

SupplementaryFile10: 240228_WGCNA_EM1_EM2_Connectivity_Permutation_Pvalues_Corrected.csv, the p- values for the change in connectivity are found in this csv. The cell clusters are in the columns and the gene networks are in rows.

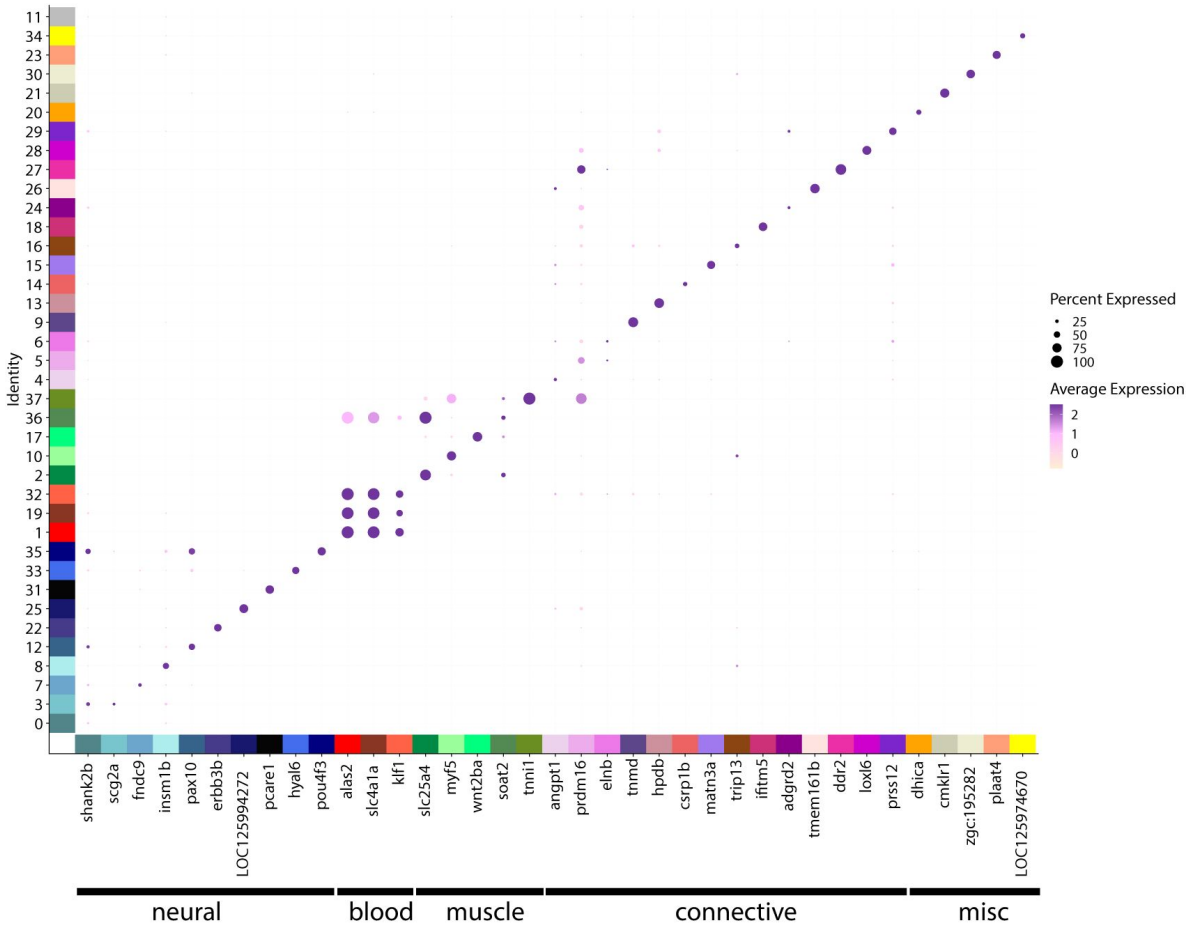

Figure S1

Marker genes define cell identity and are a resource for cell cluster exploration.

This dotplot shows the marker genes on the x-axis, and the cell cluster on the y-axis. The size of each dot indicates the percentage of cells that express the gene in each cluster. The darkness of each dot represents the average expression of the gene.

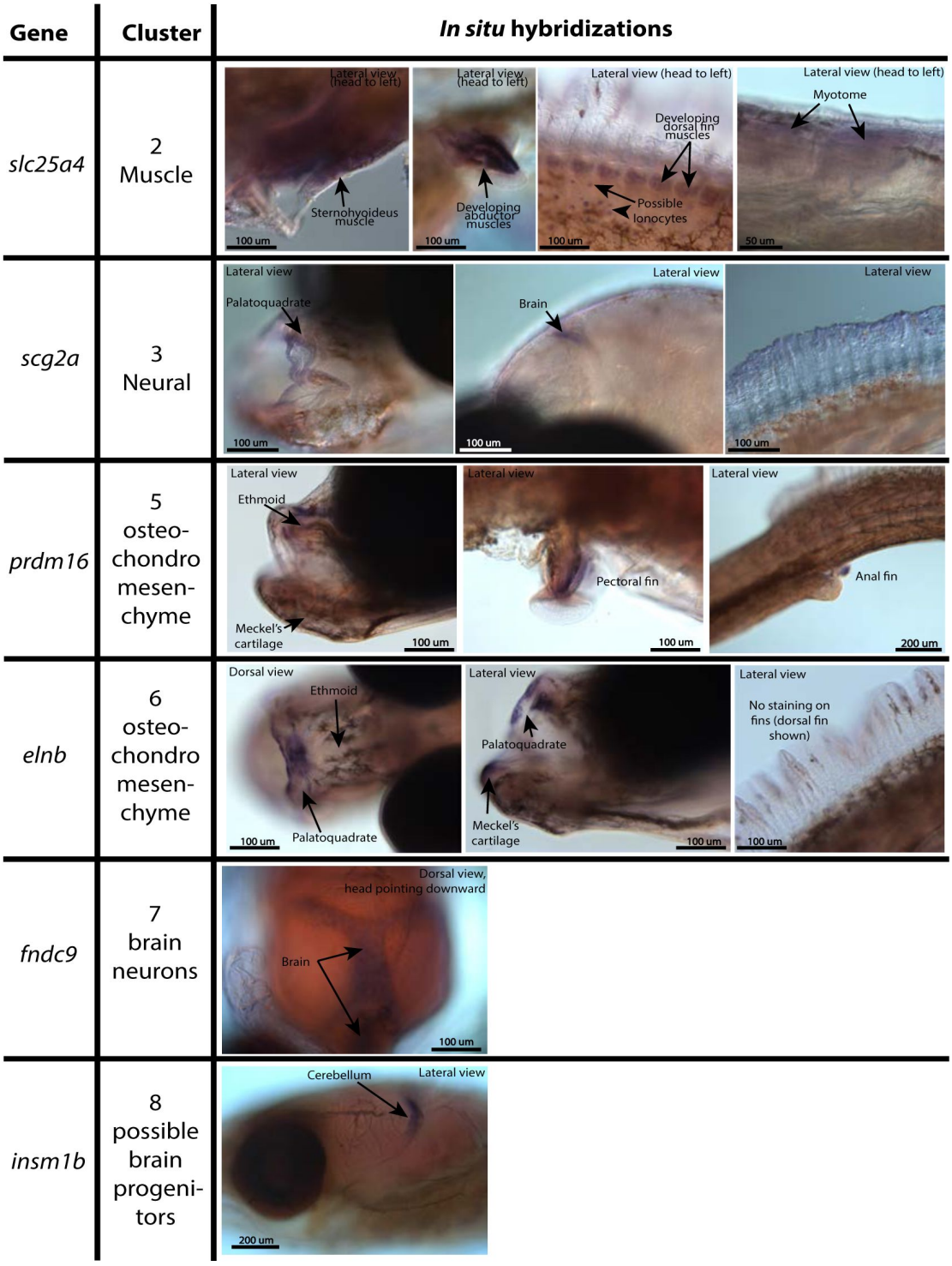

Figure S2

Marker genes examined through in situ hybridization reveal spatial patterning of cell types.

Informative marker gene *in situ* hybridization results, including *slc25a4*, *scg2a*, *prdm16*, *elnb*, *fndc9*, and *insm1b*. Embryos are staged to highlight staining of each gene. Relavant anatomical features are distinguished with arrows. One day post spawn Gulf pipefish larvae were used for *fndc9 in situs*, five day post spawn Gulf pipefish larvae were used for *insm1b in situs*, and wild caught bay pipefish at the onset of craniofacial elongation were used for all other probes.

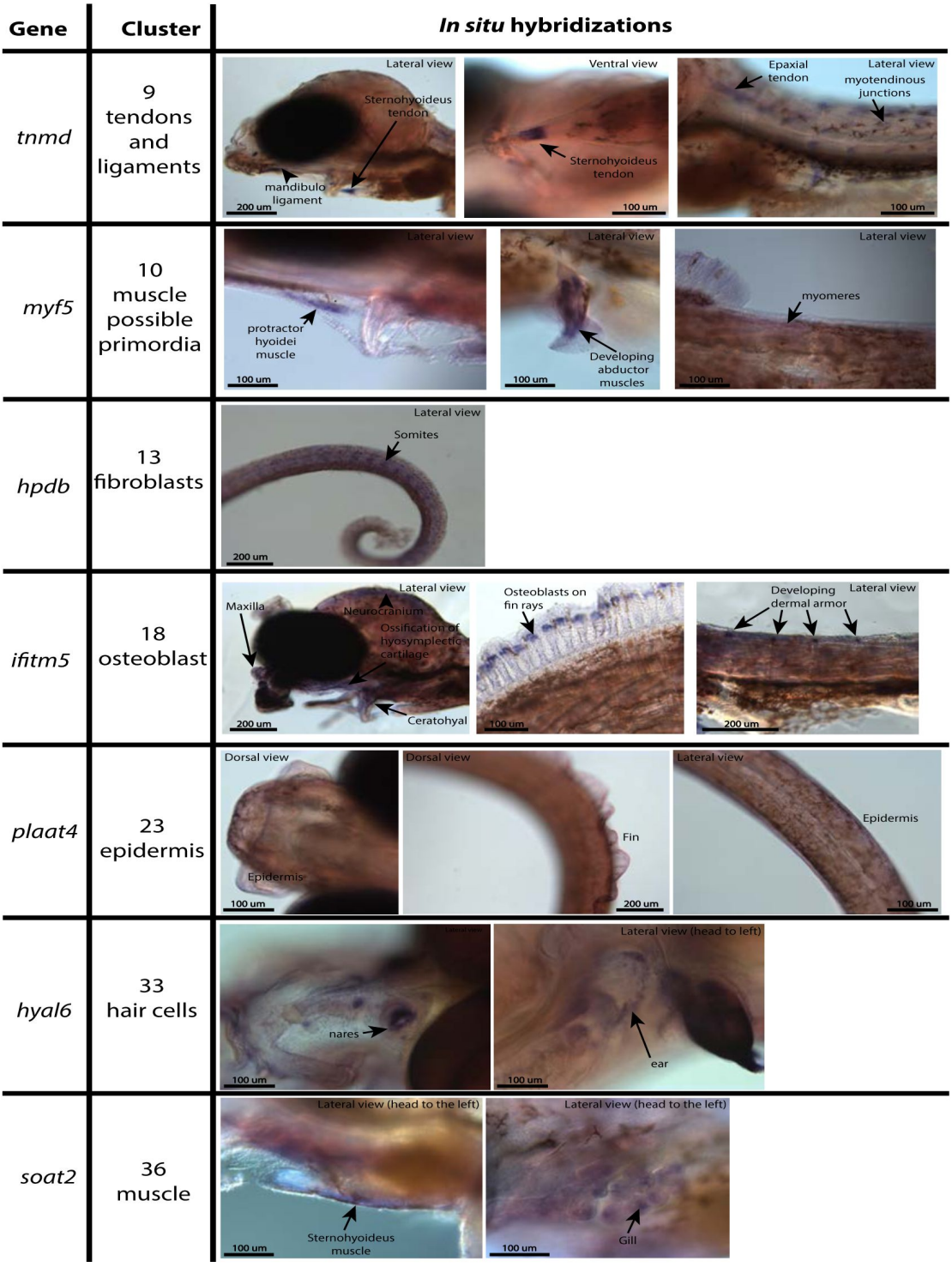

Figure S3

Marker genes examined through in situ hybridization reveal spatial patterning of cell types.

Informative marker gene *in situ* hybridization results, including *tnmd*, *myf5*, *hpbd*, *ifitm5*, *plaat4*, *hyal6*, and *soat2*. Embryos are staged to highlight staining of each gene. Relavant anatomical features are distinguished with arrows. All *in situs* were completed with wild caught bay pipefish: one day post spawn larvae were used for the *hyal6* probe, mid-somitogenesis embryos were used for the *hpbd*, and embryos at the onset of craniofacial elongation were used for all other probes.

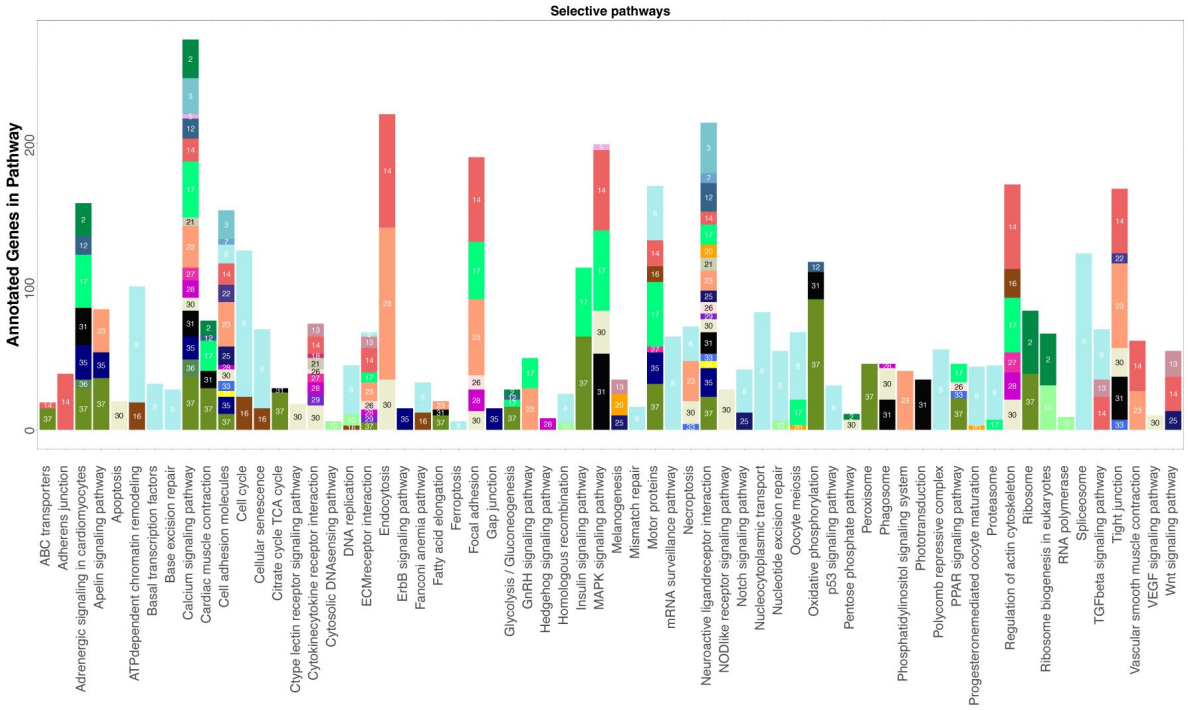

Figure S4

KEGG analysis reveals pathways signficiantly enriched in cell clusters.

KEGG pathways are listed on the x-axis. For each cluster, the number of genes that belong to the specific KEGG pathway is represented by the height of each bar. Cell clusters are labeled by both color and their designated number. For clarity, only cell clusters with more than three annotated genes in each pathway are shown and certain pathways involving infection and metabolism are removed.

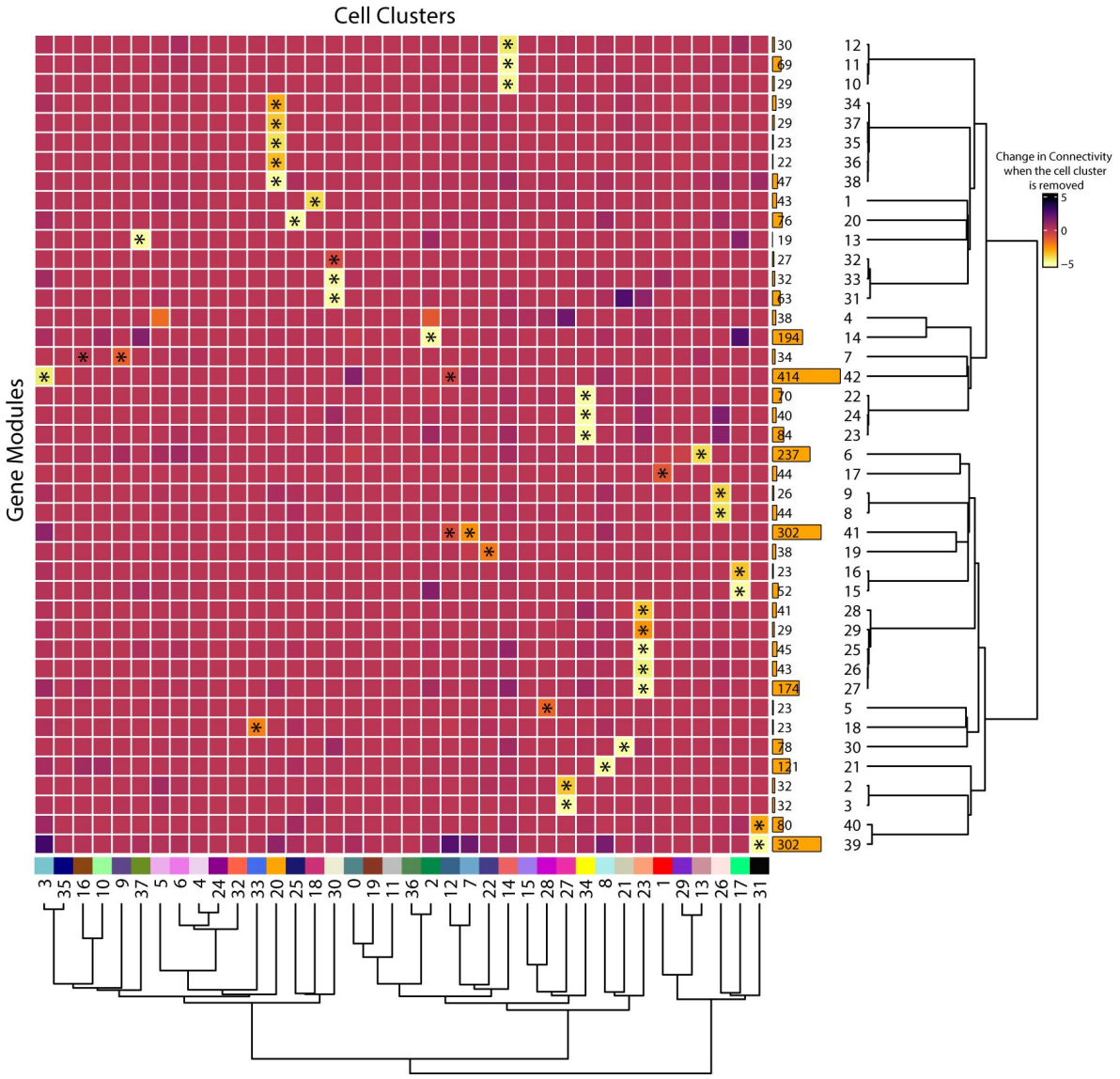

Figure S5

Cell clusters drive gene module connectivity.

The heatmap shows the change in gene network connectivity when individual cell clusters are removed. Cell clusters are in the columns and gene modules are in rows. The coloration scale is based on the change in connecitvity when a cell cluster is removed, with yellow indicating a large change in connectivity. The modules and clusters are clustered using the Pearson distance method. Asterisks indicate cluster-module pairs that are significant using a one way permutation test and FDR correction. The size of the gene modules are shown to the right.

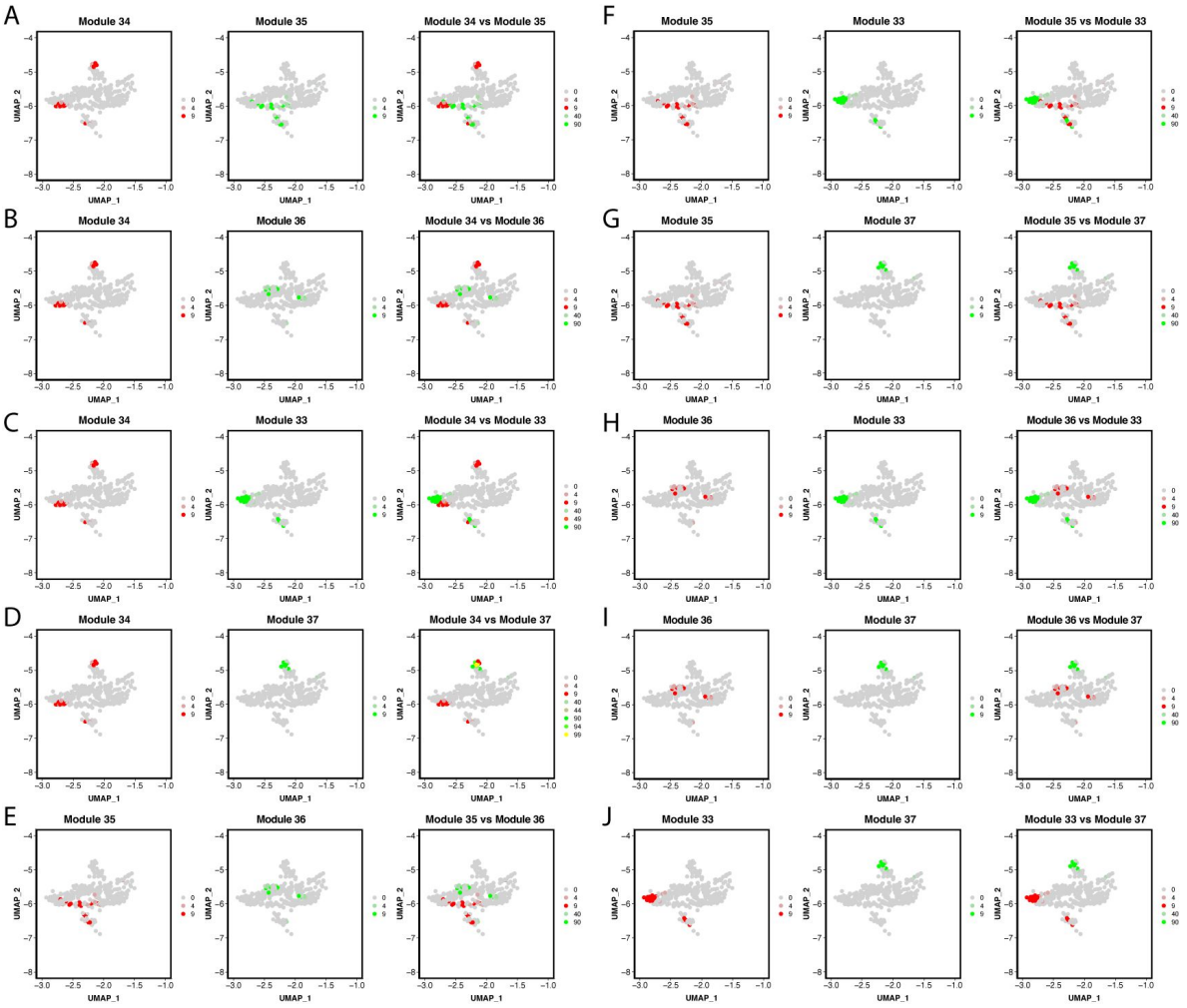

Figure S6

Gene module expression comparison with the Pigment Cluster reveals potential elusive cell types within the cluster.

Genes within the network are combined into one feature then plotted on a feature plot. In this case, UMAP dimensions one and two are the x and y-axis accordingly, dots represent each pigment cell, and the color of the dots represents the expression of the module. Cells that express gene modules are either colored red or green (when modules are plotted on separate plots). When plotted on the same plot, cells that express both gene modules are yellow. Panels A through J show each gene module comparison.

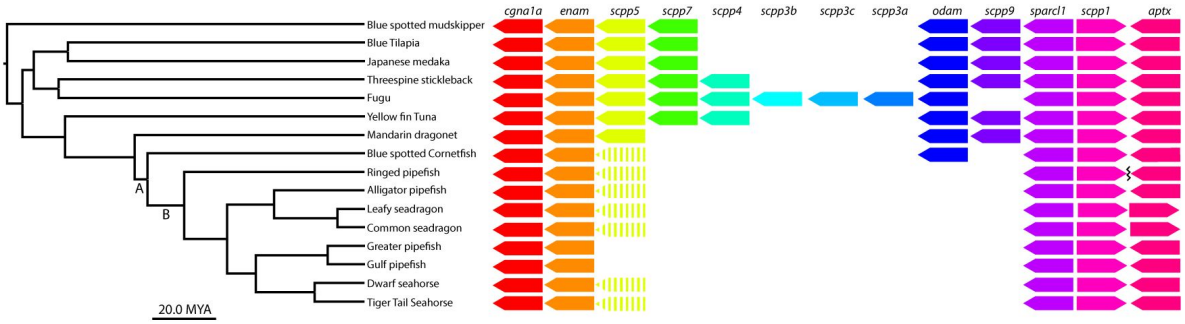

Figure S7

Scpp gene cluster analysis identifies most Scpp losses are not unique to syngnathids.

A) denotes the Syngnathoidei lineage which contains syngnathids and blue spotted cornetfish (among other fish) to highlight that this clade contains various tooth alterations. Syngnathids lack teeth completely and cornetfish have small teeth limited to the lower jaw. B) denotes solely the syngnathid clade, which completely lacks teeth. The phylogenetic tree was time calibrated using Hughes et al. 2018 and Stiller et al. 2022.

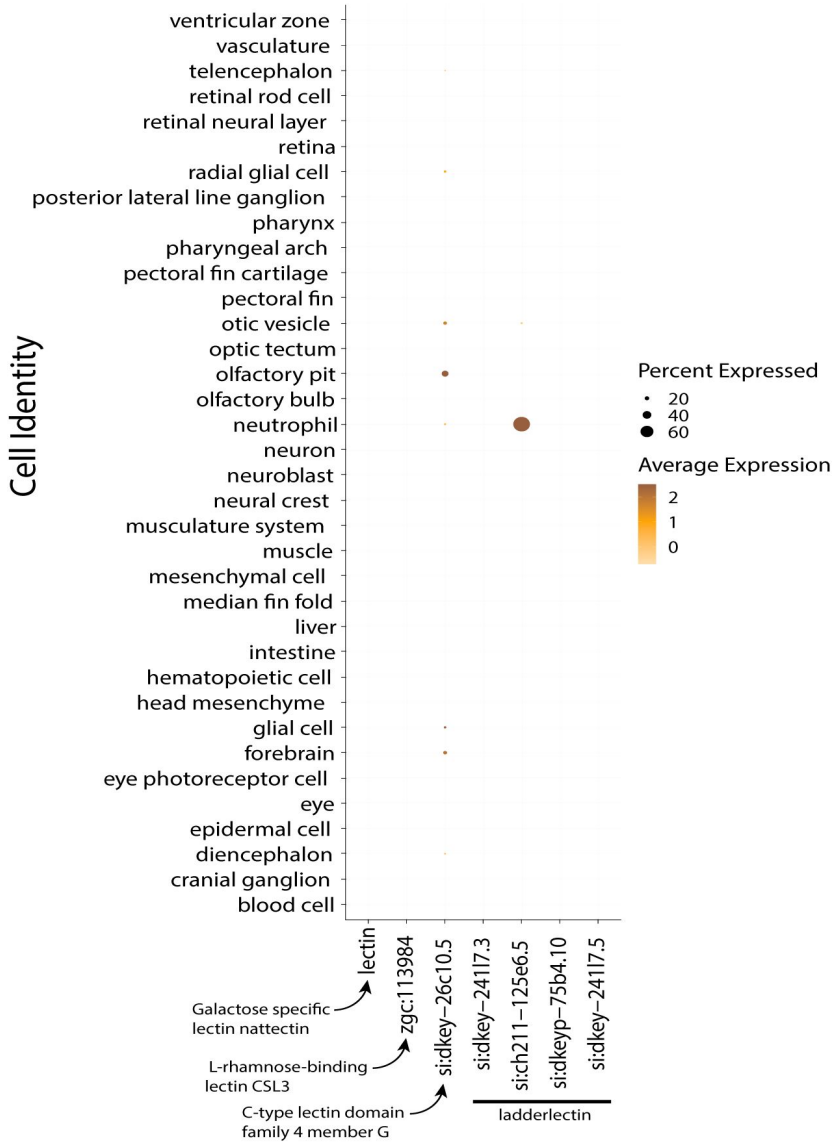

Figure S8

3dpf Zebrafish do not express lectin genes in epidermal cells.

This dotplot shows the marker genes on the x-axis and y-axis, organized by cell type. The size of the dot represents the percentage of cells that express the gene in each cluster. The darkness of the dot represents the average expression of the genes.

References

Ahi E. P., Kapralova K. H., Pálsson A., Maier V. H., Gudbrandsson J., Snorrason S. S., Jónsson Z. O., Franzdóttir S. R (2014) **Transcriptional dynamics of a conserved gene expression network associated with craniofacial divergence in Arctic charr** *EvoDevo* **5**:1–19 https://doi.org/10.1186/2041-9139-5-40/FIGURES/5

Ahi E. P., Lecaudey L. A., Ziegelbecker A., Steiner O., Glabonjat R., Goessler W., Hois V., Wagner C., Lass A., Sefc K. M (2020) **Comparative transcriptomics reveals candidate carotenoid color genes in an East African cichlid fish** *BMC Genomics* **2020** https://doi.org/10.1186/S12864-020-6473-8

Anderson A. P., Jones A. G (2019) **Choosy Gulf pipefish males ignore age but prefer active females with deeply keeled bodies** *Animal Behaviour* **155**:37–44 https://doi.org/10.1016/J.ANBEHAV.2019.05.018

Arendt D. *et al.* (2016) **The origin and evolution of cell types** *Nature Reviews Genetics* **17**:774–757 https://doi.org/10.1038/nrg.2016.127

Bergen D. J. M. *et al.* (2022) **Regenerating zebrafish scales express a subset of evolutionary conserved genes involved in human skeletal disease** *BMC Biology* **20** https://doi.org/10.1186/S12915-021-01209-8

Berglund A., Rosenqvist G., Svensson I (1986) **Reversed sex roles and parental energy investment in zygotes of two pipefish (Syngnathidae) species** *Marine Ecology Progress Series* **29**:209–215

Bjork B. C., Turbe-Doan A., Prysak M., Herron B. J., Beier D. R (2010) **Prdm16 is required for normal palatogenesis in mice** *Human Molecular Genetics* **19**:774–789 https://doi.org/10.1093/HMG/DDP543

Brown G. D., Willment J. A., Whitehead L (2018) **C-type lectins in immunity and homeostasis** *Nature Reviews Immunology* **18**:374–380 https://doi.org/10.1038/S41577-018-0004-8

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 https://doi.org/10.1038/nbt.4096

Cãestro C., Yokoi H., Postlethwait J. H. (2007) **Evolutionary developmental biology and genomics** *Nature Reviews Genetics* :932–942 https://doi.org/10.1038/nrg2226

Carcupino M (2002) **Functional significance of the male brood pouch in the reproductive strategies of pipefishes and seahorses: a morphological and ultrastructural comparative study on three anatomically different pouches** *Journal of Fish Biology* **61**:1465–1480 https://doi.org/10.1006/jfbi.2002.2160

Chari T. *et al.* (2021) **Whole-animal multiplexed single-cell RNA-seq reveals transcriptional shifts across Clytia medusa cell types** *Science Advances* **7** https://doi.org/10.1126/SCIADV.ABH1683/SUPPL_FILE/SCIADV.ABH1683_TABLES_

Chen J. W., Galloway J. L (2014) **The development of zebrafish tendon and ligament progenitors** *Development* **141**:2035–2045 https://doi.org/10.1242/DEV.104067

Chen Y. P., Zhang Y., Jiang T. X., Barlow A. J., St. Amand T. R., Hu Y., Heaney S., West Francis -, Chuong P., Maas C. M. (2000) **Conservation of early odontogenic signaling pathways in Aves** *Proceedings of the National Academy of Sciences* **97**:10044–10049 https://doi.org/10.1073/PNAS.160245097

Crump J. G., Maves L., Lawson N. D., Weinstein B. M., Kimmel C. B (2004) **An essential role for Fgfs in endodermal pouch formation influences later craniofacial skeletal patterning** *Development* **131**:5703–5716 https://doi.org/10.1242/dev.01444

Ding H. L., Clouthier D. E., Artinger K. B (2013) **Redundant roles of PRDM family members in zebrafish craniofacial development** *Developmental Dynamics* **242**:67–79 https://doi.org/10.1002/DVDY.23895

Docheva D., Hunziker E. B., Fässler R., Brandau O (2005) **Tenomodulin Is Necessary for Tenocyte Proliferation and Tendon Maturation** *Molecular and Cellular Biology* **25**:699–705 https://doi.org/10.1128/MCB.25.2.699-705.2005

Du J., Miller A. J., Widlund H. R., Horstmann M. A., Ramaswamy S., Fisher D. E (2003) **MLANA/MART1 and SILV/PMEL17/GP100 Are Transcriptionally Regulated by MITF in Melanocytes and Melanoma** *The American Journal of Pathology* **163**:333–343 https://doi.org/10.1016/S0002-9440(10)63657-7

Dudley J. S., Hannaford P., Dowland S. N., Lindsay L. A., Thompson M. B., Murphy C. R., Van Dyke J. U., Whittington C. M. (2021) **Structural changes to the brood pouch of male pregnant seahorses (Hippocampus abdominalis) facilitate exchange between father and embryos** *Placenta* **114**:115–123 https://doi.org/10.1016/J.PLACENTA.2021.09.002

Fabian P., Tseng K.-C., Thiruppathy M., Arata C., Chen H.-J., Smeeton J., Nelson N., Crump J. G (2022) **Lifelong single-cell profiling of cranial neural crest diversification in zebrafish** *Nature Communications* **2022** https://doi.org/10.1038/s41467-021-27594-w

Farnsworth D. R., Saunders L. M., Miller A. C (2020) **A single-cell transcriptome atlas for zebrafish development** *Developmental Biology* **459**:100–108 https://doi.org/10.1016/j.ydbio.2019.11.008

Farrell J. A., Wang Y., Riesenfeld S. J., Shekhar K., Regev A., Schier A. F (2018) **Single- cell reconstruction of developmental trajectories during zebrafish embryogenesis** *Science* **360** https://doi.org/10.1126/science.aar3131

Farrell J. A., Wang Y., Riesenfeld S. J., Shekhar K., Regev A., Schier A. F (2018) **Single- cell reconstruction of developmental trajectories during zebrafish embryogenesis** *Science* **360** https://doi.org/10.1126/science.aar3131

Feregrino C., Sacher F., Parnas O., Tschopp P (2019) **A single-cell transcriptomic atlas of the developing chicken limb** *BMC Genomics* **20**:1–15 https://doi.org/10.1186/s12864-019-5802-2

Fuess L. E., Bolnick D. I (2023) **GBE Single-Cell RNA Sequencing Reveals Microevolution of the Stickleback Immune System** *Genome Biology and Evolution* **15** https://doi.org/10.1093/gbe/evad053

Fuiten A. M., Cresko W. A (2021) **Evolutionary divergence of a Hoxa2b hindbrain enhancer in syngnathids mimics results of functional assays** *Development Genes and Evolution* **231**:57–71 https://doi.org/10.1007/S00427-021-00676-X/METRICS

Gibert Y., Samarut E., Ellis M. K., Jackman W. R., Laudet V (2019) **The first formed tooth serves as a signalling centre to induce the formation of the dental row in zebrafish** *Proceedings of the Royal Society B: Biological Sciences* **286** https://doi.org/10.1098/rspb.2019.0401

Goldsmith M. I., Fisher S., Waterman R., Johnson S. L (2003) **Saltatory control of isometric growth in the zebrafish caudal fin is disrupted in long fin and rapunzel mutants** *Developmental Biology* **259**:303–317 https://doi.org/10.1016/S0012-1606(03)00186-6

Hafemeister C., Satija R (2019) **Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression** *Genome Biology* **20** https://doi.org/10.1186/s13059-019-1874-1

Hain D. *et al.* (2022) **Molecular diversity and evolution of neuron types in the amniote brain** *Science* **377**

Harada A., Shiota R., Okubo R., Yorifuji M., Sogabe A., Motomura H., Hiroi J., Yasumasu S., Kawaguchi M (2022) **Brood pouch evolution in pipefish and seahorse based on histological observation** *Placenta* **120**:88–96 https://doi.org/10.1016/J.PLACENTA.2022.02.014

Harlin-Cognato A., Hoffman E. A., Jones A. G (2006) **Gene cooption without duplication during the evolution of a male-pregnancy gene in pipefish** *Proceedings of the National Academy of Sciences of the United States of America* **103**:19407–19412 https://doi.org/10.1073/pnas.0603000103

Healey H. M., Bassham S., Cresko W. A (2022) **Single-cell Iso-Sequencing enables rapid genome annotation for scRNAseq analysis** *Genetics* **220** https://doi.org/10.1093/GENETICS/IYAC017

Hong X., Wang Y., Wang K., Wei C., Li W., Yu L., Xu H., Zhu J., Zhu X., Liu X (2023) **Single-Cell Atlas Reveals the Hemocyte Subpopulations and Stress Responses in Asian Giant Softshell Turtle during Hibernation** *Biology* **12** https://doi.org/10.3390/BIOLOGY12070994/S1

Howe D. G. *et al.* (2013) **ZFIN, the Zebrafish Model Organism Database: Increased support for mutants and transgenics** *Nucleic Acids Research* **41**:D854–D860 https://doi.org/10.1093/NAR/GKS938

Hughes L. C. *et al.* (2018) **Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data** *Proceedings of the National Academy of Sciences of the United States of America* **115**:6249–6254

Jain I., Stroka C., Yan J., Huang W. M., Iovine M. K (2007) **Bone growth in zebrafish fins occurs via multiple pulses of cell proliferation** *Developmental Dynamics* **236**:2668–2674 https://doi.org/10.1002/DVDY.21270

Johnson S. L., Nguyen A. T. N., Lister J. A (2011) **mitfa is required at multiple stages of melanocyte differentiation but not to establish the melanocyte stem cell** *Developmental Biology* **350**:405–413 https://doi.org/10.1016/J.YDBIO.2010.12.004

Jungerson H. F. E. (1910) **Ichthyotomical Contributions: The structure of the Aulostomidae, Syngnathidae, and Solenostomidae**

Kaneda-Nakashima K. *et al.* (2022) **Role of Mel1/Prdm16 in bone differentiation and morphology** *Experimental Cell Research* **410** https://doi.org/10.1016/J.YEXCR.2021.112969

Kawasaki K (2009) **The SCPP gene repertoire in bony vertebrates and graded differences in mineralized tissues** *Development Genes and Evolution* **219**:147–157 https://doi.org/10.1007/S00427-009-0276-X/FIGURES/5

Kimura T., Takehana Y., Naruse K (2017) **Pnp4a Is the causal gene of the medaka iridophore Mutant guanineless** *G* **3** https://doi.org/10.1534/G3.117.040675/-/DC1

Koiwai K., Kondo H., Hirono I (2023) **scRNA-seq Analysis of Hemocytes of Penaeid Shrimp Under Virus Infection** *Marine Biotechnology* **25**:488–502 https://doi.org/10.1007/S10126-023-10221-8/METRICS

Kuo R. I., Tseng E., Eory L., Paton I. R., Archibald A. L., Burt D. W (2017) **Normalized long read RNA sequencing in chicken reveals transcriptome complexity similar to human** *BMC Genomics* **18**:1–19 https://doi.org/10.1186/S12864-017-3691-9

Kvarnemo C., Mobley K. B., Partridge C., Jones A. G., Ahnesjö I (2011) **Evidence of paternal nutrient provisioning to embryos in broad-nosed pipefish Syngnathus typhle** *Journal of Fish Biology* **78**:1725–1737 https://doi.org/10.1111/j.1095-8649.2011.02989.x

Lamason R. L. *et al.* (2005) **Genetics: SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans** *Science* **310**:1782–1786 https://doi.org/10.1126/SCIENCE.1116238/SUPPL_FILE/LAMASON.SOM.PDF

Lange M. *et al.* (2023) **Zebrahub – Multimodal Zebrafish Developmental Atlas Reveals the State-Transition Dynamics of Late-Vertebrate Pluripotent Axial Progenitors** *BioRxiv* **2023**:3–6 https://doi.org/10.1101/2023.03.06.531398

Langfelder P., Horvath S (2008) **WGCNA: An R package for weighted correlation network analysis** *BMC Bioinformatics* **9**:1–13

Leerberg D. M., Hopton R. E., Draper B. W (2019) **Fibroblast growth factor receptors function redundantly during zebrafish embryonic development** *Genetics* **212**:1301–1319 https://doi.org/10.1534/genetics.119.302345

Leysen H., Roos G., Adriaens D (2011) **Morphological variation in head shape of pipefishes and seahorses in relation to snout length and developmental growth** *Journal of Morphology* **272**:1259–1270 https://doi.org/10.1002/jmor.10982

Lin Q. *et al.* (2016) **The seahorse genome and the evolution of its specialized morphology** *Nature* **540**:395–399 https://doi.org/10.1038/nature20595

Lin Q. *et al.* (2016) **The seahorse genome and the evolution of its specialized morphology** *Nature* **540**:395–399 https://doi.org/10.1038/nature20595

Liu Z. *et al.* (2016) **The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts** *Nature Communications* **7** https://doi.org/10.1038/NCOMMS11757

Manning C. G., Foster S. J., Vincent A. C. J (2019) **A review of the diets and feeding behaviours of a family of biologically diverse marine fishes (Family Syngnathidae)** *Reviews in Fish Biology and Fisheries* **2019** https://doi.org/10.1007/S11160-019-09549-Z

McCarthy N., Sidik A., Bertrand J. Y., Eberhart J. K (2016) **An Fgf-Shh signaling hierarchy regulates early specification of the zebrafish skull** *Developmental Biology* **415**:261–277 https://doi.org/10.1016/j.ydbio.2016.04.005

Melamed P., Xue Y., Poon J. F. D., Wu Q., Xie H., Yeo J., Foo T. W. J., Chua H. K (2005) **The male seahorse synthesizes and secretes a novel C-type lectin into the brood pouch during early pregnancy** *The FEBS Journal* **272**:1221–1235 https://doi.org/10.1111/J.1742-4658.2005.04556.X

Miao M., Bruce A. E. E., Bhanji T., Davis E. C., Keeley F. W (2007) **Differential expression of two tropoelastin genes in zebrafish** *Matrix Biology* **26**:115–124 https://doi.org/10.1016/j.matbio.2006.09.011

Miyake A., Itoh N (2013) **Fgf22 regulated by Fgf3/Fgf8 signaling is required for zebrafish midbrain development** *Biology Open* **2**:515–524 https://doi.org/10.1242/BIO.20134226

Mori S., Nakamura T (2022) **Redeployment of odontode gene regulatory network underlies dermal denticle formation and evolution in suckermouth armored catfish** *Scientific Reports* **2022** https://doi.org/10.1038/s41598-022-10222-y

Moriyama Y., Ito F., Takeda H., Yano T., Okabe M., Kuraku S., Keeley F. W., Koshiba-Takeuchi K (2016) **Evolution of the fish heart by sub/neofunctionalization of an elastin gene** *Nature Communications* **2016** https://doi.org/10.1038/ncomms10397

Morrison J. A. *et al.* (2017) **Single-cell transcriptome analysis of avian neural crest migration reveals signatures of invasion and molecular transitions** *ELife* **6** https://doi.org/10.7554/eLife.28415

Parker J., Guslund N. C., Jentoft S., Roth O (2022) **Characterization of Pipefish Immune Cell Populations Through Single-Cell Transcriptomics** *Frontiers in Immunology* **13** https://doi.org/10.3389/FIMMU.2022.820152/BIBTEX

Partridge C., Shardo J., Boettcher A (2007) **Osmoregulatory role of the brood pouch in the euryhaline Gulf pipefish** *Syngnathus scovelli. Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology* https://doi.org/10.1016/j.cbpa.2007.02.007

Potts H. G., Lemieux M. E., Rice E. S., Warren W., Choudhury R. P., Mommersteeg M. T. M (2022) **Discordant Genome Assemblies Drastically Alter the Interpretation of Single-Cell RNA Sequencing Data Which Can Be Mitigated by a Novel Integration Method** *Cells* **11** https://doi.org/10.3390/CELLS11040608/S1

Qu M. *et al.* (2021) **Seadragon genome analysis provides insights into its phenotype and sex determination locus** *Science Advances* **7** https://doi.org/10.1126/SCIADV.ABG5196

Ramesh B., Small C. M., Healey H., Johnson B., Barker E., Currey M., Bassham S., Myers M., Cresko W. A., Jones A. G (2023) **Improvements to the Gulf pipefish Syngnathus scovelli genome** *GigaByte* **2023** https://doi.org/10.46471/GIGABYTE.76

Ripley J.L., Williams P. S., Foran C. M (2010) **Morphological and quantitative changes in paternal brood-pouch vasculature during embryonic development in two syngnathus pipefishes** *Journal of Fish Biology* **77**:67–79

Ripley Jennifer L., Foran C. M (2006) **Differential parental nutrient allocation in two congeneric pipefish species (Syngnathidae: Syngnathus spp** *Journal of Experimental Biology* **209**:1112–1121 https://doi.org/10.1242/jeb.02119

Ripley Jennifer L., Foran C. M (2006) **Differential parental nutrient allocation in two congeneric pipefish species (Syngnathidae: Syngnathus spp** *Journal of Experimental Biology* **209**:1112–1121 https://doi.org/10.1242/JEB.02119

Rose E., Johnson B. D., Armitage V., Jones A. G (2023) **The response of the brood pouch transcriptome to synthetic estrogen exposure in the Gulf pipefish (Syngnathus scovelli)** *Frontiers in Marine Science* **10**:1–12 https://doi.org/10.3389/fmars.2023.1138597

Roth O., Klein V., Beemelmanns A., Scharsack J. P., Reusch T. B. H (2012) **Male pregnancy and biparental immune priming** *American Naturalist* **180**:802–814 https://doi.org/10.1086/668081

Roth O. *et al.* (2020) **Evolution of male pregnancy associated with remodeling of canonical vertebrate immunity in seahorses and pipefishes** *Proceedings of the National Academy of Sciences* **201916251** https://doi.org/10.1073/pnas.1916251117

Roth O. *et al.* (2020) **Evolution of male pregnancy associated with remodeling of canonical vertebrate immunity in seahorses and pipefishes** *Proceedings of the National Academy of Sciences of the United States of America* **117**:9431–9439 https://doi.org/10.1073/PNAS.1916251117/-/DCSUPPLEMENTAL

Royan M. R., Siddique K., Csucs G., Puchades M. A., Nourizadeh-Lillabadi R., Bjaalie J. G., Henkel C. V., Weltzien F. A., Fontaine R (2021) **3D Atlas of the Pituitary Gland of the Model Fish Medaka (Oryzias latipes)** *Frontiers in Endocrinology* **12** https://doi.org/10.3389/FENDO.2021.719843/BIBTEX

Salamanca-Díaz D. A., Schulreich S. M., Cole A. G., Wanninger A (2022) **Single-Cell RNA Sequencing Atlas From a Bivalve Larva Enhances Classical Cell Lineage Studies** *Frontiers in Ecology and Evolution* **9** https://doi.org/10.3389/FEVO.2021.783984/BIBTEX

Schilling T. F., Kimmel C. B (1997) **Musculoskeletal patterning in the pharyngeal segments of the zebrafish embryo** *Development* **124**:2945–2960 https://doi.org/10.1242/DEV.124.15.2945

Schneider R. F., Gunter H. M., Salewski I., Woltering J. M., Meyer A (2023) **Growth dynamics and molecular bases of evolutionary novel jaw extensions in halfbeaks and needlefishes (Beloniformes)** *Molecular Ecology* **32**:5798–5811 https://doi.org/10.1111/MEC.17143

Schneider R. F., Woltering J. M., Adriaens D., Roth O (2023) **A comparative analysis of the ontogeny of syngnathids (pipefishes and seahorses) reveals how heterochrony contributed to their diversification** *Developmental Dynamics* **252**:553–588 https://doi.org/10.1002/DVDY.551

Shema E., Bernstein B. E., Buenrostro J. D (2019) **Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution** *Nature Genetics* **51**:19–25 https://doi.org/10.1038/s41588-018-0290-x

Shull Lomeli C., Lencer E. S., Kim H. M., Goyama S., Kurokawa M., Costello J. C., Jones K., Artinger K. B (2022) **PRDM paralogs antagonistically balance Wnt/β-catenin activity during craniofacial chondrocyte differentiation** *Development (Cambridge, England)* **149** https://doi.org/10.1242/DEV.200082/VIDEO-2

Shull Lomeli Carpio, Sen R., Menzel J., Goyama S., Kurokawa M., Artinger K. B (2020) **The conserved and divergent roles of Prdm3 and Prdm16 in zebrafish and mouse craniofacial development** *Developmental Biology* **461**:132–144 https://doi.org/10.1016/J.YDBIO.2020.02.006

Sire J. Y., Delgado S. C., Girondot M (2008) **Hen's teeth with enamel cap: From dream to impossibility** *BMC Evolutionary Biology* **8**:1–11 https://doi.org/10.1186/1471-2148-8-246

Small C. M., Bassham S., Catchen J., Amores A., Fuiten A. M., Brown R. S., Jones A. G., Cresko W. A (2016) **The genome of the Gulf pipefish enables understanding of evolutionary innovations** *Genome Biology* **17** https://doi.org/10.1186/s13059-016-1126-6

Small Clayton M., Harlin-Cognato A. D., Jones A. G (2013) **Functional similarity and molecular divergence of a novel reproductive transcriptome in two male-pregnant Syngnathus pipefish species** *Ecology and Evolution* **3**:4092–4108 https://doi.org/10.1002/ECE3.763

Small Clayton M., Healey H. M., Currey M. C., Beck E. A., Catchen J., Lin A. S. P., Cresko W. A., Bassham S (2022) **Leafy and weedy seadragon genomes connect genic and repetitive DNA features to the extravagant biology of syngnathid fishes** *Proceedings of the National Academy of Sciences of the United States of America* **119** https://doi.org/10.1073/PNAS.2119602119/SUPPL_FILE/PNAS.2119602119.SD05.X

Soldatov R. *et al.* (2019) **Spatiotemporal structure of cell fate decisions in murine neural crest** *Science* **364** https://doi.org/10.1126/science.aas9536

Sommer S., Whittington C. M., Wilson A. B (2012) **Standardised classification of pre- release development in male-brooding pipefish, seahorses, and seadragons (Family Syngnathidae)** *BMC Developmental Biology* **12**:12–15 https://doi.org/10.1186/1471-213X-12-39

Songco-Casey J. O., Coffing G. C., Piscopo D. M., Pungor J. R., Kern A. D., Miller A. C., Niell C. M (2022) **Cell types and molecular architecture of the Octopus bimaculoides visual system** *Current Biology : CB* **32** https://doi.org/10.1016/J.CUB.2022.10.015

Steger J., Cole A. G., Denner A., Lebedeva T., Genikhovich G., Ries A., Reischl R., Taudes E., Lassnig M., Technau U (2022) **Single-cell transcriptomics identifies conserved regulators of neuroglandular lineages** *Cell Reports* **40** https://doi.org/10.1016/J.CELREP.2022.111370

Stiller J., Short G., Hamilton H., Saarman N., Longo S., Wainwright P., Rouse G. W., Simison W. B (2022) **Phylogenomic analysis of Syngnathidae reveals novel relationships, origins of endemic diversity and variable diversification rates** *BMC Biology* **2022** https://doi.org/10.1186/S12915-022-01271-W

Stock D. W., Jackman W. R., Trapani J (2006) **Developmental genetic mechanisms of evolutionary tooth loss in cypriniform fishes** *Development* **133**:3127–3137 https://doi.org/10.1242/DEV.02459

Stölting K. N., Wilson A. B (2007) **Male pregnancy in seahorses and pipefish: Beyond the mammalian model** *BioEssays* **29**:884–896 https://doi.org/10.1002/bies.20626

Swartz M. E., Sheehan-Rooney K., Dixon M. J., Eberhart J. K (2011) **Examination of a palatogenic gene program in zebrafish** *Developmental Dynamics* **240**:2204–2220 https://doi.org/10.1002/DVDY.22713

Szabo-Rogers H. L., Geetha-Loganathan P., Nimmagadda S., Fu K. K., Richman J. M (2008) **FGF signals from the nasal pit are necessary for normal facial morphogenesis** *Developmental Biology* **318**:289–302 https://doi.org/10.1016/j.ydbio.2008.03.027

Tardaguila M. *et al.* (2018) **SQANTI: Extensive characterization of long-read transcript sequences for quality control in full-length transcriptome identification and quantification** *Genome Research* **28** https://doi.org/10.1101/GR.222976.117

Teichmann S. A., Babu M. M (2004) **Gene regulatory network growth by duplication** *Nature Genetics* **36**:492–496 https://doi.org/10.1038/ng1340

Tengfei D. *et al.* (2021) **Multiple Fetal Nutritional Patterns Before Parturition in Viviparous Fish Sebastes schlegelii (Hilgendorf, 1880)** *Frontiers in Marine Science* **7** https://doi.org/10.3389/FMARS.2020.571946

Thisse B., Thisse C. (2004) **Fast Release Clones: A High Throughput Expression Analysis** *ZFIN Direct Data Submission*

Thisse C., Thisse B (2007) **High-resolution in situ hybridization to whole-mount zebrafish embryos** *Nature Protocols* **2008** https://doi.org/10.1038/nprot.2007.514

Tokita M., Chaeychomsri W., Siruntawineti J (2013) **Developmental basis of toothlessness in turtles: Insight into convergent evolution of vertebrate morphology** *Evolution* **67**:260–273 https://doi.org/10.1111/j.1558-5646.2012.01752.x

Ton, Mai Linh N, Guibentif C., Göttgens B. (2020) **Single cell genomics and developmental biology: moving beyond the generation of cell type catalogues** *In Current Opinion in Genetics and Development* **64**:66–71 https://doi.org/10.1016/j.gde.2020.05.033

Ton *et al.* (2023) **An atlas of rabbit development as a model for single-cell comparative genomics** *Nature Cell Biology* **2023** https://doi.org/10.1038/s41556-023-01174-0

Tseng E (2021) **Tseng, E. (2021). cDNA cupcake (v27.0.0).**

Tucker A. S., Headon D. J., Schneider P., Ferguson B. M., Overbeek P., Tschopp J., Sharpe P. T (2000) **Edar/Eda interactions regulate enamel knot formation in tooth morphogenesis** *Development* **127**:4691–4700 https://doi.org/10.1242/DEV.127.21.4691

Tucker A., Sharpe P (2004) **The cutting-edge of mammalian development; how the embryo makes teeth** *Nature Reviews Genetics* **2004** https://doi.org/10.1038/nrg1380

Umemori H., Linhoff M. W., Ornitz D. M., Sanes J. R (2004) **FGF22 and its close relatives are presynaptic organizing molecules in the mammalian brain** *Cell* **118**:257–270 https://doi.org/10.1016/j.cell.2004.06.025

Van Wassenbergh S., Roos G., Genbrugge A., Leysen H., Aerts P., Adriaens D., Herrel A. (2009) **Suction is kid's play: extremely fast suction in newborn seahorses** *Biology Letters* **5**:200–203 https://doi.org/10.1098/rsbl.2008.0765

Van Wassenbergh S., Strother J. A., Flammang B. E., Ferry-Graham L. A., Aerts P. (2007) **Extremely fast prey capture in pipefish is powered by elastic recoil** *Journal of The Royal Society Interface* **5**:285–296 https://doi.org/10.1098/RSIF.2007.1124

Verreijdt L., Debiais-Thibaud M., Borday-Birraux V., Van Der Heyden C., Sire J. Y., Huysseune A. (2006) **Expression of the dlx gene family during formation of the cranial bones in the zebrafish (Danio rerio): Differential involvement in the visceral skeleton and braincase** *Developmental Dynamics* **235**:1371–1389 https://doi.org/10.1002/DVDY.20734

Vonk A. C. *et al.* (2023) **Single-cell analysis of lizard blastema fibroblasts reveals phagocyte-dependent activation of Hedgehog-responsive chondrogenesis** *Nature Communications* **2023** https://doi.org/10.1038/s41467-023-40206-z

Wagner A (2011) **The molecular origins of evolutionary innovations** *Trends in Genetics* **27**:397–410 https://doi.org/10.1016/J.TIG.2011.06.002

Wagner D. E., Weinreb C., Collins Z. M., Briggs J. A., Megason S. G., Klein A. M (2018) **Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo** *Science* **360**:981–987 https://doi.org/10.1126/science.aar4362

Wagner G. P., Lynch V. J. (2010) **Evolutionary novelties** *Current Biology* **20**:R48–R52 https://doi.org/10.1016/j.cub.2009.11.010

Walshe J., Mason I (2003) **Fgf signalling is required for formation of cartilage in the head** *Developmental Biology* **264**:522–536 https://doi.org/10.1016/j.ydbio.2003.08.010

Wang W. *et al.* (2024) **The TET-Sall4-BMP regulatory axis controls craniofacial cartilage development** *Cell Reports* **43** https://doi.org/10.1016/J.CELREP.2024.113873

Wetzel J. T., Wourms J. P (2004) **Embryogenesis in the Dwarf Seahorse, Hippocampus zosterae (Syngnathidae)** *Gulf and Caribbean Research* **16**:27–35 https://doi.org/10.18785/gcr.1601.04

Whittington C. M., Friesen C. R (2020) **The evolution and physiology of male pregnancy in syngnathid fishes** *Biological Reviews* **95**:1252–1272 https://doi.org/10.1111/brv.12607

Whittington C. M., Griffith O. W., Qi W., Thompson M. B., Wilson A. B (2015) **Seahorse Brood Pouch Transcriptome Reveals Common Genes Associated with Vertebrate Pregnancy** *Molecular Biology and Evolution* **32** https://doi.org/10.1093/molbev/msv177

Whittington C. M., Griffith O. W., Qi W., Thompson M. B., Wilson A. B (2015) **Seahorse Brood Pouch Transcriptome Reveals Common Genes Associated with Vertebrate Pregnancy** *Molecular Biology and Evolution* **32**:3114–3131 https://doi.org/10.1093/MOLBEV/MSV177

Williams R. M., Candido-Ferreira I., Repapi E., Gavriouchkina D., Senanayake U., Ling I. T. C., Telenius J., Taylor S., Hughes J., Sauka-Spengler T (2019) **Reconstruction of the Global Neural Crest Gene Regulatory Network In Vivo** *Developmental Cell* **51**:255–276 https://doi.org/10.1016/j.devcel.2019.10.003

Wolf M. *et al.* (2024) **Near chromosome-level and highly repetitive genome assembly of the snake pipefish Entelurus aequoreus (Syngnathiformes: Syngnathidae)** *GigaByte* **2024**:1–13 https://doi.org/10.46471/GIGABYTE.105

Woronowicz K. C., Schneider R. A (2019) **Molecular and cellular mechanisms underlying the evolution of form and function in the amniote jaw** *In EvoDevo* **10**:1–21 https://doi.org/10.1186/s13227-019-0131-8

Wourms J. P (1981) **Viviparity: The Maternal-Fetal Relationship in Fishes** *American Zoologist* **21**:473–515

Woych J. *et al.* (2022) **Cell-type profiling in salamanders identifies innovations in vertebrate forebrain evolution** *Science* **377** https://doi.org/10.1126/SCIENCE.ABP9186/SUPPL_FILE/SCIENCE.ABP9186_MDAR_

Young M. D., Behjati S (2020) **SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data** *GigaScience* **9**:1–10 https://doi.org/10.1093/GIGASCIENCE/GIAA151

Zhang Y., Ravi V., Qin G., Dai H., Zhang H., Han F., Wang X., Liu Y (2020) **Comparative genomics reveal shared genomic changes in syngnathid fishes and signatures of genetic convergence with placental mammals** *National Science Review* **7**:964–977 https://doi.org/10.1093/nsr/nwaa002

Zhou J., Gao Y., Lan Y., Jia S., Jiang R (2013) **Pax9 regulates a molecular network involving Bmp4, Fgf10, Shh signaling and the Osr2 transcription factor to control palate morphogenesis** *Development* **140**:4709–4718 https://doi.org/10.1242/DEV.099028

Editors

Reviewing Editor
Frank Chan
University of Groningen, Groningen, Netherlands

Senior Editor
Claude Desplan
New York University, New York, United States of America

Reviewer #1 (Public Review):

Syngnathid fishes (seahorses, pipefishes, and seadragons) present very particular and elaborated features among teleosts and a major challenge is to understand the cellular and molecular mechanisms that permitted such innovations and adaptations. The study provides a valuable new resource to investigate the morphogenetic basis of four main traits characterizing syngnathids, including the elongated snout, toothlessness, dermal armor, and male pregnancy. More particularly, the authors have focused on a late stage of pipefish organogenesis to perform single-cell RNA-sequencing (scRNA-seq) completed by in situ hybridization analyses to identify molecular pathways implicated in the formation of the different specific traits.

The first set of data explores the scRNA-seq atlas composed of 35,785 cells from two samples of gulf pipefish embryos that authors have been able to classify into major cell types characterizing vertebrate organogenesis, including epithelial, connective, neural, and muscle progenitors. To affirm identities and discover potential properties of clusters, authors primarily use KEGG analysis that reveals enriched genetic pathways in each cell types. While the analysis is informative and could be useful for the community, some interpretations appear superficial and data must be completed to confirm identities and properties. Notably, supplementary information should be provided to show quality control data corresponding to the final cell atlas including the UMAP showing the sample source of the cells, violin plots of gene count, UMI count, and mitochondrial fraction for the overall dataset and by cluster, and expression profiles on UMAP of selected markers characterizing cluster identities.

The second set of data aims to correlate the scRNA-seq analysis with in situ hybridizations (ISH) in two different pipefish (gulf and bay) species to identify and characterize markers spatially, and validate cell types and signaling pathways active in them. While the approach is rational, the authors must complete the data and optimize labeling protocols to support their statements. One major concern is the quality of ISH stainings and images; embryos show a high degree of pigmentation that could hide part of the expression profile, and only subparts and hardly detectable tissues/stainings are presented. The authors should provide clear and good-quality images of ISH labeling on whole-mount specimens, highlighting the magnification regions and all other organs/structures (positive controls) expressing the marker of interest along the axis. Moreover, ISH probes have been designed and produced on gulf pipefish genome and cDNA respectively, while ISH labeling has been performed indifferently on bay or gulf pipefish embryos and larvae. The authors should specify stages and species on figure panels and should ensure sequence alignment of the probe-targeted sequences in the two species to validate ISH stainings in the bay pipefish. Moreover, spatiotemporal gene expression being a very dynamic process during embryogenesis, interpretations based on undefined embryonic and larval stages of pipefish development and compared to 3dpf zebrafish are insufficient to hypothesize on developmental specificities of pipefish features, such as on the absence of tooth primordia that could represent a very discrete and transient cell population. The ISH analyses would require a clean and precise spatiotemporal expression comparison of markers at the level of the entire pipefish and zebrafish specimens at well-defined stages, otherwise, the arguments proposed on teleost innovations and adaptations turn out to be very speculative.

To conclude, whereas the scRNA-seq dataset in this unconventional model organism will be useful for the community, the spatiotemporal and comparative expression analyses have to be thoroughly pushed forward to support the claims. Addressing these points is absolutely necessary to validate the data and to give new insights to understand the extraordinary evolution of the Syngnathidae family.

https://doi.org/10.7554/eLife.97764.1.sa2

Reviewer #2 (Public Review):

Summary:

The authors present the first single-cell atlas for syngathid fishes, providing a resource for future evolution & development studies in this group.

Strengths:

The concept here is simple and I find the manuscript to be well written. I like the in situ hybridization of marker genes - this is really nice. I also appreciate the gene co-expression analysis to identify modules of expression. There are no explicit hypotheses tested in the manuscript, but the discovery of these cell types should have value in this organism and in the determination of morphological novelties in seahorses and their relatives.

Weaknesses:

I think there are a few computational analyses that might improve the generality of the results.

(1) The cell types: The authors use marker gene analysis and KEGG pathways to identify cell types. I'd suggest a tool like SAMap (https://elifesciences.org/articles/66747) which compares single-cell data sets from distinct organisms to identify 'homologous' cell types -- I imagine the zebrafish developmental atlases could serve as a reasonable comparative reference.

(2) Trajectory analyses: The authors suggest that their analyses might identify progenitor cell states and perhaps related differentiated states. They might explore cytoTRACE and/or pseudotime-based trajectory analyses to more fully delineate these ideas.

(3) Cell-cell communication: I think it's very difficult to identify 'tooth primordium' cell types, because cell types won't be defined by an organ in this way. For instance, dental glia will cluster with other glia, and dental mesenchyme will likely cluster with other mesenchymal cell types. So the histology and ISH is most convincing in this regard. Having said this, given the known signaling interactions in the developing tooth (and in development generally) the authors might explore cell-cell communication analysis (e.g., CellChat) to identify cell types that may be interacting.

https://doi.org/10.7554/eLife.97764.1.sa1

Reviewer #3 (Public Review):

Summary:

This study established a single-cell RNA sequencing atlas of pipefish embryos. The results obtained identified unique gene expression patterns for pipefish-specific characteristics, such as fgf22 in the tip of the palatoquadrate and Meckel's cartilage, broadly informing the genetic mechanisms underlying morphological novelty in teleost fishes. The data obtained are unique and novel, potentially important in understanding fish diversity. Thus, I would enthusiastically support this manuscript if the authors improve it to generate stronger and more convincing conclusions than the current forms.

Weaknesses:

Regarding the expression of sfrp1a and bmp4 dorsal to the elongating ethmoid plate and surrounding the ceratohyal: are their expression patterns spatially extended or broader compared to the pipefish ancestor? Is there a much closer species available to compare gene expression patterns with pipefish? Did the authors consider using other species closely related to pipefish for ISH? Sfrp1a and bmp4 may be expressed in the same regions of much more closely related species without face elongation. I understand that embryos of such species are not always accessible, but it is also hard to argue responsible genes for a specific phenotype by only comparing gene expression patterns between distantly related species (e.g., pipefish vs. zebrafish). Due to the same reason, I would not directly compare/argue gene expression patterns between pipefish and mice, although I should admit that mice gene expression patterns are sometimes helpful to make a hypothesis of fish evolution. Alternatively, can the authors conduct ISH in other species of pipefish? If the expression patterns of sfrp1a and bmp4 are common among fishes with face elongation, the conclusion would become more solid. If these embryos are not available, is it possible to reduce the amount of Wnt and BMP signal using Crispr/Cas, MO, or chemical inhibitor? I do think that there are several ways to test the Wnt and/or BMP hypothesis in face elongation.

https://doi.org/10.7554/eLife.97764.1.sa0