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

Running title: Novel Syngnathid Developmental Atlas

Hope M. Healey[1]*, Hayden B. Penn[1], Clayton M. Small[1,2], Susan Bassham[1], Vithika Goyal[1], Micah A. Woods[1], and William A. Cresko[1,3*]

1. Institute of Ecology and Evolution, University of Oregon
2. School of Computer and Data Science, University of Oregon
3. Knight Campus for Accelerating Scientific Impact, University of Oregon

*Corresponding authors

ORCID IDs

Hope M. Healey:	0000-0001-9978-1553
Hayden B. Penn:	0009-0000-3090-5423
Clayton M. Small:	0000-0003-1615-7590
Susan Bassham:	0000-0002-7309-2095
Vithika Goyal:	0009-0009-8601-8474
Micah A. Woods:	0009-0004-2156-3352
William A. Cresko:	0000-0002-3496-8074

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

Keywords: Evolution of Development, Syngnathid, Evolutionary Novelty, scRNAseq, Non-traditional Model

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

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.

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

64 protected by a bony dermal armor rather than scales (Jungerson, 1910). Other skeletal
65 differences include a lack of ribs and pelvic fins, and an expansion in the number of
66 vertebrae (Schneider, Woltering, et al., 2023). Finally, syngnathids exhibit male pregnancy,
67 which has involved the evolution of specialized brooding tissues and structures
68 (Whittington & Friesen, 2020). Paternal investment varies among lineages; for example,
69 seadragons tether embryos externally to their tails while seahorses and some pipefishes
70 have enclosed brood pouches proposed to support embryos through nutrient transfer and
71 osmotic regulation (Carcupino, 2002; Melamed et al., 2005; Ripley & Foran, 2006).

72 Despite advances in understanding the ecology and evolution of syngnathid
73 novelties, the developmental genetic basis for these traits is largely unknown. The recent
74 production of high quality syngnathid genome assemblies (Qu et al., 2021; Ramesh et al.,
75 2023; Small et al., 2022; Wolf et al., 2024) provides initial clues for the developmental
76 genetic basis of some evolutionary changes. Studies have found that syngnathids lack
77 several genes with deeply conserved roles in vertebrate development, including
78 pharyngeal arch development (*fgf3*), tooth development (*fgf3*, *fgf4*, and *eve1*, and most scpp
79 genes), fin development (*tbx4*), and immune function (MHC pathway components) (Lin et
80 al., 2016; Qu et al., 2021; Small et al., 2016; Small et al., 2022; Zhang et al., 2020). Though
81 these gene losses are highly suggestive of leading to unique changes, exploration of the
82 actual developmental consequences of their losses is needed.

83 To fill this gap in knowledge, we used single cell RNA sequencing (scRNAseq) to
84 investigate how these striking genomic changes have affected the developmental genetic
85 and cellular basis of syngnathids' evolutionary derived traits. The Gulf pipefish (*Syngnathus*
86 *scovelli*) is an attractive model for this study (Figure 1A). This species has a high-quality
87 reference genome annotated by NCBI and is amenable to laboratory culture (Anderson &
88 Jones, 2019; Ramesh et al., 2023). Furthermore, species from the *Syngnathus* genus are
89 used worldwide to address questions about syngnathid evolution in microbial,
90 developmental, histological, transcriptomic, ecotoxicological, and genomic studies
91 (Berglund et al., 1986; Carcupino, 2002; Fuiten & Cresko, 2021; Harada et al., 2022; Harlin-
92 Cognato et al., 2006; Partridge et al., 2007; Ripley & Foran, 2006a; Rose et al., 2023; Roth et
93 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.

124 # Results

125 ## Valuable scRNAseq atlas for studying syngnathid development

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

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

141 In total, our atlas contained 13,027 connective tissue cells (excluding cells from
142 blood, immune, and the digestive system) from 14 clusters, 10,112 nervous system cells
143 from ten clusters, 4,363 muscle cells from five clusters, 4,133 blood cells from three
144 clusters, 650 immune cells from two clusters, 432 pigment cells from one cluster, 370
145 epidermal cells from one cluster, and 137 gut cells from one cluster. Within the connective
146 tissue cell types, we also identified cartilage (302 cells), developing bone (442 cells), fins
147 (253 cells), and notochord (693 cells).

148 ## Discovery of cell cluster function and state using KEGG analysis

149 To affirm identities and discover potential properties of each cluster, we completed
150 a KEGG pathway analysis for each cluster using Seurat's marker genes (Figure S4, Figure

151 2B). For eight of the clusters (1, 4, 6, 9, 11, 15, 19, and 24), we did not find any significantly
152 enriched pathways, possibly due to similar gene expression profiles across cell types that
153 reduced the number of identified markers. However, we found one or more significantly
154 enriched pathways for the other 29 cell clusters. We observed enriched pathway terms that
155 supported cluster annotations. For example, phototransduction in the retina cluster,
156 melanogenesis in the pigment cluster, cardiac muscle contraction in muscle clusters, and
157 neuroactive ligand receptor interaction in neuronal clusters.

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

172 Commonalities of cell clusters, unique networks, and elusive cell types identified 173 in network analysis

174 We built gene networks/modules from 3,000 variable genes using Weighted Gene
175 Network Correlation Analysis (WGCNA; Langfelder & Horvath, 2008). This produced 43
176 gene modules in total (Supplementary File 5,6), assessed for each cluster-module pair for
177 their strength of association (Figure 3A, Supplementary File 7,8) and every module's
178 dependence on each cluster for their network connectivity (Figure S5, Supplementary File
179 9,10). Using the genes from each network, we completed a KEGG pathway analysis to

180 identify whether gene modules indicated specific cellular pathways or states (Figure 3B).
181 We initially explored whether these network-cluster associations could reveal
182 commonalities between cell clusters or identify whether particular clusters contained
183 multiple cell identities.

184 First, we asked whether gene modules that associate with three or more cell clusters
185 signify commonalities between clusters that have similar cell types. We identified seven
186 gene modules (6, 7, 21, 14, 17, 41, and 42) that each associate with three or more cell
187 clusters. These gene modules do connect clusters of similar identities or cell states. For
188 example, Modules 6 and 7 associate with connective tissue cells. Module 6, the larger of the
189 two modules, contains numerous other KEGG pathways found in most connective cell
190 clusters, such as 'cytokine-cytokine receptor interaction' (Figure 3C). Interestingly, Module
191 21, associated with clusters 8, 10, 16, and 25, contains genes from KEGG pathways related
192 to the 'cell cycle' and 'cellular senescence', supporting the results found in our cluster based
193 KEGG pathway analysis (Figure 3D). Module 21 also contained 'Notch signaling' genes,
194 possibly due to similar correlations with cell cycle genes; however, these were only
195 expressed in the neural cell clusters (8 and 25).

196 Where one cell cluster is associated with multiple gene networks, we wondered if
197 multiple cell identities existed within the cell cluster. We explored this possibility by
198 examining the pigment cell cluster and whether the five different correlated gene modules
199 (#34, 35, 36, 37, and 38) are expressed in distinct subgroups of cells within the cluster. We
200 found cases of more than one genetic network (modules 35 and 38) expressed in the same
201 cells. Modules 35 and 38 contain conserved pigment genes, *pmela*, *mlana*, and *dct* in
202 module 35 and *tryp1b* and *pmel* in module 38, that mark melanocytes (Figure S6; Du et al.,
203 2003; Johnson et al., 2011; Lamason et al., 2005; B. Thisse & Thisse, 2004). However, we
204 also found non-overlapping expression of networks, notably modules 36 and 37. We
205 inferred that module 36 is associated with xanthophores and xanthoblasts due to the
206 presence of *plin6* and *scarb1*, genes involved with lipid binding and activity in
207 xanthophores (Ahi et al., 2020). On the other hand, module 37 likely represents iridophores
208 and iridoblasts because it contains *pnp4a,* which is involved in purine-nucleoside
209 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*

239 had additional expression domains in those models: tongue, teeth, and Meckel's cartilage in
240 mice, and in hindbrain and all pharyngeal arches in zebrafish (Bjork et al., 2010; Ding et al.,
241 2013). Our observations are consistent with the literature, except *prdm16* has not been
242 observed above the ethmoid in zebrafish.

243 We next examined signaling pathways active in craniofacial development. Our KEGG
244 pathway analysis revealed that MAPK, Wnt, and TGF-beta signaling pathways were
245 significantly enriched in one or more craniofacial contributing cell clusters (module #6;
246 Figure 2B, Figure 3C, Figure 4D). We chose three genes, one from each major pathway, for
247 *in situ* hybridizations: *fgf22* (a fgf ligand present in the actively dividing cells module),
248 *bmp4* (a bmp ligand present in the largest connective tissue module), and *sfrp1a* (a Wnt
249 pathway enabler present in a cartilage gene module). We found *fgf22* is expressed at the tip
250 of the palatoquadrate and Meckel's cartilage and in the pharyngeal arches (figure 4E). Our
251 finding is surprising given that *fgf22* is primarily expressed in the nervous system and has
252 never been observed in craniofacial skeletal development (Miyake & Itoh, 2013; Umemori
253 et al., 2004). We observed *bmp4* and *sfrp1a* expressed above the ethmoid plate and along
254 the ceratohyal (figure 4F-G). *Bmp4* has a conserved role in craniofacial development,
255 particularly important at later stages for driving chondrocyte differentiation (Wang et al.,
256 2024; Zhou et al., 2013). However, contrary to pipefish, *bmp4* in zebrafish is minimally
257 expressed in the ethmoid region during cartilage formation (Schilling & Kimmel, 1997;
258 Swartz et al., 2011; Wang et al., 2024). Interestingly, *sfrp1a* has not been observed in the
259 palate of mice or ethmoid region of zebrafish, but *sfrp1a* craniofacial expression has been
260 observed in other fishes with derived craniofacial structures (Ahi et al., 2014; Schneider,
261 Gunter, et al., 2023).

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

263 Previous papers have identified possible candidate genes for the loss of teeth in
264 syngnathid fishes including those from genes that initiate tooth bud formation (*fgf4, eve1*),
265 regulate tooth morphogenesis (*fgf3, fgf4*), and synthesize tooth minerals (*scpp4, scpp7,*
266 *scpp9, odam*, and *scpp5*; Lin et al., 2016; Qu et al., 2021; Small et al., 2016; Small et al., 2022;
267 Zhang et al., 2020). However, it is unknown whether syngnathid tooth development

268 initiates then halts or whether it never begins. We searched for signs of early tooth
269 primordia within our atlas to ask whether tooth development might initiate in syngnathids.
270 Additionally, we examined whether genes present in mature teeth are still expressed in
271 syngnathids and what types of cells express them.

272 Our cell cluster annotations did not identify a tooth primordium cluster. We
273 therefore searched for tooth primordia by examining expression of specific fodontogenesis
274 marker genes (*aldh1a2*, *bmp4*, *dlx2a*, *dlx3b*, *lef1*, *lhx6a*, *lhx8*, *msx1a*, *msx2*, and *wnt10a*;
275 Figure 5A). We observed several primordium genes expressed in our atlas. However, there
276 was no cluster with every marker gene expressed in over 10% of cells. Several markers
277 (*dlx3b*, *lef1*, *msx1a*, *msx2*, and *wnt10a*) were expressed in cluster #28, a fin cluster
278 distinguished by *hoxa13a* and *hoxa13b* expression. Since we previously noted that cluster
279 #16 seems to be a primordial connective tissue cluster, we wondered if it could contain
280 tooth primordial cells. In this cluster, we found the following percentage of expression of
281 each gene in the cells: *bmp4* in 46%, *aldh1a2* in 29%, *lef1* in 24%, *dlx2a* in 18%, *dlx3b* in
282 16.4%, *eda* in 12.7%, *msx1a* in 10.6%, *pitx2* in 10%, *lhx6a* in 8%, *wnt10a* in 4.6%, *msx2* in
283 2.65%, *lhx8* in 1.8%, and *shha* in 1%.

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

296 We next investigated Scpp genes (*enam*, *scpp1*, and *spp1*), which are expressed in
297 late tooth development and tooth maintenance in other vertebrates (Figure 5A). These
298 genes also have some expression outside of teeth such as in dental bone (*scpp1*, *spp1;*
299 Kawasaki, 2009) and fins (*enam;* Jain et al., 2007). We identified *spp1* and *scpp1* expression
300 in 54.5% and 24% of bone cells, respectfully, and sparse *spp1* expression (<3%) in other
301 connective tissue cell types. We found minimal expression of *enam* (in 11.4%) of epidermal
302 cells and in less than 8% of muscle cells. In wholemount *in situ* hybridization, we found that
303 *scpp1* is expressed in all developing pipefish bones, both endochondral and dermal (Figure
304 5G). Since Scpp gene losses observed in syngnathids have been hypothesized to be
305 responsible for their tooth loss (Lin et al., 2016a; Qu et al., 2021; Zhang et al., 2020a), we
306 explored Scpp gene family content in the close toothed relative to syngnathids, the blue
307 spotted cornetfish, and we found several gene losses (*scpp4*, *scpp7*, and *scpp9*) and a likely
308 pseudogene (*scpp5*) (Figure S7; Hughes et al., 2018; Stiller et al., 2022).

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

312 **Tooth and skeletal genes are expressed during dermal armor development**

313 Syngnathid dermal armor is mineralized dermal bone underneath the skin (Figure
314 6D, G, H). It is unknown when dermal armor primordia initiate and how they are patterned.
315 Spatial expression analysis of *pitx2* and *dlx3b* in search of tooth primordia instead revealed
316 expression of these genes in possible dermal armor primordia (Figure 6A,B, E,F). We found
317 *pitx2* staining localized dorsally to striated muscle underneath developing dermal armor.
318 *dlx3b* is expressed in a repeating pattern along the body in the epidermal and dermal
319 tissues. Both staining patterns were in discrete regions of the muscle and epidermal layers
320 rather than being continuously expressed across the tissues. We did not observe the
321 expression of other tooth primordium genes (*bmp4, lhx6a*, and *dlx2a*) in this region.

322 Because the embryos from our atlas had not begun dermal armor mineralization,
323 the atlas cannot be directly used for discovery of genes active in dermal armor. However,
324 the atlas contains osteoblasts from craniofacial bones which we used to create osteoblast

325 specific gene networks. We therefore asked whether these osteoblast genes were present
326 in mineralizing dermal armor at later stages. *In situ* hybridization expression analysis
327 revealed that *scpp1*, an osteoblast and tooth mineralization gene, and *ifitm5*, an osteoblast
328 gene, were expressed in the dermal armor at the onset of mineralization (Figure 6C; Figure
329 S3).

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

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

337 Within our larger KEGG analysis, we asked whether nutrient absorption KEGG terms
338 were among the enriched pathways for the epidermal cells. We identified 106 enriched
339 genes in the endocytosis pathway (p-value=.036; Figure 7A). Four metabolism pathways
340 (galactose, glutathione, sphingolipid, and starch and sucrose) are also enriched. For
341 comparison, we investigated whether these KEGG terms are also enriched in epidermal
342 cells of non-brooding fishes, we completed a KEGG pathway analysis on a comparably
343 staged zebrafish single cell RNA sequencing atlas (3dpf; Lange et al., 2023). The zebrafish
344 epidermal cells did not have a significant enrichment of the endocytosis pathway (23
345 genes, p-value=.99) or any metabolism pathway. However, 11 of the 23 endocytosis genes
346 expressed in the zebrafish epidermis were also present in the pipefish epidermis.

347 We next examined epidermal gene networks. We found that the largest epidermal
348 gene network (#16) contained a striking enrichment of C type lectin genes, carbohydrate
349 binding proteins that possess antimicrobial properties (Figure 7B). This network contained
350 14 total lectin genes expressed in the epidermal cells: five galactose-specific lectin
351 nattectin, 2 alpha-N-acetylgalactosamine-specific lectins, one L-rhamnose-binding lectins,
352 four ladderlectin, one C-type lectin 37Dd-like, and one C-type lectin domain family 4
353 member G-like. Through examining lectin gene expression in the entire dataset, we found

354 that these genes were specific to epidermal cells. Interestingly, previous literature has
355 identified an upregulation of C-type lectins in brood pouch tissues throughout different
356 stages of syngnathid pregnancy (Roth et al., 2020; Small et al., 2013; Whittington et al.,
357 2015).

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

369 Discussion

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

380 Our atlas represents a novel resource for Evo-Devo research

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

394 For syngnathids specifically, this atlas represents an important step towards
395 understanding the genetic nature of unique syngnathid traits. Numerous developmental
396 genetic changes can lead to evolutionary innovations, including the evolution of novel
397 genes, gene duplications, gene losses, gene family expansions or contractions, evolution of
398 regulatory elements, co-option of gene regulatory networks, re-wiring of gene networks,
399 assembly of novel gene networks, and/or the emergence of novel cells (Arendt et al., 2016;
400 Cãestro et al., 2007; Teichmann & Babu, 2004; A. Wagner, 2011; G. P. Wagner & Lynch,
401 2010). Considering these possibilities, we examined select syngnathid traits and speculated
402 on developmental genetic mechanisms influencing their evolution. Continuing to
403 investigate these proposed mechanisms through expanded scRNAseq atlases and other
404 studies will be critical for understanding syngnathid evolution.

405 Conserved pathways may contribute to derived syngnathid heads

406 Syngnathids have highly derived heads including an elongated ethmoid region,
407 uniquely shaped hyoid, and altered muscles and tendons to support specialized "pivot
408 feeding". The developmental underpinnings of these derived traits have remained
409 underexplored. In our atlas, we identified cell types present in the developing pipefish head

410 and genetic pathways active in those cell types. We identified numerous cells that were
411 present in the developing face: cartilage, bone, tendons, ligaments, osteochondrogenic
412 mesenchyme, fibroblasts, and unclassified connective tissue cells. Overall, we did not find
413 any unrecognizable cell types, suggesting that genetic modifications within conserved cell
414 types may drive craniofacial modifications.

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

429 We found *fgf22* expressed in the mandible and pharyngeal arches using *in situ*
430 hybridizations, but not in the elongating regions of the head. Interestingly, *fgf22* has not
431 been reported in craniofacial development of any other species (Miyake & Itoh, 2013). Fgf
432 signaling, however, is a conserved and essential pathway for craniofacial development
433 (Crump et al., 2004; Leerberg et al., 2019; McCarthy et al., 2016; Szabo-Rogers et al., 2008;
434 Walshe & Mason, 2003; Woronowicz & Schneider, 2019), raising the possibility that *fgf22*
435 has been co-opted into a role played by different Fgf genes in other species. Future work
436 should investigate whether *fgf22* craniofacial expression is observed in other species and
437 consider a relationship between *fgf22*'s novel craniofacial expression and the loss of Fgf
438 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., 2020a), 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

468 the retained genes *scpp1*, *spp1*, and *enam,* are expressed in structures outside of tooth
469 development, suggesting developmental pleiotropic constraint. Specifically, we found *spp1*
470 and *scpp1* expressed in osteoblasts which is consistent with zebrafish (Bergen et al., 2022;
471 Kawasaki, 2009; Liu et al., 2016) and *enam* expressed in the epidermis which has not been
472 reported in zebrafish (Goldsmith et al., 2003; Jain et al., 2007; Liu et al., 2016). Through
473 examining the conservation of the scpp genes in close syngnathid relatives, we found that
474 most scpp gene losses (*scpp4*, *scpp7*, and *scpp9* and a functional *scpp5*) are shared with a
475 tooth-bearing outgroup to the family, and likely occurred prior to the loss of teeth in
476 syngnathids. Overall, our analysis favors the hypothesis that pleiotropic Scpp genes were
477 retained in syngnathids while other Scpp genes were lost due to relaxed selection.

478 Redeployment of the bone gene network to build dermal armor

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

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

492 Instead, dermal armor might originate from the co-option of existing bone
493 development gene regulatory networks. *Dlx3b* expression has been observed in epithelia
494 and mesenchyme during dermal and perichondral bone development in zebrafish
495 (Verreijdt et al., 2006). In addition to *dlx3b*, we observed bone development genes *scpp1*
496 and *ifitm5* expressed in the ossifying dermal armor. Future studies could test our

497 hypothesis that the dermal armor evolved through re-deployed osteoblast networks by
498 examining osteoblast gene network expression over time.

499 Signatures of embryonic interactions within the novel pouch environment

500 Syngnathid embryos are reared within the brood pouch, a novel structure and
501 environment composed of male derived tissues (epithelium and placental-like tissues that
502 include specialized cell types) that harbors a pouch microbiome (Stölting & Wilson, 2007).
503 During pregnancy, the male brood pouch undergoes numerous changes including increased
504 vascularization and altered expression of immune genes (Harada et al., 2022; Ripley et al.,
505 2010; Roth et al., 2020; Small et al., 2013; Whittington et al., 2015). Researchers predict
506 that these changes relate to nutrient and waste transfer and prevention of embryonic
507 rejection and bacterial infection (Dudley et al., 2021; Whittington & Friesen, 2020).
508 However, there are few studies that examine whether and how embryos interact with the
509 brood pouch environment (Kvarnemo et al., 2011; Ripley & Foran, 2006). To consider
510 whether embryos have specializations for life in this brood pouch environment, we asked
511 about cell type-specific expression of nutrient acquisition and/or immune genes.

512 Pipefish embryos uptake paternally derived carbohydrates, proteins, and lipids
513 (Kvarnemo et al., 2011; Ripley & Foran, 2006). Our data suggests that this uptake could
514 occur through the embryonic epidermis. Specifically, we noticed an enrichment of
515 endocytosis and metabolism genes in epidermal cells. Epidermal absorption of maternally
516 derived nutrients has been suggested in viviparous fishes (Tengfei et al., 2021; Wourms,
517 1981). Interestingly, microvilli, a type of cellular projection, have been observed on the
518 anal fin of developing seahorses (Wetzel & Wourms, 2004). Possibly, these microvilli may
519 be functionally equivalent to those on the small intestine, maximizing nutrient absorption
520 from the environment.

521 During pregnancy, the male brood pouch increases expression of C-type lectin genes
522 (Roth et al., 2020; Small et al., 2013; Whittington et al., 2015). These genes are
523 transmembrane or secreted receptors that sense self or non-self and are primarily studied
524 for their role in innate and adaptive immunity (Brown et al., 2018). We identified 14 C-type
525 lectin genes expressed in the embryonic epidermis. Our work suggests that lectin genes are

526 produced by both the father and the embryos, but their function is still unclear. Syngnathid
527 research has primarily suggested that lectin genes are produced to prevent bacterial
528 infection (Melamed et al., 2005), though they could be important for male-embryo
529 recognition.

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

537 Conclusions

538 Our study represents the first scRNAseq developmental atlas in syngnathids, and
539 one of the first non-model developmental scRNAseq atlases, providing a major step
540 forward for evo-devo research. We used our atlas to begin addressing questions on the
541 evolution and development of syngnathid innovations including their unique craniofacial
542 structure, loss of teeth, dermal armor, and development within the male brood pouch. By
543 combining scRNAseq analysis with spatial expression data from *in situ* hybridization, we
544 made important discoveries in cell type identity and distribution as well as spatial
545 expression of marker and signaling genes. We found that syngnathids express conserved
546 signaling pathways during craniofacial development, suggesting that alterations within
547 these pathways may be important for the evolution of their craniofacial skeletons. We did
548 not find evidence for tooth primordia within syngnathids and propose that genetic changes
549 early in tooth development could have led to their loss of teeth. We propose that the re-
550 deployment of bone gene networks, but probably not tooth networks, could play a role in
551 the dermal armor development. Finally, we observed an enrichment of endocytosis genes
552 and many C-type lectin genes in epidermal cells, which suggests ways these cells might
553 interact with the brood pouch environment. Our atlas pushes forward our understanding of

554 syngnathid development and evolution, and provides resources for developmental genetic
555 analysis in nascent evo-devo model species.

556

557 # Methods

558 ## Single Cell RNA Sequencing Libraries Preparation

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

566 We dissociated the embryos using 460ul of .25% trypsin in water and 40ul 100
567 mg/mL collagenase I (Sigma C0130-200mg) for 16 minutes. We filtered cells using a 40uM
568 cell strainer (Thomas Scientific #1181X52). We quantified cell concentrations using the
569 TC20 Automated Cell Counter (Biorad) then diluted the samples to 800 cells/ul in 4% BSA
570 in PBS. The University of Oregon Genomics and Cell Characterization Core (GC3F;
571 https://gc3f.uoregon.edu) prepared single cell libraries for each sample using 10X
572 Genomics Single Cell 3' Genome Expression mRNAseq kit with NextGEM v3.1 chemistry.
573 We sequenced these libraries on an S4 lane on the NovaSeq 6000 at the GC3F. To improve
574 the 3' UTR genome annotations, we also prepared scISOrSeq libraries from the first
575 embryonic sample and from dissociated pouch cells from pregnant and nonpregnant males.
576 These libraries were produced in accordance with (Healey et al., 2022). Embryonic,
577 pregnant pouch and non-pregnant pouch libraries were sequenced separately on PacBio
578 Sequel II - SMRT Cells 8M.

579 To turn the scISOrSeq reads into gene models, we followed the pipeline from Healey
580 et al. 2022. We ran our custom script (scISOr_Seq_processing.py) to remove barcodes,
581 identify cell barcodes, and demultiplex with the single cell flag and appropriate barcodes
582 (5' CCCATGTACTCTGCGTTGATACCACTGCT and 3' CTACACGACGCTCTTCCGATCT). We
583 aligned the reads to the 2022 Gulf pipefish genome GenBank: GCA_024217435.2) using
584 minimap v2.9 (cite: Li 2018). We filtered the reads using cDNA cupcake to remove

585 duplicate transcripts (Tseng, 2021). We used SQANTI3 to identify gene models and filter
586 them (Tardaguila et al., 2018). We merged the SQANTI3 annotations with the Gulf pipefish
587 genome (NCBI: ROL_Ssco_1.1) using TAMA merge (Kuo et al., 2017). Since the Gulf pipefish
588 genome does not contain mitochondrial genes, we appended the genome files with the Gulf
589 pipefish mitochondrial genome (NCBI RefSeq: NC_065499.1).

590 ## Single Cell Atlas Construction

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

597 To remove extraneous RNA counts from the dataset, we used SoupX (v1.5.2; Young
598 & Behjati, 2020). We identified doublet scores for our dataset using scrublet (v0.2.3). The
599 doublet removal step reduced the first sample by 114 cells (from 20,733 cells to 20,619
600 cells) and second sample by 167 cells (from 17,626 cells to 17,459 cells). We finally
601 removed cells with less than 500 features, greater than 9000 features, greater than 1E5
602 RNA counts, with a scrublet score greater than the detected threshold (.76 for sample 2 and
603 .21 for sample 2), or greater than 10% mitochondrial reads. The second filtering step
604 removed 727 cells from sample one (20,619 cells to 19,892 cells) and 1,566 cells from
605 sample two (from 17,459 cells to 15,893 cells).

606 We normalized the datasets with SCTransform (v0.3.3). We used Seurat's
607 integration tools, `SelectIntegrationFeatures` using 3,000 feature genes,
608 `FindIntegrationAnchors` using `SCT` normalization, and `IntegrateData` using `SCT`
609 normalization, to integrate the two datasets. After integration, our combined atlas had
610 35,785 cells (Supplementary File 1,2). We then used the integrated dataset to complete
611 PCA analysis. We tested using a variety of principle components for further analysis and
612 chose 30 PCs for our analysis based on the clear delineation of major cell types. We next
613 clustered the cells using 30 PCs and plotted the data on a UMAP with Seurat.

614 ## Single Cell Atlas Cluster Identification

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

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

634 *Syngnathus scovelli* used for *in situs* were purchased from Alyssa's Seahorse Saavy
635 and Gulf Specimens Marine Lab then reared in our facility at 25 degrees C water and 25-28
636 PPT Salinity. Since *Syngnathus leptorhynchus* does not have a published genome, we
637 designed probes using NCBI Primer Blast with the Gulf pipefish genome and produced
638 these probes using Gulf pipefish embryonic cDNA pools. To create the probes, we
639 completed two rounds of PCR. The first round used a gene specific forward primer with a
640 reverse gene specific primer that had 10 nucleotides of T7 promotor sequence attached.
641 PCR products were cleaned with Zymo clean and concentrator DNA kit and eluted in 15 ul
642 of elution buffer. Round two of PCR used the same gene specific forward primer with a

643 modified T7 promoter sequence (TGGACTAATACGACTCACTATAGGG) as the reverse
644 primer, and finally the product was cleaned again with Zymo clean and concentrator DNA
645 kit and eluted in 15 ul of elution buffer.

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

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

666 **Single Cell KEGG Analysis**

667 To identify pathways upregulated in cell clusters, we completed a KEGG analysis.
668 We downloaded gulf pipefish KEGG pathways (https://www.kegg.jp). For the KEGG
669 analysis, we used the marker genes identified from our `FindAllMarkers` list as the input
670 genes. We converted these gene ids using `keggConv` to KEGG ids. We used a Wilcoxon

671 enrichment test to ask whether cluster marker genes were enriched for each KEGG
672 pathway.

673 ### Single Cell Atlas Network Analysis

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

682 We calculated the dissimilarity matrix by calculating TOM similarity of the
683 adjacency matrix and subtracting it from one. We then created a gene tree through `hclust`
684 of the dissimilarity matrix with the average method. From the gene tree, we created
685 modules using `cutreeDynamic` with deepSplit of two and the minClusterSize of 15 (setting
686 the smallest cluster size to 15). We calculated module eigengenes using `moduleEigengenes`
687 of the adjacency matrix then calculated the dissimilarity of the module eigengenes with `cor`
688 of the module eigengenes subtracted from one and clustered the module eigengenes with
689 `hclust`. We then chose a dissimilarity of 35% as the cut off for merging modules; however,
690 no modules had dissimilarity scores with each other below 40%. Since the module
691 eigengenes are calculated for each cell in the dataset, we used these values to calculate the
692 t-statistic for each cell cluster (considered each sample) in the module (all of the module
693 eigengene scores were used as the population).

694 We next tested the hypothesis that certain cell clusters were strongly associated
695 with specific gene modules through a two-way permutation test (1000 permutations) and
696 corrected p-values to control the False Discovery Rate (FDR). Additionally, we tested
697 whether specific cell clusters drove underlying gene network structure by measuring
698 network connectivity. The network connectivity was first calculated using the entire
699 dataset, then each cell cluster was progressively dropped from the dataset and the

700 connectivity was remeasured. This resulted in change in connectivity scores for each
701 module-cluster pair. To assess whether these changes were significant, we completed 1000
702 permutations whereby cells were randomly dropped (the number of cells dropped was
703 equal to the cell cluster size of the focal cluster), connectivity measured, and the change in
704 connectivity recorded. The p-value is the number of instances where the change in
705 connectivity is greater in the permutations than in the focal cell cluster run. P-values were
706 corrected to control the FDR.

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

713 *In situ* Hybridization

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

728 ## Bone and Cartilage Staining

729 We used alcian and alizarin stains to mark cartilage and bones. We followed the
730 protocol from Walker and Kimmel (2007) with minor alterations. We stored samples in
731 50% glycerol/0.1% KOH at 4C and imaged them in 100% glycerol. After imaging, we white
732 balanced the photographs using Photoshop (v23.4.2) levels tool.

733 ## Gene Cluster Analysis

734 To examine close syngnathid outgroups, we downloaded the 2023 mandarin dragonet
735 genome from NCBI (GenBank assembly accession: GCA_027744825.1) and used an
736 unpublished highly contiguous cornetfish genome. Since these genomes were unannotated,
737 we manually identified scpp genes. We searched for scpp genes using BLASTN with medaka
738 and additional fish sequences as the query. We also searched for these genes with mVISTA
739 plots across conserved gene synteny regions (LAGAN alignment using translated anchoring)
740 with medaka as the focal species. To identify scpp genes in additional species, we gathered
741 cluster information from NCBI and ensembl. Additionally, we used mVISTA plots to further
742 search for unannotated scpp genes.

743

744 ## Acknowledgements:

745 We are indebted to Emily Rose for her collection of the Gulf pipefish specimens for
746 the single cell libraries. We are grateful for everyone in the Cresko Lab for their ideas and
747 participation in early morning pipefish collection trips. We are additionally thankful for
748 Emily Beck for her assistance collecting pipefish. In particular, we are immensely
749 appreciative of Mark Currey and Tiffany Thornton's valiant efforts to culture Gulf pipefish.
750 Additionally, we are grateful for Balan Ramesh and Adam Jones for collaborating with the
751 Cresko Lab to produce new Syngnathiformes genomes. We thank Tina Arredondo from the
752 UO GC3F for her preparation of our single cell RNA sequencing libraries. This work was
753 funded by the National Science Foundation Grant OPP-2015301 (to WAC, SB, and CMS),
754 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.

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.

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

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

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

785 **SupplementaryFile8:**
786 **240228_ssc_WGCNA_associat_mods_w_clus_tstat_pvalues_corrected.csv,** p-values for
787 the t-statistics of the strength of association between gene modules and cell clusters. P-
788 values are corrected for multiple testing hypotheses using fdr. The cell clusters are in the
789 rows and the gene modules are in the columns.

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

793 **SupplementaryFile10:**
794 **240228_WGCNA_EM1_EM2_Connectivity_Permutation_Pvalues_Corrected.csv,** the p-
795 values for the change in connectivity are found in this csv. The cell clusters are in the
796 columns and the gene networks are in rows.

797 ## References

798 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), 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 21:1*, *21*(1), 1–15. 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., Musser, J. M., H Baker, C. V, Bergman, A., Cepko, C., Erwin, D. H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M. D., & Wagner, G. P. (2016). *The origin and evolution of cell types*. https://doi.org/10.1038/nrg.2016.127

Bergen, D. J. M., Tong, Q., Shukla, A., Newham, E., Zethof, J., Lundberg, M., Ryan, R., Youlten, S. E., Frysz, M., Croucher, P. I., Flik, G., Richardson, R. J., Kemp, J. P., Hammond, C. L., & Metz, J. R. (2022). Regenerating zebrafish scales express a subset of evolutionary conserved genes involved in human skeletal disease. *BMC Biology 2021 20:1*, *20*(1), 1–25. 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*. *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*(5), 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*(6), 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

826 transcriptomic data across different conditions, technologies, and species. *Nature*
827 *Biotechnology*, *36*(5), 411–420. https://doi.org/10.1038/nbt.4096

828 Cãestro, C., Yokoi, H., & Postlethwait, J. H. (2007). Evolutionary developmental biology and
829 genomics. In *Nature Reviews Genetics* (Vol. 8, Issue 12, pp. 932–942). Nature
830 Publishing Group. https://doi.org/10.1038/nrg2226

831 Carcupino, M. (2002). Functional significance of the male brood pouch in the reproductive
832 strategies of pipefishes and seahorses: a morphological and ultrastructural
833 comparative study on three anatomically different pouches. *Journal of Fish Biology*,
834 *61*(6), 1465–1480. https://doi.org/10.1006/jfbi.2002.2160

835 Chari, T., Weissbourd, B., Gehring, J., Ferraioli, A., Leclère, L., Herl, M., Gao, F., Chevalier, S.,
836 Copley, R. R., Houliston, E., Anderson, D. J., & Pachter, L. (2021). Whole-animal
837 multiplexed single-cell RNA-seq reveals transcriptional shifts across Clytia medusa cell
838 types. *Science Advances*, *7*(48), 1683.
839 https://doi.org/10.1126/SCIADV.ABH1683/SUPPL_FILE/SCIADV.ABH1683_TABLES_
840 S3_AND_S5.ZIP

841 Chen, J. W., & Galloway, J. L. (2014). The development of zebrafish tendon and ligament
842 progenitors. *Development*, *141*(10), 2035–2045. https://doi.org/10.1242/DEV.104067

843 Chen, Y. P., Zhang, Y., Jiang, T. X., Barlow, A. J., St. Amand, T. R., Hu, Y., Heaney, S., Francis-
844 West, P., Chuong, C. M., & Maas, R. (2000). Conservation of early odontogenic signaling
845 pathways in Aves. *Proceedings of the National Academy of Sciences*, *97*(18), 10044–
846 10049. https://doi.org/10.1073/PNAS.160245097

847 Crump, J. G., Maves, L., Lawson, N. D., Weinstein, B. M., & Kimmel, C. B. (2004). An essential
848 role for Fgfs in endodermal pouch formation influences later craniofacial skeletal
849 patterning. *Development*, *131*(22), 5703–5716. https://doi.org/10.1242/dev.01444

850 Ding, H. L., Clouthier, D. E., & Artinger, K. B. (2013). Redundant roles of PRDM family
851 members in zebrafish craniofacial development. *Developmental Dynamics*, *242*(1), 67–
852 79. https://doi.org/10.1002/DVDY.23895

853 Docheva, D., Hunziker, E. B., Fässler, R., & Brandau, O. (2005). Tenomodulin Is Necessary for
854 Tenocyte Proliferation and Tendon Maturation. *Molecular and Cellular Biology*, *25*(2),
855 699–705. https://doi.org/10.1128/MCB.25.2.699-705.2005

856 Du, J., Miller, A. J., Widlund, H. R., Horstmann, M. A., Ramaswamy, S., & Fisher, D. E. (2003).
857 MLANA/MART1 and SILV/PMEL17/GP100 Are Transcriptionally Regulated by MITF
858 in Melanocytes and Melanoma. *The American Journal of Pathology*, *163*(1), 333–343.
859 https://doi.org/10.1016/S0002-9440(10)63657-7

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

865 Fabian, P., Tseng, K.-C., Thiruppathy, M., Arata, C., Chen, H.-J., Smeeton, J., Nelson, N., &
866 Crump, J. G. (2022). Lifelong single-cell profiling of cranial neural crest diversification
867 in zebrafish. *Nature Communications 2022 13:1*, *13*(1), 1–13.
868 https://doi.org/10.1038/s41467-021-27594-w

869 Farnsworth, D. R., Saunders, L. M., & Miller, A. C. (2020). A single-cell transcriptome atlas
870 for zebrafish development. *Developmental Biology*, *459*(2), 100–108.
871 https://doi.org/10.1016/j.ydbio.2019.11.008

872 Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018a). Single-
873 cell reconstruction of developmental trajectories during zebrafish embryogenesis.
874 *Science*, *360*(6392). https://doi.org/10.1126/science.aar3131

875 Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018b). Single-
876 cell reconstruction of developmental trajectories during zebrafish embryogenesis.
877 *Science*, *360*(6392). https://doi.org/10.1126/science.aar3131

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

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

884 Fuiten, A. M., & Cresko, W. A. (2021). Evolutionary divergence of a Hoxa2b hindbrain
885 enhancer in syngnathids mimics results of functional assays. *Development Genes and*
886 *Evolution*, *231*(3–4), 57–71. https://doi.org/10.1007/S00427-021-00676-X/METRICS

887 Gibert, Y., Samarut, E., Ellis, M. K., Jackman, W. R., & Laudet, V. (2019). The first formed
888 tooth serves as a signalling centre to induce the formation of the dental row in
889 zebrafish. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1904).
890 https://doi.org/10.1098/rspb.2019.0401

891 Goldsmith, M. I., Fisher, S., Waterman, R., & Johnson, S. L. (2003). Saltatory control of
892 isometric growth in the zebrafish caudal fin is disrupted in long fin and rapunzel
893 mutants. *Developmental Biology*, *259*(2), 303–317. https://doi.org/10.1016/S0012-
894 1606(03)00186-6

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

898 Hain, D., Gallego-Flores, T., Klinkmann, M., Macias, A., Ciirdaeva, E., Arends, A., Thum, C.,
899 Tushev, G., Kretschmer, F., Tosches, M. A., & Laurent, G. (2022). Molecular diversity
900 and evolution of neuron types in the amniote brain. *Science*, *377*(6610).

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

905 Harlin-Cognato, A., Hoffman, E. A., & Jones, A. G. (2006). Gene cooption without duplication
906 during the evolution of a male-pregnancy gene in pipefish. *Proceedings of the National*
907 *Academy of Sciences of the United States of America*, *103*(51), 19407–19412.
908 https://doi.org/10.1073/pnas.0603000103

909 Healey, H. M., Bassham, S., & Cresko, W. A. (2022). Single-cell Iso-Sequencing enables rapid
910 genome annotation for scRNAseq analysis. *Genetics*, *220*(3).
911 https://doi.org/10.1093/GENETICS/IYAC017

912 Hong, X., Wang, Y., Wang, K., Wei, C., Li, W., Yu, L., Xu, H., Zhu, J., Zhu, X., & Liu, X. (2023).
913 Single-Cell Atlas Reveals the Hemocyte Subpopulations and Stress Responses in Asian
914 Giant Softshell Turtle during Hibernation. *Biology*, *12*(7), 994.
915 https://doi.org/10.3390/BIOLOGY12070994/S1

916 Howe, D. G., Bradford, Y. M., Conlin, T., Eagle, A. E., Fashena, D., Frazer, K., Knight, J., Mani, P.,
917 Martin, R., Moxon, S. A. T., Paddock, H., Pich, C., Ramachandran, S., Ruef, B. J., Ruzicka,
918 L., Schaper, K., Shao, X., Singer, A., Sprunger, B., … Westerfield, M. (2013). ZFIN, the
919 Zebrafish Model Organism Database: Increased support for mutants and transgenics.
920 *Nucleic Acids Research*, *41*(D1), D854–D860. https://doi.org/10.1093/NAR/GKS938

921 Hughes, L. C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C. C., Thompson, A. W., Arcila, D., Betancur,
922 R., Li, C., Becker, L., Bellora, N., Zhao, X., Li, X., Wang, M., Fang, C., Xie, B., Zhoui, Z.,
923 Huang, H., Chen, S., … Shi, Q. (2018). Comprehensive phylogeny of ray-finned fishes
924 (Actinopterygii) based on transcriptomic and genomic data. *Proceedings of the*
925 *National Academy of Sciences of the United States of America*, *115*(24), 6249–6254.

926 Jain, I., Stroka, C., Yan, J., Huang, W. M., & Iovine, M. K. (2007). Bone growth in zebrafish fins
927 occurs via multiple pulses of cell proliferation. *Developmental Dynamics*, *236*(9), 2668–
928 2674. https://doi.org/10.1002/DVDY.21270

929 Johnson, S. L., Nguyen, A. T. N., & Lister, J. A. (2011). mitfa is required at multiple stages of
930 melanocyte differentiation but not to establish the melanocyte stem cell.
931 *Developmental Biology*, *350*(2), 405–413.
932 https://doi.org/10.1016/J.YDBIO.2010.12.004

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

935 Kaneda-Nakashima, K., Igawa, K., Suwanruengsri, M., Naoyuki, F., Ichikawa, T., Funamoto,
936 T., Kurogi, S., Sekimoto, T., Yamashita, Y., Chosa, E., Yamaguchi, R., & Morishita, K.

(2022). Role of Mel1/Prdm16 in bone differentiation and morphology. *Experimental Cell Research*, *410*(2), 112969. 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*(3), 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. *G3: Genes, Genomes, Genetics*, *7*(4), 1357–1363. 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*(3), 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), 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*(6), 1725–1737. https://doi.org/10.1111/j.1095-8649.2011.02989.x

Lamason, R. L., Mohideen, M. A. P. K., Mest, J. R., Wong, A. C., Norton, H. L., Aros, M. C., Jurynec, M. J., Mao, X., Humphreville, V. R., Humbert, J. E., Sinha, S., Moore, J. L., Jagadeeswaran, P., Zhao, W., Ning, G., Makalowska, I., McKeigue, P. M., O'Donnell, D., Kittles, R., … Cheng, K. C. (2005). Genetics: SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*, *310*(5755), 1782–1786. https://doi.org/10.1126/SCIENCE.1116238/SUPPL_FILE/LAMASON.SOM.PDF

Lange, M., Granados, A., VijayKumar, S., Bragantini, J., Ancheta, S., Santhosh, S., Borja, M., Kobayashi, H., McGeever, E., Solak, A. C., Yang, B., Zhao, X., Liu, Y., Detweiler, A. M., Paul, S., Mekonen, H., Lao, T., Banks, R., Kim, Y.-J., … Royer, L. A. (2023). Zebrahub – Multimodal Zebrafish Developmental Atlas Reveals the State-Transition Dynamics of

Late-Vertebrate Pluripotent Axial Progenitors. *BioRxiv*, 2023.03.06.531398. 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), 1–13.

Leerberg, D. M., Hopton, R. E., & Draper, B. W. (2019). Fibroblast growth factor receptors function redundantly during zebrafish embryonic development. *Genetics*, *212*(4), 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*(10), 1259–1270. https://doi.org/10.1002/jmor.10982

Lin, Q., Fan, S., Zhang, Y., Xu, M., Zhang, H., Yang, Y., Lee, A. P., Woltering, J. M., Ravi, V., Gunter, H. M., Luo, W., Gao, Z., Lim, Z. W., Qin, G., Schneider, R. F., Wang, X., Xiong, P., Li, G., Wang, K., … Venkatesh, B. (2016a). The seahorse genome and the evolution of its specialized morphology. *Nature*, *540*(7633), 395–399. https://doi.org/10.1038/nature20595

Lin, Q., Fan, S., Zhang, Y., Xu, M., Zhang, H., Yang, Y., Lee, A. P., Woltering, J. M., Ravi, V., Gunter, H. M., Luo, W., Gao, Z., Lim, Z. W., Qin, G., Schneider, R. F., Wang, X., Xiong, P., Li, G., Wang, K., … Venkatesh, B. (2016b). The seahorse genome and the evolution of its specialized morphology. *Nature*, *540*(7633), 395–399. https://doi.org/10.1038/nature20595

Liu, Z., Liu, S., Yao, J., Bao, L., Zhang, J., Li, Y., Jiang, C., Sun, L., Wang, R., Zhang, Y., Zhou, T., Zeng, Q., Fu, Q., Gao, S., Li, N., Koren, S., Jiang, Y., Zimin, A., Xu, P., … Waldbieser, G. C. (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 29:2*, *29*(2), 197–221.

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*(2), 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*(5), 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*(2), 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*(5), 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 12:1*, *12*(1), 1–16. 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 7:1*, *7*(1), 1–10. https://doi.org/10.1038/ncomms10397

Morrison, J. A., McLennan, R., Wolfe, L. A., Gogol, M. M., Meier, S., McKinney, M. C., Teddy, J. M., Holmes, L., Semerad, C. L., Box, A. C., Li, H., Hall, K. E., Perera, A. G., & Kulesa, P. M. (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

1020 Parker, J., Guslund, N. C., Jentoft, S., & Roth, O. (2022). Characterization of Pipefish Immune
1021 Cell Populations Through Single-Cell Transcriptomics. *Frontiers in Immunology*, *13*,
1022 820152. https://doi.org/10.3389/FIMMU.2022.820152/BIBTEX

1023 Partridge, C., Shardo, J., & Boettcher, A. (2007). Osmoregulatory role of the brood pouch in
1024 the euryhaline Gulf pipefish, Syngnathus scovelli. *Comparative Biochemistry and*
1025 *Physiology - A Molecular and Integrative Physiology*, *147*(2 SPEC. ISS.), 556–561.
1026 https://doi.org/10.1016/j.cbpa.2007.02.007

1027 Potts, H. G., Lemieux, M. E., Rice, E. S., Warren, W., Choudhury, R. P., & Mommersteeg, M. T.
1028 M. (2022). Discordant Genome Assemblies Drastically Alter the Interpretation of
1029 Single-Cell RNA Sequencing Data Which Can Be Mitigated by a Novel Integration
1030 Method. *Cells*, *11*(4), 608. https://doi.org/10.3390/CELLS11040608/S1

1031 Qu, M., Liu, Y., Zhang, Y., Wan, S., Ravi, V., Qin, G., Jiang, H., Wang, X., Zhang, H., Zhang, B.,
1032 Gao, Z., Huysseune, A., Zhang, Z., Zhang, H., Chen, Z., Yu, H., Wu, Y., Tang, L., Li, C., … Lin,
1033 Q. (2021). Seadragon genome analysis provides insights into its phenotype and sex
1034 determination locus. *Science Advances*, *7*(34), eabg5196.
1035 https://doi.org/10.1126/SCIADV.ABG5196

1036 Ramesh, B., Small, C. M., Healey, H., Johnson, B., Barker, E., Currey, M., Bassham, S., Myers,
1037 M., Cresko, W. A., & Jones, A. G. (2023). Improvements to the Gulf pipefish Syngnathus
1038 scovelli genome. *GigaByte*, *2023*. https://doi.org/10.46471/GIGABYTE.76

1039 Ripley, J.L., Williams, P. S., & Foran, C. M. (2010). Morphological and quantitative changes in
1040 paternal brood-pouch vasculature during embryonic development in two syngnathus
1041 pipefishes. *Journal of Fish Biology*, *77*, 67–79.

1042 Ripley, Jennifer L., & Foran, C. M. (2006a). Differential parental nutrient allocation in two
1043 congeneric pipefish species (Syngnathidae: Syngnathus spp.). *Journal of Experimental*
1044 *Biology*, *209*(6), 1112–1121. https://doi.org/10.1242/jeb.02119

1045 Ripley, Jennifer L., & Foran, C. M. (2006b). Differential parental nutrient allocation in two
1046 congeneric pipefish species (Syngnathidae: Syngnathus spp.). *Journal of Experimental*
1047 *Biology*, *209*(6), 1112–1121. https://doi.org/10.1242/JEB.02119

1048 Rose, E., Johnson, B. D., Armitage, V., & Jones, A. G. (2023). The response of the brood pouch
1049 transcriptome to synthetic estrogen exposure in the Gulf pipefish (Syngnathus
1050 scovelli). *Frontiers in Marine Science*, *10*(June), 1–12.
1051 https://doi.org/10.3389/fmars.2023.1138597

1052 Roth, O., Klein, V., Beemelmanns, A., Scharsack, J. P., & Reusch, T. B. H. (2012). Male
1053 pregnancy and biparental immune priming. *American Naturalist*, *180*(6), 802–814.
1054 https://doi.org/10.1086/668081

1055 Roth, O., Solbakken, M. H., Tørresen, O. K., Bayer, T., Matschiner, M., Baalsrud, H. T., Hoff, S.
1056 N. K., Brieuc, M. S. O., Haase, D., Hanel, R., Reusch, T. B. H., & Jentoft, S. (2020a).
1057 Evolution of male pregnancy associated with remodeling of canonical vertebrate
1058 immunity in seahorses and pipefishes. *Proceedings of the National Academy of Sciences*,
1059 201916251. https://doi.org/10.1073/pnas.1916251117

1060 Roth, O., Solbakken, M. H., Tørresen, O. K., Bayer, T., Matschiner, M., Baalsrud, H. T., Hoff, S.
1061 N. K., Brieuc, M. S. O., Haase, D., Hanel, R., Reusch, T. B. H., & Jentoft, S. (2020b).
1062 Evolution of male pregnancy associated with remodeling of canonical vertebrate
1063 immunity in seahorses and pipefishes. *Proceedings of the National Academy of Sciences*
1064 *of the United States of America*, *117*(17), 9431–9439.
1065 https://doi.org/10.1073/PNAS.1916251117/-/DCSUPPLEMENTAL

1066 Royan, M. R., Siddique, K., Csucs, G., Puchades, M. A., Nourizadeh-Lillabadi, R., Bjaalie, J. G.,
1067 Henkel, C. V., Weltzien, F. A., & Fontaine, R. (2021). 3D Atlas of the Pituitary Gland of
1068 the Model Fish Medaka (Oryzias latipes). *Frontiers in Endocrinology*, *12*, 719843.
1069 https://doi.org/10.3389/FENDO.2021.719843/BIBTEX

1070 Salamanca-Díaz, D. A., Schulreich, S. M., Cole, A. G., & Wanninger, A. (2022). Single-Cell RNA
1071 Sequencing Atlas From a Bivalve Larva Enhances Classical Cell Lineage Studies.
1072 *Frontiers in Ecology and Evolution*, *9*, 783984.
1073 https://doi.org/10.3389/FEVO.2021.783984/BIBTEX

1074 Schilling, T. F., & Kimmel, C. B. (1997). Musculoskeletal patterning in the pharyngeal
1075 segments of the zebrafish embryo. *Development*, *124*(15), 2945–2960.

1076 https://doi.org/10.1242/DEV.124.15.2945

1077 Schneider, R. F., Gunter, H. M., Salewski, I., Woltering, J. M., & Meyer, A. (2023). Growth
1078 dynamics and molecular bases of evolutionary novel jaw extensions in halfbeaks and
1079 needlefishes (Beloniformes). *Molecular Ecology*, *32*(21), 5798–5811.
1080 https://doi.org/10.1111/MEC.17143

1081 Schneider, R. F., Woltering, J. M., Adriaens, D., & Roth, O. (2023). A comparative analysis of
1082 the ontogeny of syngnathids (pipefishes and seahorses) reveals how heterochrony
1083 contributed to their diversification. *Developmental Dynamics*, *252*(5), 553–588.
1084 https://doi.org/10.1002/DVDY.551

1085 Shema, E., Bernstein, B. E., & Buenrostro, J. D. (2019). Single-cell and single-molecule
1086 epigenomics to uncover genome regulation at unprecedented resolution. *Nature*
1087 *Genetics*, *51*(1), 19–25. https://doi.org/10.1038/s41588-018-0290-x

1088 Shull, Lomeli C., Lencer, E. S., Kim, H. M., Goyama, S., Kurokawa, M., Costello, J. C., Jones, K., &
1089 Artinger, K. B. (2022). PRDM paralogs antagonistically balance Wnt/β-catenin activity
1090 during craniofacial chondrocyte differentiation. *Development (Cambridge, England)*,
1091 *149*(4). https://doi.org/10.1242/DEV.200082/VIDEO-2

1092 Shull, Lomeli Carpio, Sen, R., Menzel, J., Goyama, S., Kurokawa, M., & Artinger, K. B. (2020).
1093 The conserved and divergent roles of Prdm3 and Prdm16 in zebrafish and mouse
1094 craniofacial development. *Developmental Biology*, *461*(2), 132–144.
1095 https://doi.org/10.1016/J.YDBIO.2020.02.006

1096 Sire, J. Y., Delgado, S. C., & Girondot, M. (2008). Hen's teeth with enamel cap: From dream to
1097 impossibility. *BMC Evolutionary Biology*, *8*(1), 1–11. https://doi.org/10.1186/1471-
1098 2148-8-246

1099 Small, C. M., Bassham, S., Catchen, J., Amores, A., Fuiten, A. M., Brown, R. S., Jones, A. G., &
1100 Cresko, W. A. (2016). The genome of the Gulf pipefish enables understanding of
1101 evolutionary innovations. *Genome Biology*, *17*(1), 258.
1102 https://doi.org/10.1186/s13059-016-1126-6

1103 Small, Clayton M., Harlin-Cognato, A. D., & Jones, A. G. (2013). Functional similarity and
1104 molecular divergence of a novel reproductive transcriptome in two male-pregnant
1105 Syngnathus pipefish species. *Ecology and Evolution*, *3*(12), 4092–4108.
1106 https://doi.org/10.1002/ECE3.763

1107 Small, Clayton M., Healey, H. M., Currey, M. C., Beck, E. A., Catchen, J., Lin, A. S. P., Cresko, W.
1108 A., & Bassham, S. (2022). Leafy and weedy seadragon genomes connect genic and
1109 repetitive DNA features to the extravagant biology of syngnathid fishes. *Proceedings of*
1110 *the National Academy of Sciences of the United States of America*, *119*(26),
1111 e2119602119.
1112 https://doi.org/10.1073/PNAS.2119602119/SUPPL_FILE/PNAS.2119602119.SD05.X
1113 LSX

1114 Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L.,
1115 Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino, M. L.,
1116 Boismoreau, F., Hilscher, M. M., Yokota, C., Qian, X., Nilsson, M., Bronner, M. E., …
1117 Adameyko, I. (2019). Spatiotemporal structure of cell fate decisions in murine neural
1118 crest. *Science*, *364*(6444). https://doi.org/10.1126/science.aas9536

1119 Sommer, S., Whittington, C. M., & Wilson, A. B. (2012). Standardised classification of pre-
1120 release development in male-brooding pipefish, seahorses, and seadragons (Family
1121 Syngnathidae). *BMC Developmental Biology*, *12*, 12–15. https://doi.org/10.1186/1471-
1122 213X-12-39

1123 Songco-Casey, J. O., Coffing, G. C., Piscopo, D. M., Pungor, J. R., Kern, A. D., Miller, A. C., & Niell,
1124 C. M. (2022). Cell types and molecular architecture of the Octopus bimaculoides visual
1125 system. *Current Biology : CB*, *32*(23), 5031.
1126 https://doi.org/10.1016/J.CUB.2022.10.015

1127 Steger, J., Cole, A. G., Denner, A., Lebedeva, T., Genikhovich, G., Ries, A., Reischl, R., Taudes, E.,
1128 Lassnig, M., & Technau, U. (2022). Single-cell transcriptomics identifies conserved
1129 regulators of neuroglandular lineages. *Cell Reports*, *40*(12).
1130 https://doi.org/10.1016/J.CELREP.2022.111370

1131 Stiller, J., Short, G., Hamilton, H., Saarman, N., Longo, S., Wainwright, P., Rouse, G. W., &
1132 Simison, W. B. (2022). Phylogenomic analysis of Syngnathidae reveals novel
1133 relationships, origins of endemic diversity and variable diversification rates. *BMC*
1134 *Biology 2022 20:1*, *20*(1), 1–21. https://doi.org/10.1186/S12915-022-01271-W

1135 Stock, D. W., Jackman, W. R., & Trapani, J. (2006). Developmental genetic mechanisms of
1136 evolutionary tooth loss in cypriniform fishes. *Development*, *133*(16), 3127–3137.
1137 https://doi.org/10.1242/DEV.02459

1138 Stölting, K. N., & Wilson, A. B. (2007). Male pregnancy in seahorses and pipefish: Beyond
1139 the mammalian model. *BioEssays*, *29*(9), 884–896.
1140 https://doi.org/10.1002/bies.20626

1141 Swartz, M. E., Sheehan-Rooney, K., Dixon, M. J., & Eberhart, J. K. (2011). Examination of a
1142 palatogenic gene program in zebrafish. *Developmental Dynamics*, *240*(9), 2204–2220.
1143 https://doi.org/10.1002/DVDY.22713

1144 Szabo-Rogers, H. L., Geetha-Loganathan, P., Nimmagadda, S., Fu, K. K., & Richman, J. M.
1145 (2008). FGF signals from the nasal pit are necessary for normal facial morphogenesis.
1146 *Developmental Biology*, *318*(2), 289–302.
1147 https://doi.org/10.1016/j.ydbio.2008.03.027

1148 Tardaguila, M., Fuente, L. de la, Marti, C., Pereira, C., Pardo-Palacios, F. J., Risco, H. del,
1149 Ferrell, M., Mellado, M., Macchietto, M., Verheggen, K., Edelmann, M., Ezkurdia, I.,
1150 Vazquez, J., Tress, M., Mortazavi, A., Martens, L., Rodriguez-Navarro, S., Moreno-
1151 Manzano, V., & Conesa, A. (2018). SQANTI: Extensive characterization of long-read
1152 transcript sequences for quality control in full-length transcriptome identification and
1153 quantification. *Genome Research*, *28*(3), 396. https://doi.org/10.1101/GR.222976.117

1154 Teichmann, S. A., & Babu, M. M. (2004). Gene regulatory network growth by duplication.
1155 *Nature Genetics*, *36*(5), 492–496. https://doi.org/10.1038/ng1340

1156 Tengfei, D., Yongshuang, X., Haixia, Z., Li, Z., Qinghua, L., Xueying, W., Jun, L., Shihong, X.,
1157 Yanfeng, W., Jiachen, Y., Lele, W., Yunong, W., & Guang, G. (2021). Multiple Fetal
1158 Nutritional Patterns Before Parturition in Viviparous Fish Sebastes schlegelii

(Hilgendorf, 1880). *Frontiers in Marine Science*, *7*, 571946. https://doi.org/10.3389/FMARS.2020.571946/BIBTEX

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 3:1*, *3*(1), 59–69. 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*(1), 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* (Vol. 64, pp. 66–71). Elsevier Ltd. https://doi.org/10.1016/j.gde.2020.05.033

Ton, Mai Linh Nu, Keitley, D., Theeuwes, B., Guibentif, C., Ahnfelt-Rønne, J., Andreassen, T. K., Calero-Nieto, F. J., Imaz-Rosshandler, I., Pijuan-Sala, B., Nichols, J., Benito-Gutiérrez, È., Marioni, J. C., & Göttgens, B. (2023). An atlas of rabbit development as a model for single-cell comparative genomics. *Nature Cell Biology 2023 25:7*, *25*(7), 1061–1072. https://doi.org/10.1038/s41556-023-01174-0

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*(21), 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 5:7*, *5*(7), 499–508. https://doi.org/10.1038/nrg1380

1186 Umemori, H., Linhoff, M. W., Ornitz, D. M., & Sanes, J. R. (2004). FGF22 and its close relatives
1187 are presynaptic organizing molecules in the mammalian brain. *Cell*, *118*(2), 257–270.
1188 https://doi.org/10.1016/j.cell.2004.06.025

1189 Van Wassenbergh, S., Roos, G., Genbrugge, A., Leysen, H., Aerts, P., Adriaens, D., & Herrel, A.
1190 (2009). Suction is kid's play: extremely fast suction in newborn seahorses. *Biology*
1191 *Letters*, *5*(2), 200–203. https://doi.org/10.1098/rsbl.2008.0765

1192 Van Wassenbergh, S., Strother, J. A., Flammang, B. E., Ferry-Graham, L. A., & Aerts, P. (2007).
1193 Extremely fast prey capture in pipefish is powered by elastic recoil. *Journal of The*
1194 *Royal Society Interface*, *5*(20), 285–296. https://doi.org/10.1098/RSIF.2007.1124

1195 Verreijdt, L., Debiais-Thibaud, M., Borday-Birraux, V., Van Der Heyden, C., Sire, J. Y., &
1196 Huysseune, A. (2006). Expression of the dlx gene family during formation of the
1197 cranial bones in the zebrafish (Danio rerio): Differential involvement in the visceral
1198 skeleton and braincase. *Developmental Dynamics*, *235*(5), 1371–1389.
1199 https://doi.org/10.1002/DVDY.20734

1200 Vonk, A. C., Zhao, X., Pan, Z., Hudnall, M. L., Oakes, C. G., Lopez, G. A., Hasel-Kolossa, S. C.,
1201 Kuncz, A. W. C., Sengelmann, S. B., Gamble, D. J., & Lozito, T. P. (2023). Single-cell
1202 analysis of lizard blastema fibroblasts reveals phagocyte-dependent activation of
1203 Hedgehog-responsive chondrogenesis. *Nature Communications 2023 14:1*, *14*(1), 1–20.
1204 https://doi.org/10.1038/s41467-023-40206-z

1205 Wagner, A. (2011). The molecular origins of evolutionary innovations. *Trends in Genetics*,
1206 *27*(10), 397–410. https://doi.org/10.1016/J.TIG.2011.06.002

1207 Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., & Klein, A. M. (2018).
1208 Single-cell mapping of gene expression landscapes and lineage in the zebrafish
1209 embryo. *Science*, *360*(6392), 981–987. https://doi.org/10.1126/science.aar4362

1210 Wagner, G. P., & Lynch, V. J. (2010). Evolutionary novelties. In *Current Biology* (Vol. 20,
1211 Issue 2, pp. R48–R52). Cell Press. https://doi.org/10.1016/j.cub.2009.11.010

1212 Walshe, J., & Mason, I. (2003). Fgf signalling is required for formation of cartilage in the

head. *Developmental Biology*, *264*(2), 522–536. https://doi.org/10.1016/j.ydbio.2003.08.010

Wang, W., Yang, N., Wang, L., Zhu, Y., Chu, X., Xu, W., Li, Y., Xu, Y., Gao, L., Zhang, B., Zhang, G., Sun, Q., Wang, W., Wang, Q., Zhang, W., & Chen, D. (2024). The TET-Sall4-BMP regulatory axis controls craniofacial cartilage development. *Cell Reports*, *43*(3), 113873. 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*(1), 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*(5), 1252–1272. https://doi.org/10.1111/brv.12607

Whittington, C. M., Griffith, O. W., Qi, W., Thompson, M. B., & Wilson, A. B. (2015a). Seahorse Brood Pouch Transcriptome Reveals Common Genes Associated with Vertebrate Pregnancy. *Molecular Biology and Evolution*, *32*(12), msv177. https://doi.org/10.1093/molbev/msv177

Whittington, C. M., Griffith, O. W., Qi, W., Thompson, M. B., & Wilson, A. B. (2015b). Seahorse Brood Pouch Transcriptome Reveals Common Genes Associated with Vertebrate Pregnancy. *Molecular Biology and Evolution*, *32*(12), 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*(2), 255-276.e7. https://doi.org/10.1016/j.devcel.2019.10.003

Wolf, M., Ferrette, B. L. da S., Coimbra, R. T. F., Jong, M. de, Nebenführ, M., Prochotta, D., Schöneberg, Y., Zapf, K., Rosenbaum, J., Intyre, H. A. M., Maier, J., Souza, C. C. S. de, Gehlhaar, L. M., Werner, M. J., Oechler, H., Wittekind, M., Sonnewald, M., Nilsson, M. A., Janke, A., & Winter, S. (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* (Vol. 10, Issue 1, pp. 1–21). BioMed Central Ltd. https://doi.org/10.1186/s13227-019-0131-8

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

Woych, J., Gurrola, A. O., Deryckere, A., Jaeger, E. C. B., Gumnit, E., Merello, G., Gu, J., Araus, A. J., Leigh, N. D., Yun, M., Simon, A., & Tosches, M. A. (2022). Cell-type profiling in salamanders identifies innovations in vertebrate forebrain evolution. *Science*, *377*(6610). https://doi.org/10.1126/SCIENCE.ABP9186/SUPPL_FILE/SCIENCE.ABP9186_MDAR_REPRODUCIBILITY_CHECKLIST.PDF

Young, M. D., & Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *GigaScience*, *9*(12), 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. (2020a). Comparative genomics reveal shared genomic changes in syngnathid fishes and signatures of genetic convergence with placental mammals ! Phylogenomics and evolutionary rate Transposable elements in syngnathids Expansion and contraction of gene families in t. *National Science Review*, *7*(6), 964–977. https://doi.org/10.1002/0471250953.bi0403s18.56.

Zhang, Y., Ravi, V., Qin, G., Dai, H., Zhang, H., Han, F., Wang, X., & Liu, Y. (2020b). *Comparative genomics reveal shared genomic changes in syngnathid fishes and signatures of genetic convergence with placental mammals ! Phylogenomics and evolutionary rate Transposable elements in syngnathids Expansion and contraction of gene families in t*. 56–58. https://doi.org/10.1002/0471250953.bi0403s18.56.

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*(23), 4709–4718. https://doi.org/10.1242/DEV.099028

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

Figures

March 11, 2024

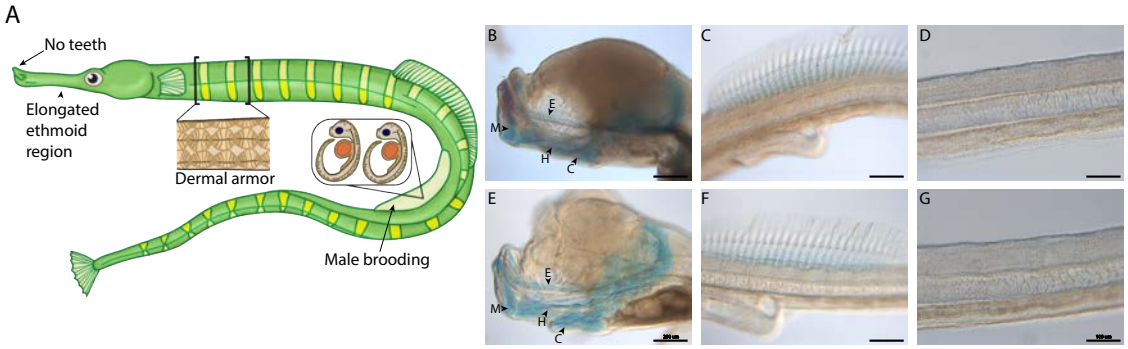

Figure 1: *Gulf pipefish exemplify syngnathid derived traits.* Gulf pipefish have elongate snouts, have lost teeth on their oral and pharyngeal jaws, possess dermal armor, and have brood pouches in males (panel A). Cartilage (alcian) and bone (alizarin) stained clutch siblings of embryos from the two 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.

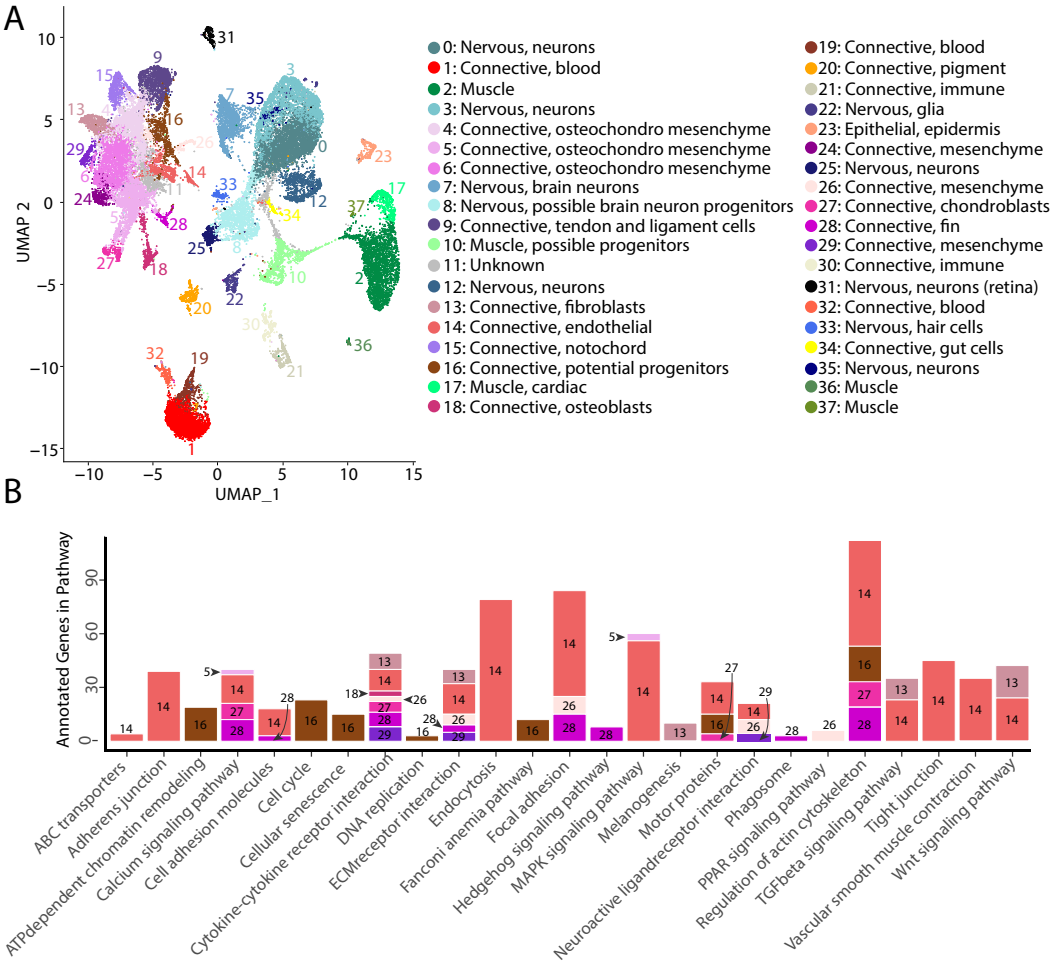

Figure 2: *Gulf pipefish single cell atlas contains cells from the entire embryo and identifies genetic pathways active in different cell types.* The UMAP plot (panel A) shows all of the cell clusters and their identities reduced to the first two UMAP dimensions. The graph in panel B displays results 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.

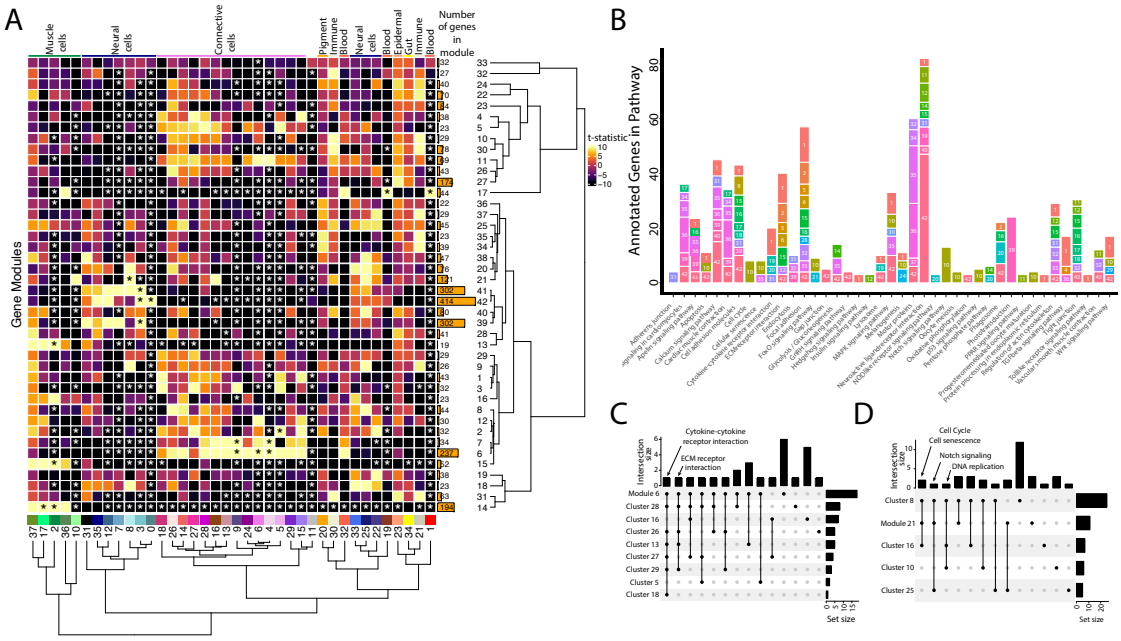

Figure 3: *Weighted Gene Network Analysis (WGCNA) identifies gene modules that define and unite cell clusters.* A) The strength of association between the gene modules and cell clusters is shown in panel A with dendrogram clustering illustrating the distance between modules and cell clusters. Gene modules 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.

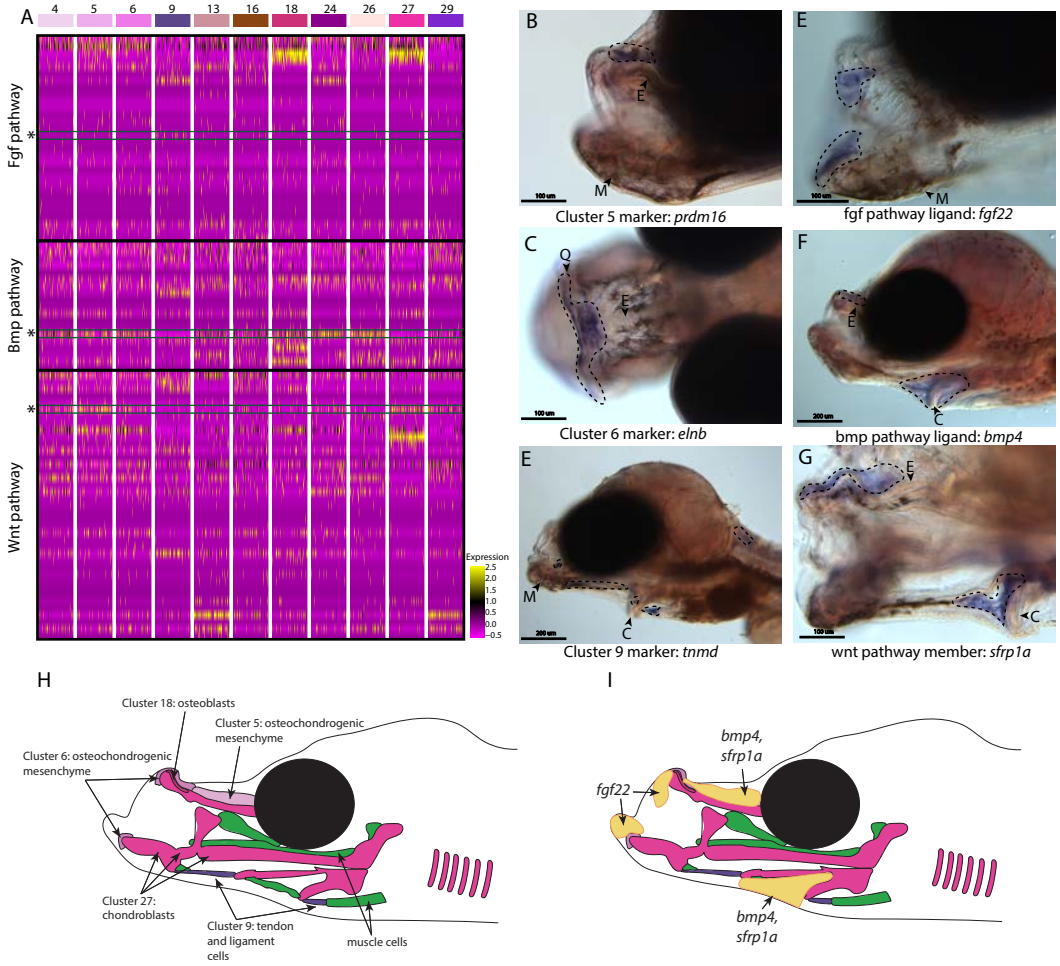

Figure 4: *Conserved cell types and gene pathways build unique faces of syngnathids.* Three main conserved signaling pathways are enriched in connective cell types, MAPK signaling (including Fgf signaling), TGF-beta signaling (including BMP signaling), and Wnt signaling. Receptors and ligands expression 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.

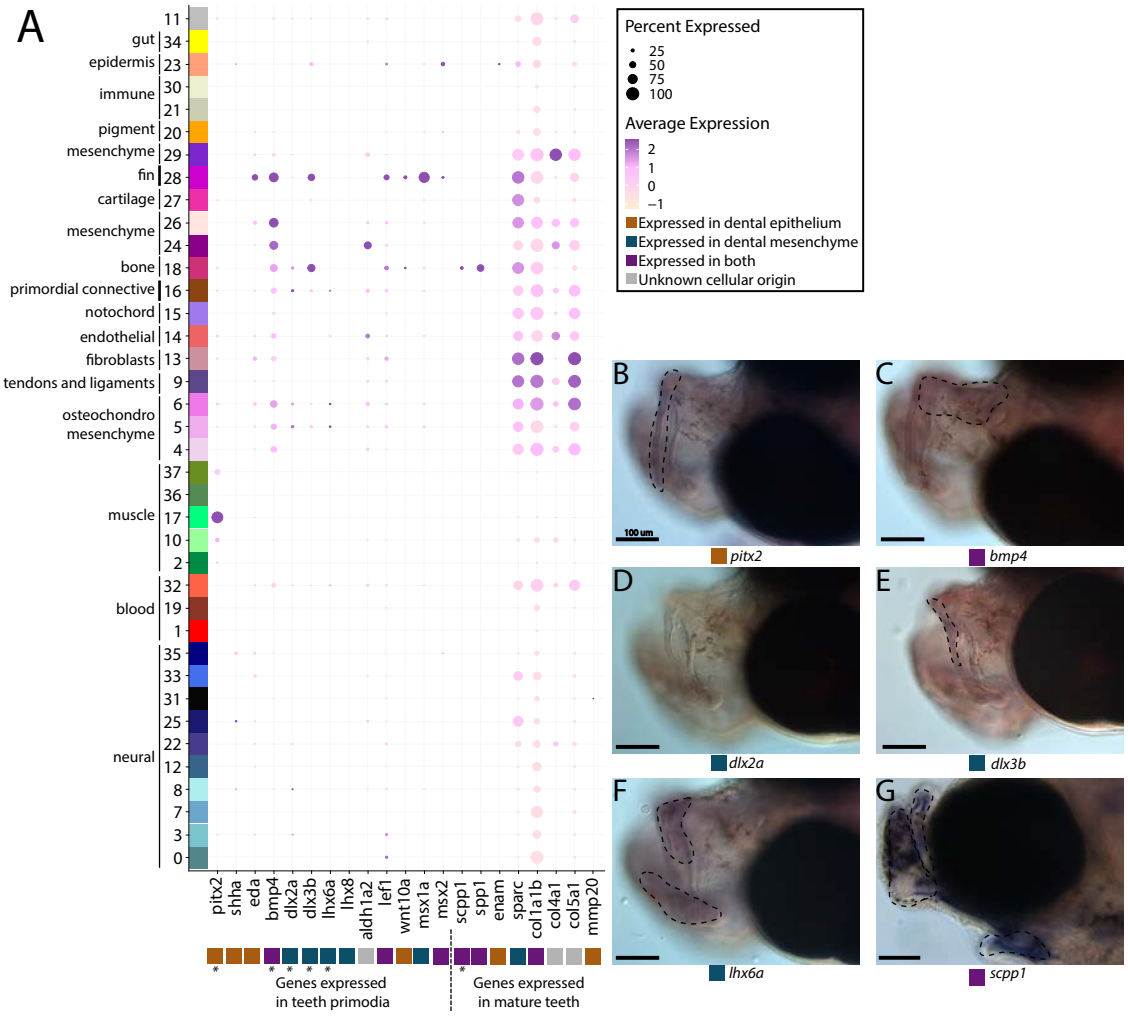

Figure 5: *Pipefish do not possess identifiable tooth primordium cells, but continue to express tooth development genes in other contexts.* Panel A presents a dotplot of genes classified by the tissue layer in which they are reported to be expressed during tooth development in other vertebrate (Tucker 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.

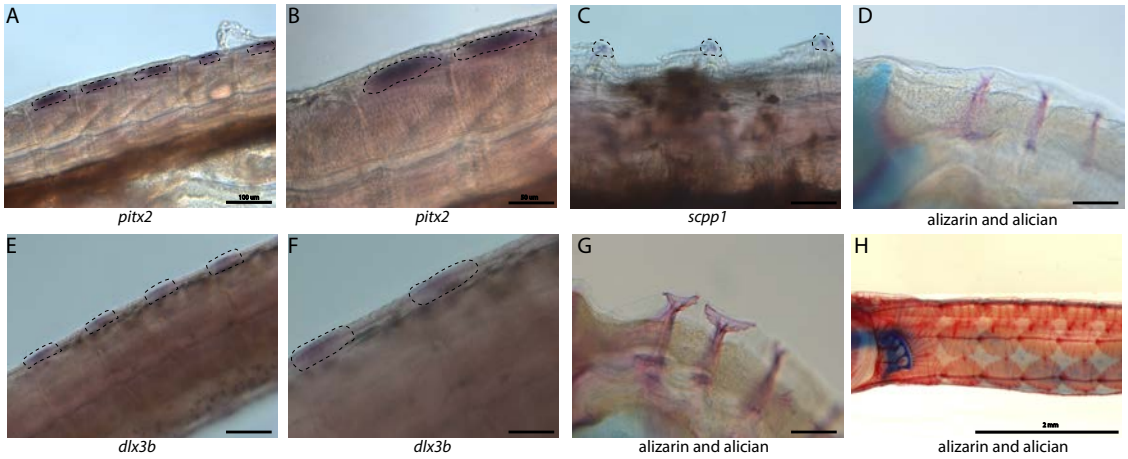

Figure 6: *Tooth and bone development genes expressed during exoskeleton development.* We discovered *pitx2* (A1 and A2) and *dlx3b* (B1 and B2) expression during possible emergence of exoskeletal primordium in wild caught bay pipefish. The embryos used for these *in situs* were the same stage as those 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.

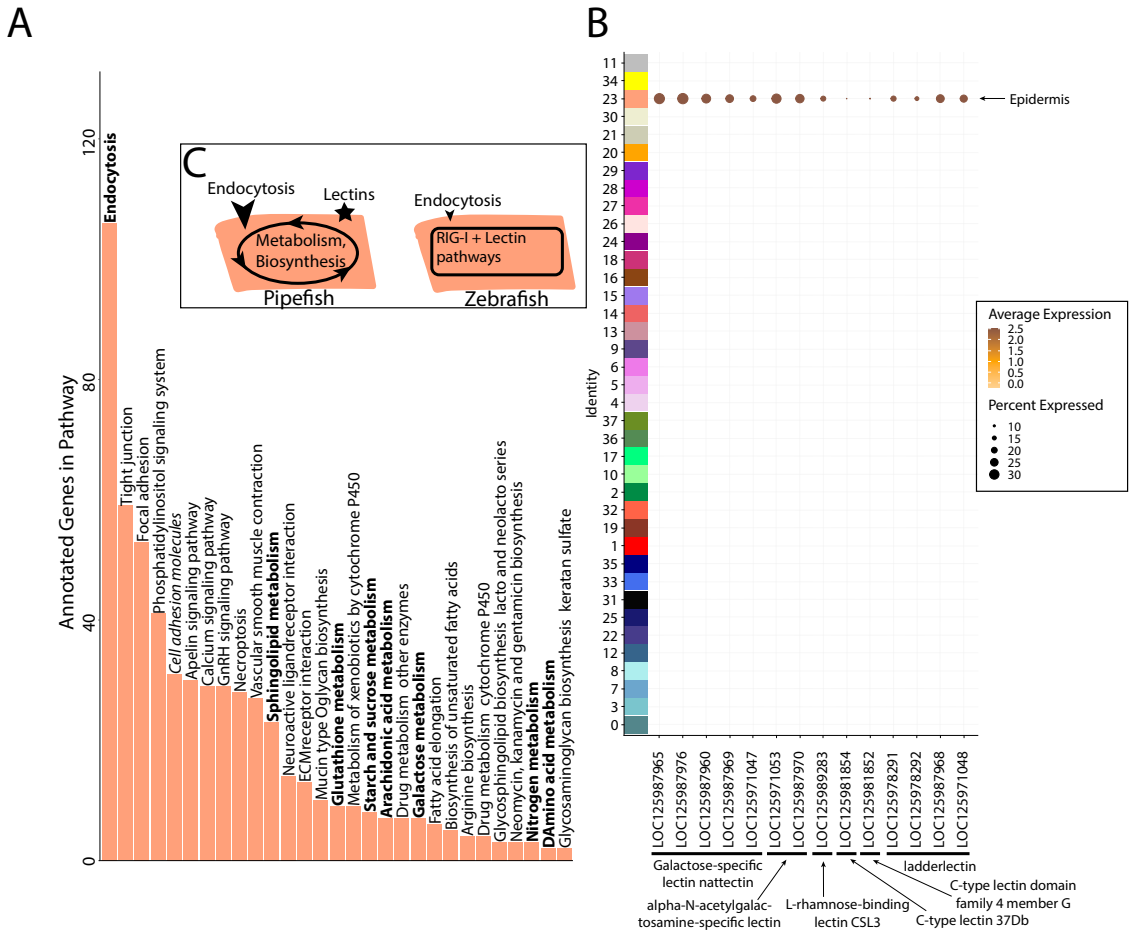

Figure 7: *Gene expression signatures suggest embryonic interactions within brood pouch environment.* Epidermal cells (panel A), with pathways that suggest increased endocytosis and metabolism pathways are in bold text. Pathways upregulated in 3dpf zebrafish epidermal cells are in italics. Pipefish epidermal 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.

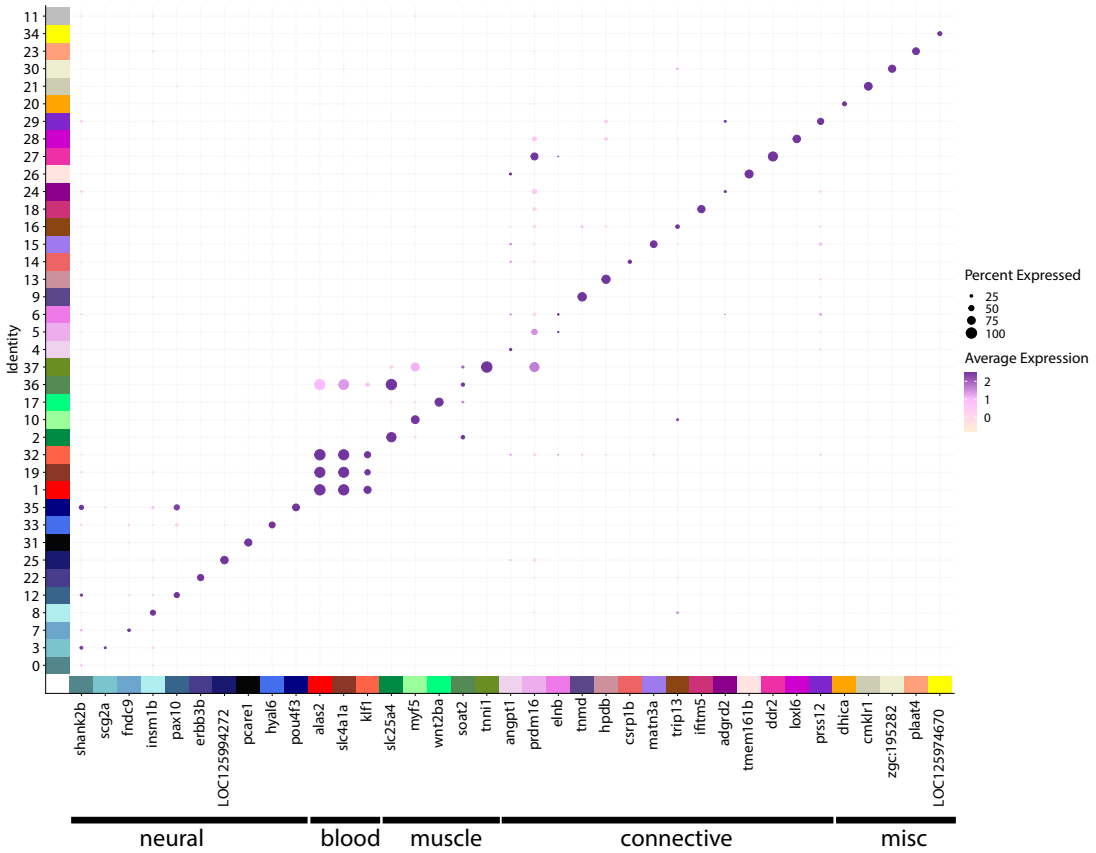

Figure S1: *Marker genes define cell identity and are a resource for cell cluster exploration.* This dotplot shows the marker genes on the x-axis, and the cell cluster on the y-axis. The size of each dot indicates the percentage of cells that express the gene in each cluster. The darkness of each dot represents the average expression of the gene.

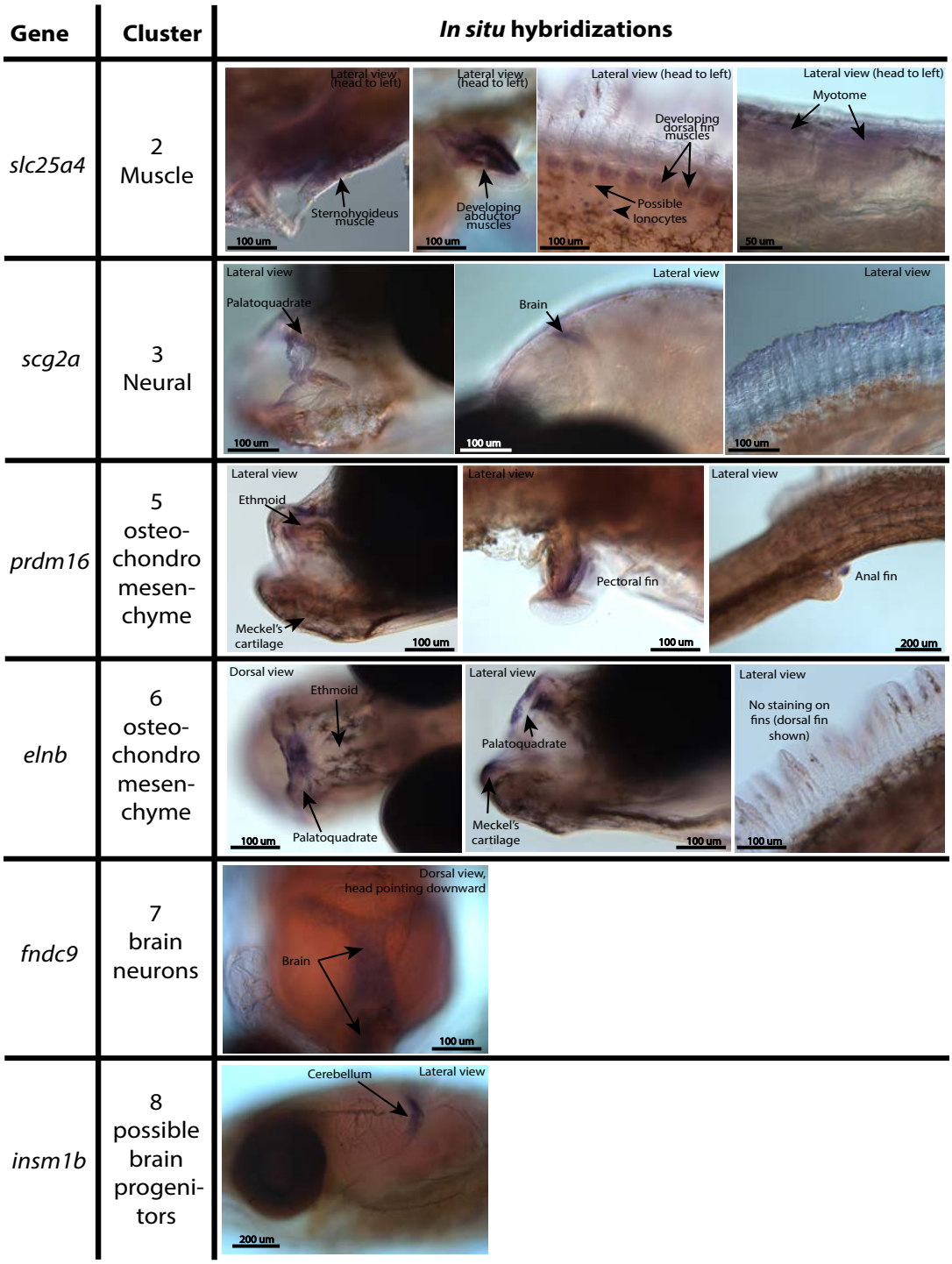

Figure S2: *Marker genes examined through in situ hybridization reveal spatial patterning of cell types.*Informative marker gene *in situ* hybridization results, including *slc25a4*, *scg2a*, *prdm16*, *elnb*, *fndc9*, and *insm1b*. Embryos are staged to highlight staining of each gene. Relavant anatomical 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.

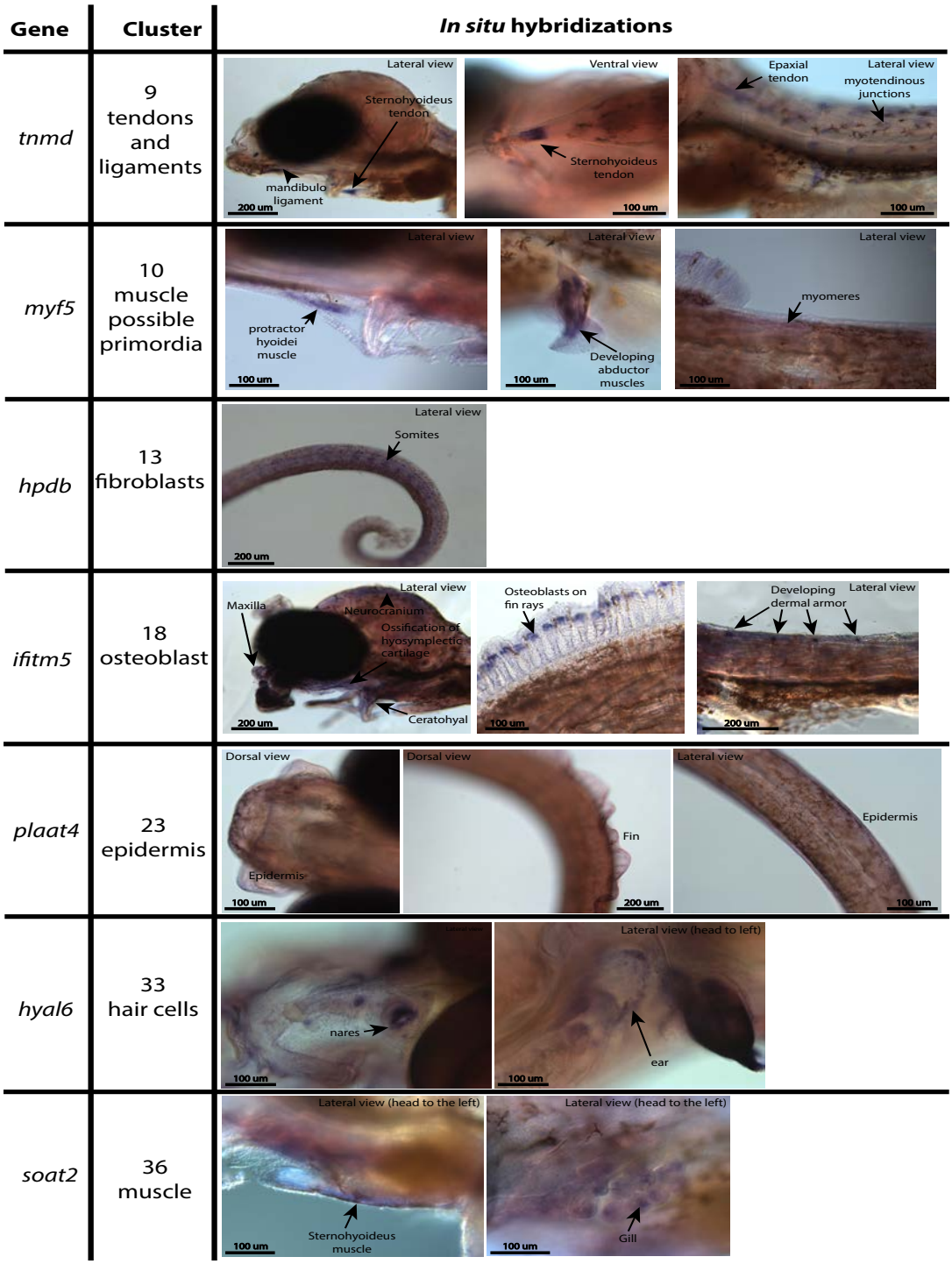

Figure S3: *Marker genes examined through in situ hybridization reveal spatial patterning of cell types.* Informative marker gene *in situ* hybridization results, including *tnmd*, *myf5*, *hpbd*, *ifitm5*, *plaat4*, *hyal6*, and *soat2*. Embryos are staged to highlight staining of each gene. Relavant 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.

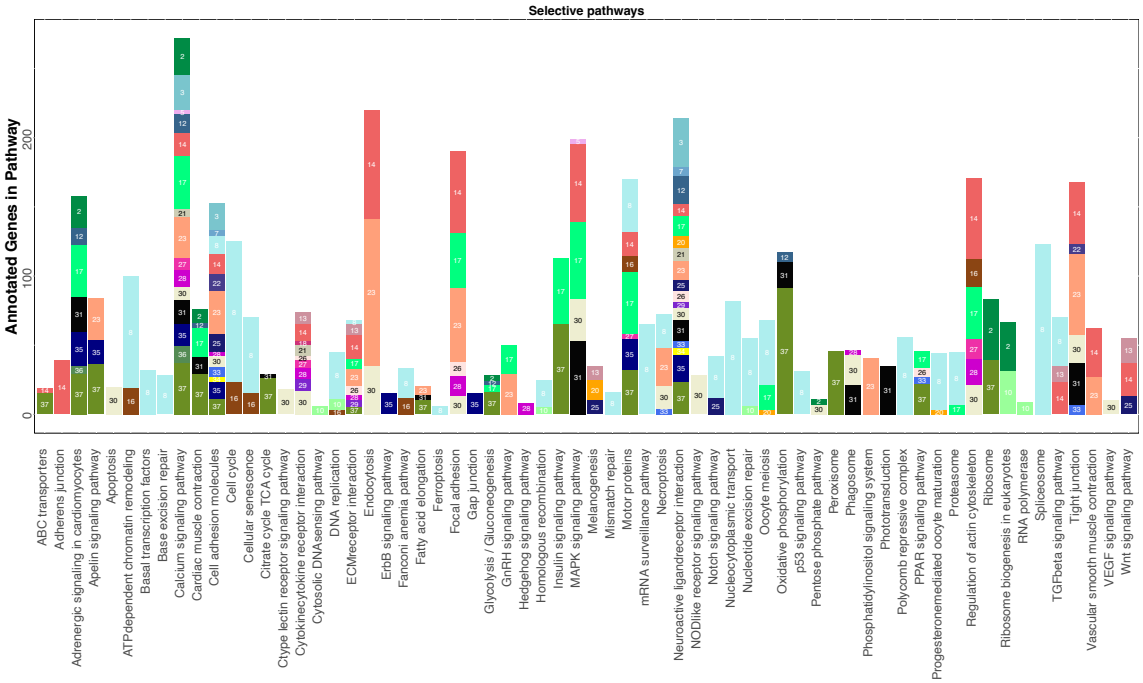

Figure S4: *KEGG analysis reveals pathways signficiantly enriched in cell clusters.* KEGG pathways are listed on the x-axis. For each cluster, the number of genes that belong to the specific KEGG pathway is represented by the height of each bar. Cell clusters are labeled by both color and their designated 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.

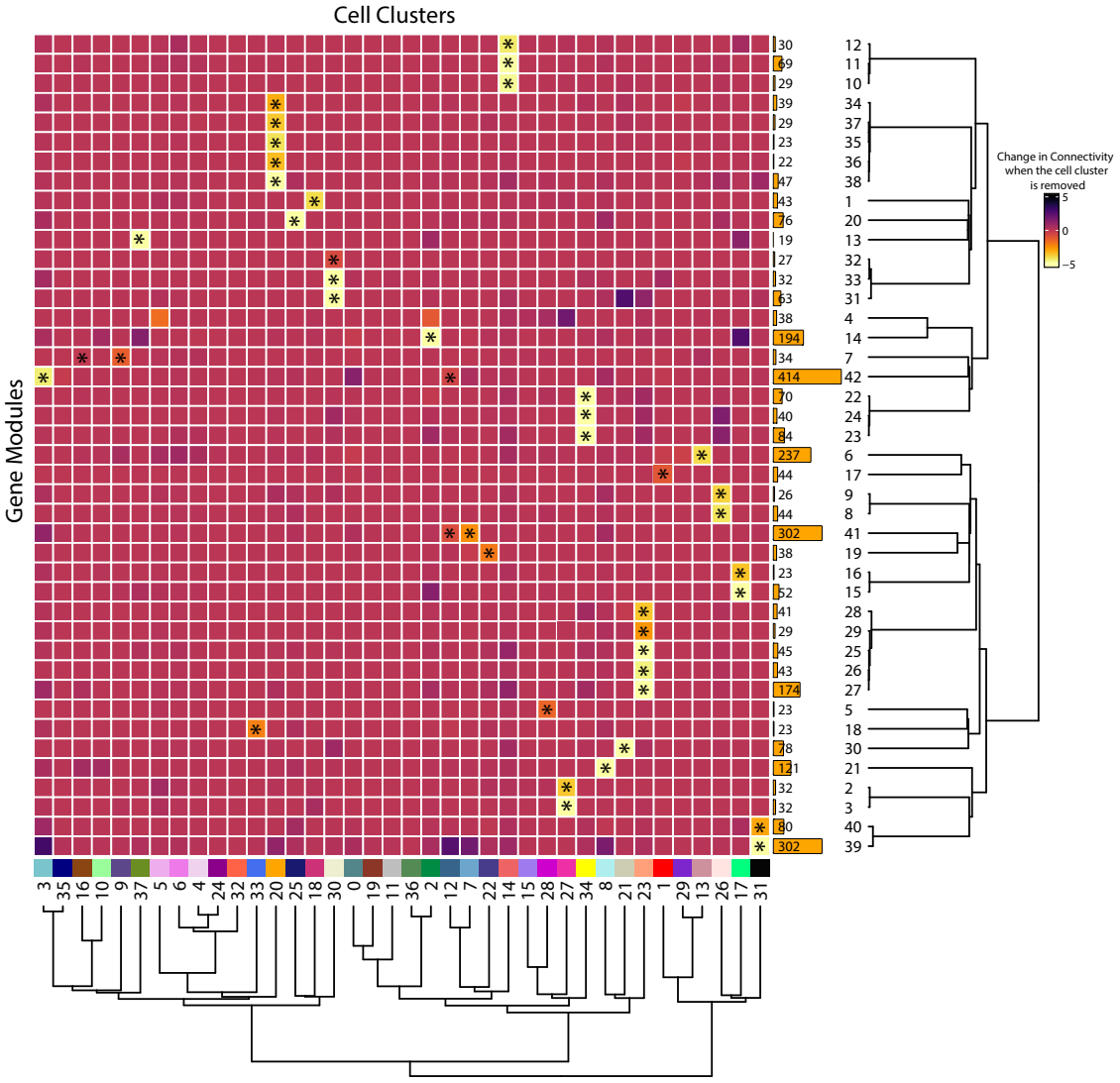

Figure S5: *Cell clusters drive gene module connectivity.* The heatmap shows the change in gene network connectivity when individual cell clusters are removed. Cell clusters are in the columns and gene modules are in rows. The coloration scale is based on the change in connecitvity when a cell cluster 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.

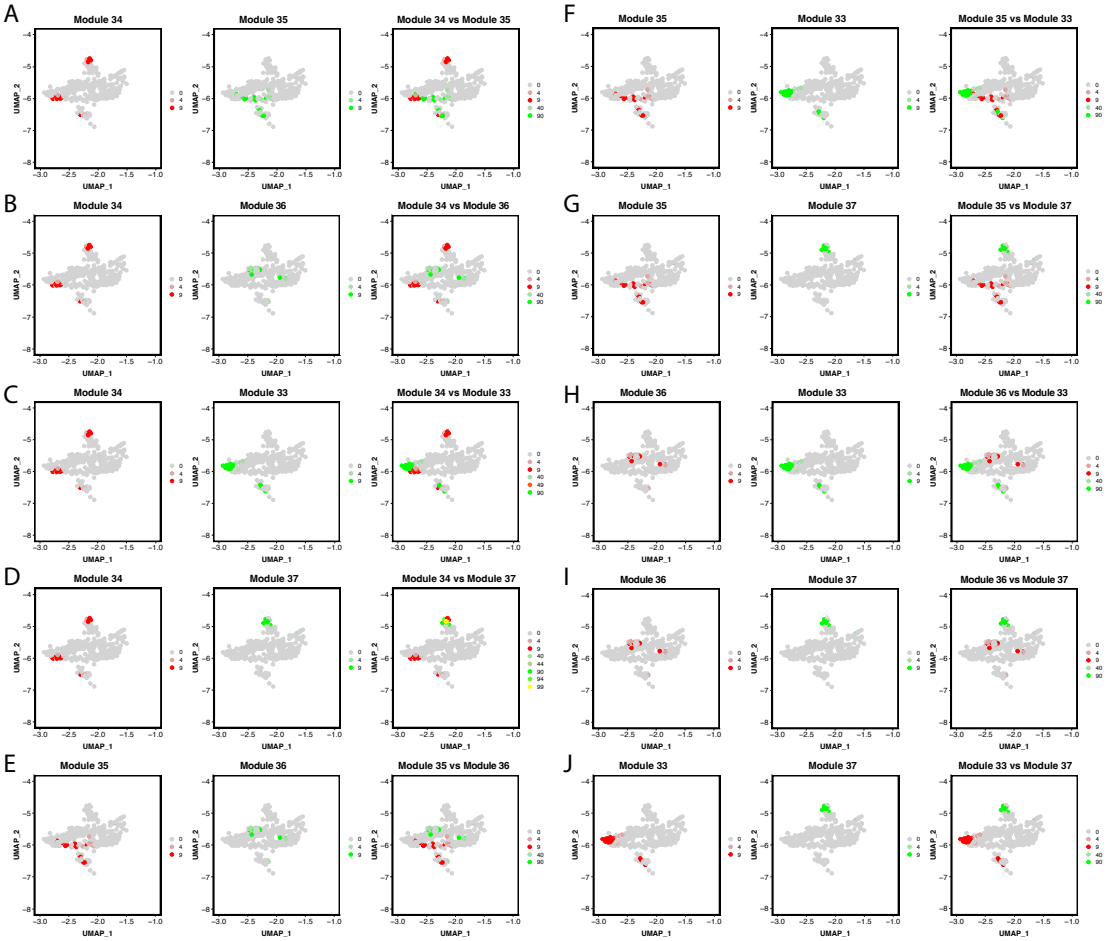

Figure S6: *Gene module expression comparison with the Pigment Cluster reveals potential elusive cell types within the cluster.* Genes within the network are combined into one feature then plotted on a feature plot. In this case, UMAP dimensions one and two are the x and y-axis accordingly, dots represent 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.

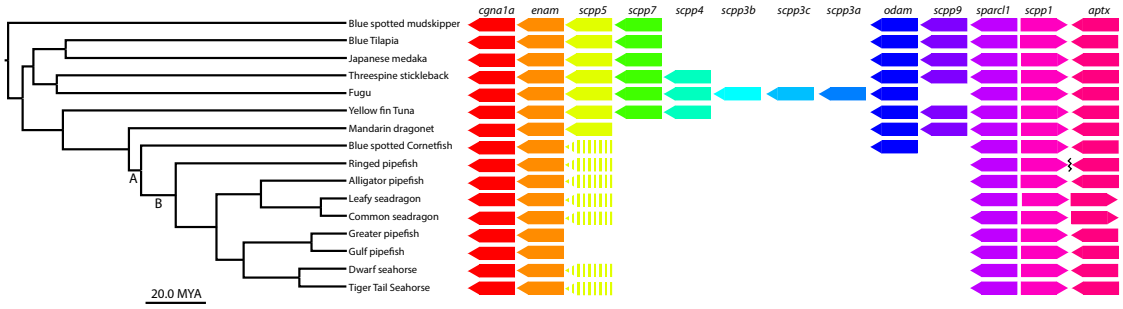

Figure S7: *Scpp gene cluster analysis identifies most Scpp losses are not unique to syngnathids.* A) denotes the Syngnathoidei lineage which contains syngnathids and blue spotted cornetfish (among other fish) to highlight that this clade contains various tooth alterations. Syngnathids lack teeth completely 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.

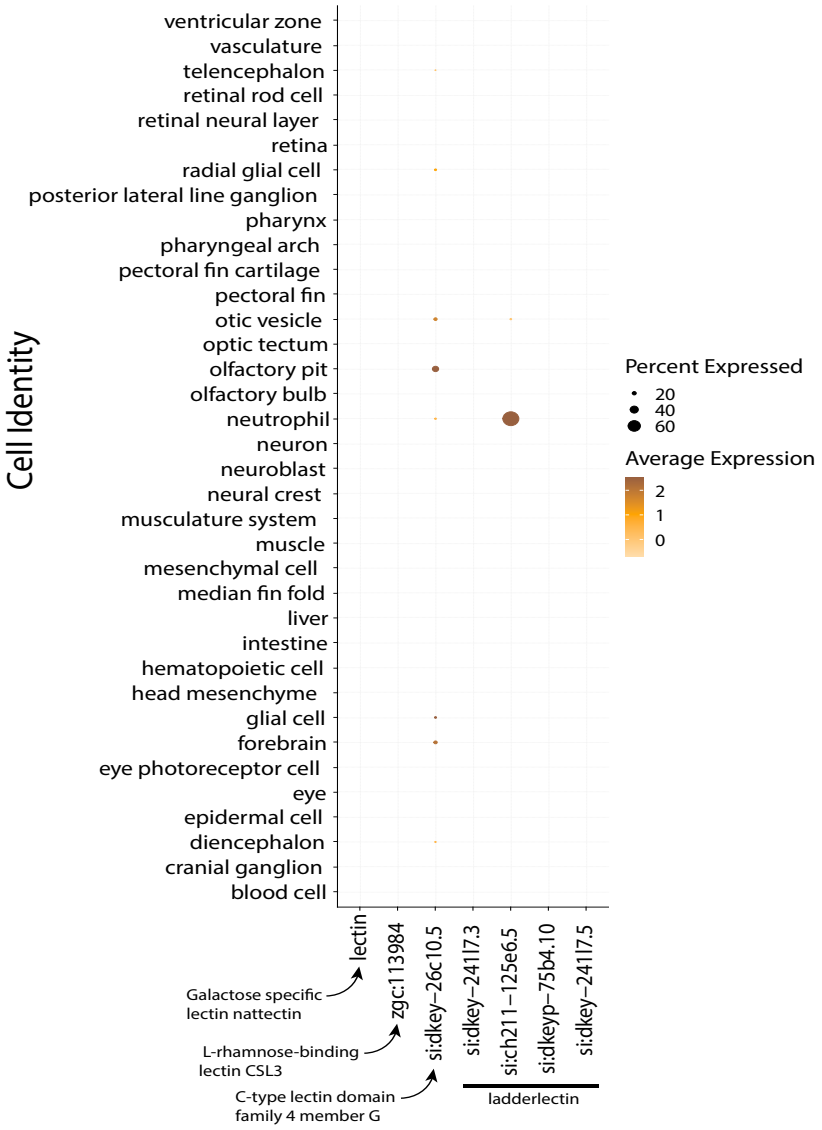

Figure S8: *3dpf Zebrafish do not express lectin genes in epidermal cells.* This dotplot shows the marker genes on the x-axis and y-axis, organized by cell type. The size of the dot represents the percentage of cells that express the gene in each cluster. The darkness of the dot represents the average expression of the genes.