Title: Single Cell Sequencing Provides Clues about the Developmental Genetic Basis of Evolutionary Adaptations in Syngnathid Fishes

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 genome assemblies 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
79 Scpp genes), fin development (*tbx4*), and immune function (MHC pathway components)
80 (Lin et al., 2016; Qu et al., 2021; Small et al., 2016; Small et al., 2022; Zhang et al., 2020).
81 Though these gene losses are highly suggestive of leading to unique changes, exploration of
82 the 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' derived traits. The Gulf pipefish (*Syngnathus scovelli*) is
86 an attractive model for this study (Figure 1A). This species has a high-quality reference
87 genome annotated by NCBI and is amenable to laboratory culture (Anderson & Jones, 2019;
88 Ramesh et al., 2023). Furthermore, species from the *Syngnathus* genus are used worldwide
89 to address questions about syngnathid evolution in microbial, developmental, functional
90 morphological, 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, 2006; 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

94 syngnathid traits, including their elongated head, toothlessness, dermal armor, and
95 development of embryos inside the brood pouch. These traits represent the diversity of
96 evolutionary changes observed in the syngnathid clade (highly altered, lost, and novel
97 traits); hypotheses from studies of model organisms suggest developmental pathways
98 involved in their evolution (Lin et al., 2016; Roth et al., 2020; Small et al., 2016; Small et al.,
99 2022).

100 scRNAseq atlases are a powerful complement to genomic analyses (Shema et al.,
101 2019; Ton et al., 2020). They can provide crucial insights into types of cells present, genes
102 that distinguish cell types (marker genes), active gene networks, and a means to identify
103 expression of genes of interest within predicted cell types (Farnsworth et al., 2020; Farrell
104 et al., 2018; Williams et al., 2019). Specifically, scRNAseq captures RNA expression profiles
105 from individual cells, allowing cell types to be inferred *post-hoc*. scRNAseq has successfully
106 been applied to syngnathid adult kidneys (Parker et al., 2022), but there are no published
107 syngnathid developmental atlases.

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

123 # Results

124 ## Valuable scRNAseq atlas for studying syngnathid development

125 We produced the first developmental scRNAseq atlas for a syngnathid from two
126 samples comprising 20 similarly staged embryos from pregnant, wild-caught Gulf pipefish
127 (*Syngnathus scovelli*) males. The samples represent a late organogenesis developmental
128 stage (Figure 1B-H). These embryos had a primarily cartilaginous skeleton with minimal
129 mineralization, including jaw cartilages that were at the onset of mineralization and
130 ethmoid elongation. The embryos also possessed cartilaginous dorsal fin pterygiophores
131 but had no signs of dermal armor mineralization. This stage is referred to as "frontal jaws"
132 in the literature on syngnathids (Sommer et al., 2012).

133 The atlas included 35,785 cells (19,892 and 15,893 cells from each sample; S1:S2),
134 which formed 38 cell clusters (Figure 2A, Supplementary File 1, 2). We classified cells into
135 4 different broad tissue types – epithelial, connective, neural, and muscle – using Seurat-
136 identified marker genes and published model organism resources (S3:S57). We next used
137 Seurat identified marker genes to pinpoint single marker genes that were most unique to
138 each cluster (Figure S58, Supplementary File 3, 4, 5). We completed *in situ* hybridization
139 using Gulf and bay pipefish embryos for cell clusters for which examining gene expression
140 would help hone and validate cluster annotations (Figure S59:S71).

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). The number of recovered cells per identity may not
148 necessarily represent organismal cellular proportions because of potential variability in
149 dissociation success for different cell types (Denisenko et al., 2020; Venema et al., 2022).

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

151 To affirm identities and discover potential properties of each cluster, we completed
152 a KEGG pathway analysis for each cluster using Seurat's marker genes (Figure S72, Figure
153 2B). For eight of the clusters (1, 4, 6, 9, 11, 15, 19, and 24), we did not find any significantly
154 enriched pathways, possibly due to similar gene expression profiles across cell types that
155 reduced the number of identified markers. However, we found one or more significantly
156 enriched pathways for the other 29 cell clusters. We observed enriched pathway terms that
157 supported cluster annotations. For example, 'phototransduction' in the retina cluster,
158 'melanogenesis' in the pigment cluster, 'cardiac muscle contraction' in muscle clusters, and
159 'neuroactive ligand receptor interaction' in neuronal clusters.

160 The inferred KEGG pathways demonstrated some commonalities across the
161 different tissue types, including in signaling pathways and cell states. Notably, our
162 identified KEGG terms delineated progenitor and differentiated cell clusters. Based on their
163 KEGG terms, we classified clusters 8, 10, and 16 as possible neural, muscle, and connective
164 tissue progenitor cells, respectively. We also detected expression of *pax3a* and *pax3b*,
165 muscle primordia markers, in cluster 10, supporting this annotation. These clusters had
166 enriched KEGG terms associated with cell division ('cell cycle', 'DNA replication',
167 'nucleotide excision repair', and 'homologous recombination'), and lacked enrichment for
168 KEGG pathways present with differentiated cell types of their lineage. Specifically, cluster 8
169 lacked the neural KEGG term 'neuroactive ligand receptor interaction', cluster 10 lacked
170 muscle KEGG terms 'adrenergic signaling in cardiomyocytes', 'calcium signaling pathways',
171 and 'cardiac muscle contraction', and cluster 16 lacked connective tissue term 'ECM
172 receptor interaction'. To complement these findings, we completed a cell differentiation
173 analysis using CytoTRACE for neural, muscle, and connective clusters (Figure S73). Clusters
174 8, 10, and 16 had the lowest scores in each respective comparison, which indicated
175 undifferentiated cell states. Thus, it is likely that clusters 8, 10, and 16 represented
176 undifferentiated cells within the major lineages of neural, muscle, and connective cells.

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

179 We built gene networks/modules from 3,000 variable genes using weighted gene
180 network correlation analysis (WGCNA; Langfelder & Horvath, 2008). This produced 43
181 gene modules in total (Supplementary File 8, 9), assessed for each cluster-module pair for
182 their strength of association (Figure 3A, Supplementary File 10, 11) and every module's
183 dependence on each cluster for their network connectivity (Figure S74, Supplementary File
184 12, 13). Using the genes from each network, we completed a KEGG pathway analysis to
185 identify whether gene modules indicated specific cellular pathways or states (Figure 3B).
186 We initially explored whether these network-cluster associations could reveal
187 commonalities between cell clusters or identify whether particular clusters contained
188 multiple cell identities.

189 First, we asked whether gene modules that associate with three or more cell clusters
190 signify commonalities between clusters that have similar cell types. We identified seven
191 gene modules (6, 7, 21, 14, 17, 41, and 42) that each associate with three or more cell
192 clusters. These gene modules do connect clusters of similar identities or cell states. For
193 example, Modules 6 and 7 associate with connective tissue cells. Module 6, the larger of the
194 two modules, contains numerous other KEGG pathways found in most connective cell
195 clusters, such as 'cytokine-cytokine receptor interaction' (Figure 3C). Interestingly, module
196 21, associated with clusters 8, 10, 16, and 25, contains genes from KEGG pathways related
197 to the 'cell cycle' and 'cellular senescence', supporting the results found in our cluster based
198 KEGG pathway analysis (Figure 3D). Module 21 also contained 'Notch signaling' genes,
199 possibly due to similar correlations with cell cycle genes; however, these were only
200 expressed in the neural cell clusters (8 and 25).

201 Where one cell cluster is associated with multiple gene networks, we wondered if
202 multiple cell identities existed within the cell cluster. We explored this possibility by
203 examining the pigment cell cluster (cluster #20) and whether the five different correlated
204 gene modules (#34, 35, 36, and 37, and 38) are expressed in distinct subgroups of cells
205 within the cluster (S75). We found cases of more than one genetic network (modules 35
206 and 38) expressed in the same cells. Modules 35 and 38 included conserved pigment genes,
207 *pmela*, *mlana*, and *dct* in module 35 and *tryp1b* and *pmel* in module 38, that mark
208 melanocytes (Figure S6; Du et al., 2003; Johnson et al., 2011; Lamason et al., 2005; Thisse &

209 Thisse, 2004).However, we also found non-overlapping expression of networks, notably
210 modules 36 and 37. We inferred that module 36 is associated with xanthophores and
211 xanthoblasts due to the presence of *plin6* and *scarb1*, genes involved with lipid binding and
212 activity in xanthophores (Ahi et al., 2020). On the other hand, module 37 likely represents
213 iridophores and iridoblasts because it contains *pnp4a,* which is involved in purine-
214 nucleoside phosphorylase activity in iridoblasts (Kimura et al., 2017).

215 Conserved signaling pathways are active during syngnathid craniofacial 216 development

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

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

230 Other genes identified here as cell markers were not uniquely craniofacial but
231 provide insights into the cell types that comprise the face. For example, *tnmd* expression
232 marked tendons and ligaments (cluster #9) throughout the face and body (Figure 4C;
233 Figure S65). Our finding is consistent with *tnmd's* role in tenocyte development in model
234 systems, namely zebrafish and mouse (Chen & Galloway, 2014; Docheva et al., 2005).
235 Identifying tenocyte cells is particularly relevant in syngnathid fishes where tendons are
236 enlarged and store elastic energy necessary for their specialized feeding (Van Wassenbergh
237 et al., 2007).

238 The last category of markers contained genes with regulatory roles important for
239 craniofacial development such as *prdm16,* the marker for osteochondro-mesenchymal cells
240 (cluster #5). The gene *prdm16* mediates the methylation of histones and regulates gene
241 expression, key for promoting craniofacial chondrocyte differentiation (Ding et al., 2013;
242 Kaneda-Nakashima et al., 2022; Lomeli Carpio Shull et al., 2020). We found *prdm16*
243 expressed in mesenchymal cells directly above the ethmoid plate and in fins (Figure 4D;
244 Figure S61). Comparison of pipefish with a related, short snouted fish (stickleback)
245 identified similarity in fin expression, but differences in craniofacial staining (S76).
246 Stickleback expressed *prdm16* in the hindbrain, gill arches, and lower jaw, consistent with
247 published zebrafish data (Ding et al., 2013), but staining was additionally detected in the
248 ethmoid region in pipefish.

249 We next examined signaling pathways active in craniofacial development. Our KEGG
250 pathway analysis revealed that MAPK, Wnt, and TGF-beta signaling pathways were
251 significantly enriched in one or more craniofacial contributing cell clusters (module #6;
252 Figure 2B, Figure 3C, Figure 4A). In addition to these KEGG findings, expression of ligands
253 and receptors that are members of these pathways (Fgf, a MAPK pathway, Wnt, and Bmp, a
254 TGF-beta pathway) was observed in cell clusters of both differentiated and
255 undifferentiated states (Figure 4A). We chose three genes, one from each major pathway,
256 for *in situ* hybridizations: *fgf22* (a Fgf ligand present in the actively dividing cells' module),
257 *bmp4* (a Bmp ligand present in the largest connective tissue module), and *sfrp1a* (a Wnt
258 pathway enabler present in a cartilage gene module). Although mouse and zebrafish
259 studies identified *fgf22* expression only in the nervous system (Miyake & Itoh, 2013;
260 Umemori et al., 2004), we found *fgf22* is expressed at the tip of the palatoquadrate and
261 Meckel's cartilage, the gill arches, fins, and brain in pipefish (figure 4E; S77). These same
262 expression patterns were observed in stickleback, suggesting a surprising co-option of fin
263 and craniofacial expression in percomorph fish.

264 We observed *bmp4* and *sfrp1a* expressed above the ethmoid plate and along the
265 ceratohyal in pipefish (figure 4F-G). The gene *bmp4* has a conserved role in craniofacial
266 development, particularly important at later stages for driving chondrocyte differentiation
267 (Wang et al., 2024; Zhou et al., 2013). However, the specificity of *bmp4* expression to

268 mesenchyme around the ceratohyal and ethmoid was not observed in stickleback fish,
269 which had broad craniofacial expression (including the jaws, tooth germs, and gill arches;
270 S78). Interestingly, *sfrp1a* expression has not been observed in the palate of mice or
271 ethmoid region of zebrafish, but *sfrp1a* craniofacial expression was identified in stickleback
272 (lower jaw and gill arches; S79) and has been observed in other fishes (Ahi et al., 2014;
273 Schneider, Gunter, et al., 2023; Schilling & Kimmel, 1997; Swartz et al., 2011; Wang et al.,
274 2024).

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

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

285 Our thorough examination of cell clusters for identity annotation did not find a tooth
286 primordium cluster. We therefore searched for tooth primordia by examining expression of
287 specific odontogenesis marker genes (*aldh1a2*, *bmp4*, *dlx2a*, *dlx3b*, *lef1*, *lhx6a*, *lhx8*, *msx1a*,
288 *msx2*, and *wnt10a*; Figure 5A). We observed several primordium genes expressed in our
289 atlas. However, there was no cluster with every marker gene expressed in over 10% of
290 cells. Several markers (*dlx3b*, *lef1*, *msx1a*, *msx2*, and *wnt10a*) were expressed in cluster #28,
291 a fin cluster distinguished by *hoxa13a* and *hoxa13b* expression. Since we previously noted
292 that cluster #16 seems to be a primordial connective tissue cluster, we wondered if it could
293 contain tooth primordial cells. In this cluster, we found the following percentage of
294 expression of each gene in the cells: *bmp4* in 46%, *aldh1a2* in 29%, *lef1* in 24%, *dlx2a* in
295 18%, *dlx3b* in 16.4%, *eda* in 12.7%, *msx1a* in 10.6%, *pitx2* in 10%, *lhx6a* in 8%, *wnt10a* in
296 4.6%, *msx2* in 2.65%, *lhx8* in 1.8%, and *shha* in 1%.

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

309 We next investigated *scpp* genes (*enam*, *scpp1*, and *spp1*), which are expressed in
310 late tooth development and tooth maintenance in other vertebrates (Figure 5A). These
311 genes also have some expression outside of teeth such as in dental bone (*scpp1*, *spp1;*
312 Kawasaki, 2009) and fins (*enam;* Jain et al., 2007). We identified *spp1* and *scpp1* expression
313 in 54.5% and 24% of bone cells, respectfully, and sparse *spp1* expression (<3%) in other
314 connective tissue cell types. We found minimal expression of *enam* (in 11.4%) in epidermal
315 cells and in less than 8% of muscle cells. In whole mount *in situ* hybridization, we found
316 that *scpp1* is expressed in all developing pipefish bones, both endochondral and dermal
317 (Figure 5G). Since *scpp* gene losses observed in syngnathids have been hypothesized to be
318 responsible for their tooth loss (Lin et al., 2016; Qu et al., 2021; Zhang et al., 2020), we
319 explored *scpp* gene family content in the close, tooth-bearing relative to syngnathids, the
320 blue spotted cornetfish, and we found several gene losses (*scpp4*, *scpp7*, and *scpp9*) and a
321 likely pseudogene (*scpp5*) (Figure S80; Hughes et al., 2018; Stiller et al., 2022).

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

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

326 Syngnathid dermal armor is mineralized dermal bone underneath the skin (Figure
327 6D, G, H). It is unknown when dermal armor primordia initiate and how they are patterned.
328 Spatial expression analysis of *pitx2* and *dlx3b* in search of tooth primordia instead revealed
329 expression of these genes in possible dermal armor primordia (Figure 6A,B, E,F). We found
330 *pitx2* staining localized dorsally to striated muscle underneath developing dermal armor.
331 The gene *dlx3b* is expressed in a repeating pattern along the body in the epidermal and
332 dermal tissues. Both staining patterns were in discrete regions of the muscle and epidermal
333 layers rather than being continuously expressed across the tissues. We did not observe the
334 expression of other tooth primordium genes (*bmp4, lhx6a*, and *dlx2a*) in this region.

335 Because the embryos from our atlas had not begun dermal armor mineralization,
336 the atlas cannot be directly used for discovery of genes active in dermal armor. However,
337 the atlas contains osteoblasts from craniofacial bones which we used to create osteoblast-
338 specific gene networks. We therefore asked whether these osteoblast genes were present
339 in mineralizing dermal armor at later stages. *In situ* hybridization expression analysis
340 revealed that *scpp1*, an osteoblast and tooth mineralization gene, and *ifitm5*, an osteoblast
341 gene, were expressed in the dermal armor at the onset of mineralization (Figure 6C; Figure
342 S68).

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

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

350 Within our larger KEGG analysis, we asked whether nutrient absorption and
351 immune KEGG terms were among the enriched pathways for the epidermal cells. We
352 identified 106 enriched genes in the 'endocytosis' pathway (p-value=.036; Figure 7A). Four
353 metabolism pathways ('galactose', 'glutathione', 'sphingolipid', and 'starch and sucrose')
354 are also enriched. No immune-related KEGG terms were enriched in the epidermis. For

355 comparison, we investigated whether these KEGG terms are also enriched in epidermal
356 cells of non-brooding fishes. We completed a KEGG pathway analysis on a comparably
357 staged zebrafish single cell RNA sequencing atlas (3 days post fertilization; Lange et al.,
358 2023). The 3dpf zebrafish epidermal cells did not have a significant enrichment of the
359 'endocytosis' pathway (23 up-regulated genes, p-value=.99) or any metabolism pathway.
360 However, there are 11 endocytosis genes up-regulated in both zebrafish and pipefish
361 epidermal cells, suggesting conserved expression of these genes in pipefish.

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

373 We did not find any expression of C-type lectin genes in zebrafish epidermal cells
374 (Figure S81), unlike in pipefish. However, it is possible there are unannotated C-type lectin
375 genes in zebrafish that remained cryptic in the atlas. Interestingly, the 'C-type lectin
376 receptor signaling pathway' was significantly enriched in zebrafish epidermal cells (13
377 genes, p-value=.04) but not in pipefish epidermal cells (34 genes, p-value=.4). Although
378 these results seem paradoxical, the KEGG term 'C-type lectin signaling pathway' does not
379 include any of the C-type lectin genes themselves. Overall, these data suggest that the
380 expression of C-type lectin genes in the pipefish embryonic epidermis is potentially unique
381 and warrants further investigation.

382
Discussion

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

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

394 Developmental single cell atlases have elucidated cell identities and genetic
395 pathways active in model organisms such as zebrafish, mice, and chick (Farnsworth et al.,
396 2020; Farrell et al., 2018; Feregrino et al., 2019; Morrison et al., 2017; Soldatov et al., 2019;
397 Wagner et al., 2018; Williams et al., 2019). In less traditional models, the majority of
398 scRNAseq atlases are produced from adult tissues, allowing investigations into cell types,
399 population differences, and genetic networks (Chari et al., 2021; Fuess & Bolnick, 2023;
400 Hain et al., 2022; Hong et al., 2023; Koiwai et al., 2023; Parker et al., 2022; Potts et al., 2022;
401 Royan et al., 2021; Songco-Casey et al., 2022; Vonk et al., 2023; Woych et al., 2022).
402 Creation of developmental scRNAseq atlases in non-model organisms is just beginning to
403 accelerate, but many emerging models still lack such a resource, limiting our understanding
404 of their development (Healey et al., 2022; Salamanca-Díaz et al., 2022; Steger et al., 2022;
405 Ton et al., 2023). Our single cell atlas is one of the first created to understand the
406 development of derived traits in a non-model organism.

407 For syngnathids specifically, this atlas represents an important step towards
408 understanding the genetic nature of unique syngnathid traits. Numerous developmental
409 genetic changes can lead to evolutionary innovations, including the evolution of novel
410 genes, gene duplications, gene losses, gene family expansions or contractions, evolution of
411 regulatory elements, co-option of gene regulatory networks, re-wiring of gene networks,

412 assembly of novel gene networks, and/or the emergence of novel cells (Arendt et al., 2016;
413 Cãestro et al., 2007; Teichmann & Babu, 2004; Wagner, 2011; Wagner & Lynch, 2010).
414 Considering these possibilities, we examined select syngnathid traits and speculated on
415 developmental genetic mechanisms influencing their evolution. Continuing to investigate
416 these proposed mechanisms through expanded scRNAseq atlases and other studies will be
417 critical for understanding syngnathid evolution.

418 Conserved pathways may contribute to derived syngnathid heads

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

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

441 region. Future studies investigating this testable hypothesis would clarify whether
442 prolonged chondrocyte differentiation broadly underlies craniofacial diversity.

443 We found *fgf22* expressed in the mandible, gill arches, and fins using *in situ*
444 hybridizations, but not in the elongating regions of the head. Interestingly, *fgf22* expression
445 has not been reported in craniofacial development of any other species (Miyake & Itoh,
446 2013), but we found similar craniofacial and fin expression in stickleback fish. Fgf signaling,
447 however, is a conserved and essential pathway for craniofacial development (Crump et al.,
448 2004; Leerberg et al., 2019; McCarthy et al., 2016; Szabo-Rogers et al., 2008; Walshe &
449 Mason, 2003; Woronowicz & Schneider, 2019), raising the possibility that *fgf22* has been
450 co-opted into a role played by different *fgf* genes in other species. Future work should
451 consider a relationship between the novel craniofacial expression of *fgf22* and the loss of
452 Fgf ligands *fgf3* and *fgf4* in syngnathids. If *fgf22* is active in existing gene networks,
453 particularly those where *fgf3* or *fgf4* is active in other species, then its novel expression
454 may indicate evolved genetic redundancy prior to the gene losses.
455 Our analysis suggests ways in which unique syngnathid craniofacial structure could
456 have evolved through genetic network evolution. Unusual expression location (e.g., *fgf22*)
457 and specificity (e.g., *sfrp1a*, *bmp4*, and *prdm16*) in pipefish compared to zebrafish and
458 stickleback suggests that changes in signaling gene deployment and/or content within
459 craniofacial gene networks, particularly genes from Wnt, Fgf, or TGF-beta families, could
460 underly the exceptionally elongated syngnathid face.

461 **Early, not late, tooth development is likely at the root of evolutionary tooth loss**

462 Tooth loss has occurred independently in numerous lineages and has often been
463 studied to understand the developmental basis of character loss. For instance, research in
464 birds and turtles found that tooth programs initiate but are subsequently truncated,
465 explaining toothlessness in mature animals (Chen et al., 2000; Tokita et al., 2013).
466 Additional studies in birds have found losses in tooth maturation genes (specifically *scpp*
467 genes; Sire et al., 2008). Since numerous primordium and maturation genes are lost in
468 syngnathids (Lin et al., 2016; Qu et al., 2021; Small et al., 2016; Small et al., 2022; Zhang et
469 al., 2020), we asked if syngnathids begin tooth development at all.

470 We found that early tooth development genes were still expressed in pipefish, which
471 is unsurprising given their pleiotropic roles, but found no convincing evidence either in the
472 atlas or in spatial gene expression analysis of tissues for tooth primordia. The possibility
473 remains, however, that a transient cell population could fall at a different developmental
474 time than assayed here. Barring this caveat, syngnathid *fgf3* and *fgf4* losses could have
475 resulted in insufficient Fgf signaling from the oral epithelium to the dental mesenchyme,
476 causing failure of tooth initiation (Small et al., 2022; Stock et al., 2006).

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

493 Redeployment of the bone gene network to build dermal armor

494 The syngnathid dermal armor is a type of evolutionary novelty, which can arise
495 through either differentiation of serially repeated elements or *de novo* origination, derived
496 from either the redeployment of existing gene networks, rewiring of existing gene
497 networks, or the assemblage of new gene networks (Wagner & Lynch, 2010). Currently,

498 there is no understanding of the developmental genetic underpinnings of the syngnathid
499 dermal armor.

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

507 Instead, dermal armor might originate from the co-option of existing bone
508 development gene regulatory networks. Expression of the gene *dlx3b* has been observed in
509 epithelia and mesenchyme during dermal and perichondral bone development in zebrafish
510 (Verreijdt et al., 2006). In addition to *dlx3b*, we observed bone development genes *scpp1*
511 and *ifitm5* expressed in the ossifying dermal armor. Future studies could test our
512 hypothesis that the dermal armor evolved through re-deployed osteoblast networks by
513 examining osteoblast gene network expression over time.

514 Signatures of embryonic interactions within the novel pouch environment

515 Syngnathid embryos are reared within the brood pouch, a novel structure and
516 environment composed of male-derived tissues (epithelium and placental-like tissues that
517 include specialized cell types) that harbors a pouch microbiome (Stölting & Wilson, 2007).
518 During pregnancy, the male brood pouch undergoes numerous changes including increased
519 vascularization and altered expression of immune genes (Harada et al., 2022; Ripley et al.,
520 2010; Roth et al., 2020; Small et al., 2013; Whittington et al., 2015). Researchers predict
521 that these changes relate to nutrient and waste transfer and prevention of embryonic
522 rejection and bacterial infection (Dudley et al., 2021; Whittington & Friesen, 2020).
523 However, there are few studies that examine whether and how embryos interact with the
524 brood pouch environment (Kvarnemo et al., 2011; Ripley & Foran, 2006). To consider
525 whether embryos have specializations for life in this brood pouch environment, we asked
526 about cell type-specific expression of nutrient acquisition and/or immune genes.

527 Pipefish embryos uptake paternally derived carbohydrates, proteins, and lipids
528 (Kvarnemo et al., 2011; Ripley & Foran, 2006). Our data suggest that this uptake could
529 occur through the embryonic epidermis. Specifically, we noticed an enrichment of
530 endocytosis and metabolism genes in epidermal cells. Epidermal absorption of maternally
531 derived nutrients has been suggested in viviparous fishes (Tengfei et al., 2021; Wourms,
532 1981). Interestingly, microvilli, a type of cellular projection, have been observed on the
533 anal fin of developing seahorses (Wetzel & Wourms, 2004). In light of our findings in
534 pipefish, possibly these seahorse microvilli could be functionally equivalent to those on the
535 small intestine, maximizing nutrient absorption from the environment.

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

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

552 Conclusions

553 Our study represents the first scRNAseq developmental atlas in syngnathids, and
554 one of the first non-model developmental scRNAseq atlases, providing a major step
555 forward for evo-devo research. We used our atlas to begin addressing questions on the

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

571 # Methods

572 ## Single Cell RNA Sequencing Libraries Preparation

573 We created scRNAseq atlases from embryos of wild caught Gulf pipefish
574 (*Syngnathus scovelli,* acquired from collaborator Emily Rose using collection permit SAL-
575 21-0182-E), and all work was performed under an approved IACUC protocol. We harvested
576 20 embryos per pouch from two wild caught male pipefish. The embryos were at a stage
577 before the tubular face was fully elongated, and while the head skeleton is cartilaginous
578 with minimal signs of mineralization of superficial intramembranous bones. This
579 corresponds to a stage termed "frontal jaws" in a recent description of pipefish
580 development (Sommer et al., 2012).

581 We dissociated the embryos using 460ul of .25% trypsin in water and 40ul 100
582 mg/mL collagenase I (Sigma C0130-200mg) for 16 minutes. We filtered cells using a 40uM

583 cell strainer (Thomas Scientific #1181X52). We quantified cell concentrations using the
584 TC20 Automated Cell Counter (Biorad) then diluted the samples to 800 cells/ul in .04%
585 BSA in PBS. The University of Oregon Genomics and Cell Characterization Core (GC3F;
586 https://gc3f.uoregon.edu) prepared single cell libraries for each sample using 10X
587 Genomics Single Cell 3' Genome Expression mRNAseq kit with NextGEM v3.1 chemistry.
588 We sequenced these libraries on an S4 lane on the NovaSeq 6000 at the GC3F. To improve
589 the 3' UTR genome annotations, we also prepared scISOrSeq libraries from the first
590 embryonic sample and from dissociated pouch cells from pregnant and nonpregnant males.
591 These libraries were produced in accordance with (Healey et al., 2022). Embryonic,
592 pregnant pouch and non-pregnant pouch libraries were sequenced separately on PacBio
593 Sequel II - SMRT Cells 8M.

594 To turn the scISOrSeq reads into gene models, we followed the pipeline from Healey
595 et al. 2022. We ran the script (scISOr_Seq_processing.py from
596 https://github.com/hopehealey/scISOseq_processing) to remove barcodes, identify cell
597 barcodes, and demultiplex with the single cell flag and appropriate barcodes (5'
598 CCCATGTACTCTGCGTTGATACCACTGCT and 3' CTACACGACGCTCTTCCGATCT). We aligned
599 the reads to the 2022 Gulf pipefish genome (GenBank: GCA_024217435.2) using minimap
600 v2.9 (Li, 2018). We filtered the reads using cDNA cupcake to remove duplicate transcripts
601 (Tseng, 2021). We used SQANTI3 to identify gene models and filter them (Tardaguila et al.,
602 2018). We merged the SQANTI3 annotations with the Gulf pipefish genome (NCBI
603 GenBank: GCF_024217435.2) using TAMA merge (Kuo et al., 2017). Since the Gulf pipefish
604 genome does not contain mitochondrial genes, we appended the annotation and fasta files
605 with the Gulf pipefish mitochondrial genome (NCBI RefSeq: NC_065499.1).

606 **Single Cell Atlas Construction**

607 We ran Cell Ranger (10X Genomics 3.0.2) using our scRNAseq reads, the Gulf
608 pipefish genome assembly with mitochondrial genome, and the modified gene annotations.
609 Cell Ranger estimated 20,733 cells for sample one, 23,682 genes expressed, and 21,039
610 mean reads per cell. For sample two, Cell Ranger predicted 17,626 cells, 23,740 genes

611 expressed, and 29,804 mean reads per cell. We analyzed Cell Ranger's output using Seurat
612 (v4.1.0) on R (v4.0.2; Butler et al., 2018; Hafemeister & Satija, 2019).

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

622 We normalized the datasets with SCTransform (v0.3.3). We used Seurat's
623 integration tools, `SelectIntegrationFeatures` using 3,000 feature genes,
624 `FindIntegrationAnchors` using `SCT` normalization, and `IntegrateData` using `SCT`
625 normalization, to integrate the two datasets. After integration, our combined atlas had
626 35,785 cells (Supplementary File 1,2). We then used the integrated dataset to complete
627 PCA analysis. We tested using a variety of principle components for further analysis and
628 chose 30 PCs for our analysis based on the clear delineation of major cell types. We next
629 clustered the cells using 30 PCs and plotted the data on a UMAP with Seurat.

630 **Single Cell Atlas Cluster Identification**

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

640 annotations. For each cluster, we examined multiple genes using DotPlots and FeaturePlots
641 to propose the cluster identity.

642 Next, we identified one gene for each cluster which marked the cluster best
643 (expressed in the most cells in the focal cluster and expressed in as few of the other
644 clusters as possible) through consulting the two marker gene lists and examining markers
645 with Dot Plots (Supplementary File 3, 4, 5). Using these markers, we completed a set of *in*
646 *situ* hybridizations to hone our cluster annotations. Due to challenges culturing Gulf
647 pipefish, we used both embryos and larvae from *Syngnathus leptorhynchus*, a pipefish from
648 the same genus that lives in Oregon coastal habitats and allowed for easy collection, and
649 from cultured Gulf pipefish for the cluster annotation *in situ* experiments. We caught
650 pregnant male *Syngnathus leptorhynchus* using a beach seine near Coos Bay, Oregon under
651 Oregon Department of Fisheries and Wildlife permit number 26987.

652 *Syngnathus scovelli* used for *in situ* experiments were purchased from Alyssa's
653 Seahorse Saavy and Gulf Specimens Marine Lab then reared in our facility at 25 degrees C
654 water and 25-28 PPT Salinity. We designed probes using NCBI Primer Blast with the Gulf
655 pipefish genome assembly and produced these probes using Gulf pipefish embryonic cDNA
656 pools. To create the probes, we completed two rounds of PCR. The first round used a gene
657 specific forward primer with a reverse gene specific primer that had 10 nucleotides of T7
658 promotor sequence attached. PCR products were cleaned with Zymo clean and
659 concentrator DNA kit and eluted in 15 ul of elution buffer. Round two of PCR used the same
660 gene specific forward primer with a modified T7 promoter sequence
661 (TGGACTAATACGACTCACTATAGGG) as the reverse primer, and finally the product was
662 cleaned again with Zymo clean and concentrator DNA kit and eluted in 15 ul of elution
663 buffer.

664 Round one PCR conditions are as followed: 95 degrees Celsius for 3:00 minutes, 40
665 cycles of denaturation (95 degrees for 30 seconds), annealing (30 seconds, annealing
666 temperature varied by probe), and extension (72 degrees for 1:00 minute), and a final
667 extension step of 3:00 minutes. Round two PCR conditions are as followed: 98 degrees
668 Celsius for 30 seconds, 30 degrees for 10 seconds, 72 degrees for 50 seconds, then 35

669 rounds of denaturation (98 degrees for 10 seconds), annealing (50 degrees for 10 seconds),
670 and extension (72 degrees for 50 seconds), and a final extension step of 10 minutes at 72
671 degrees. For round two PCRs with multiple bands (specifically, *ifitm5*), the band of the
672 expected size was excised with a razorblade and DNA was extracted with Zymoclean gel
673 DNA recovery kit.

674 The round two PCR product was sanger sequenced to confirm identity. All the
675 marker genes primers and successful PCR conditions are in Supplementary File 6.
676 Alignments of the probes to the unpublished *Syngnathus leptorhynchus* genome assembly
677 are included in Supplementary File 7. The probes were transcribed with T7 polymerase for
678 2-6 hours then cleaned with Zymo RNA clean and concentrator and eluted into 30 ul of
679 water. For the *in situ* hybridizations, we selected embryos and newly spawned larvae close
680 to the developmental stage used in the atlas. We completed *in situ* hybridizations in
681 keeping with (Thisse & Thisse, 2007), leaving the embryos in stain until background was
682 observed. For spawned larvae, we completed a bleaching step (1% H2O2 and .5% KOH for
683 8 minutes) prior to the proteinase K digest. After imaging, we used the levels tool in Adobe
684 Photoshop (v23.4.2) to white balance the pictures.

685 **Single Cell KEGG Analysis**

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

692 **Differentiation State Analysis**

693 To assess whether proposed primordial cell clusters were composed of
694 undifferentiated cells relative to other clusters from their lineages, we completed a
695 differentiation analysis using CytoTRACE (v0.3.3, Gulati et al., 2020). Clusters from similar
696 lineages (neural: 0, 3, 7, 8, 12, 22, 25, 33, and 35; muscle: 2, 10, 17, and 37; connective: 4, 5,

697 6, 9, 15, 16, 18, 24, 27, 28, and 29) were isolated using subset. Cell counts were gathered
698 using as.Matrix(GetAssayData), then CytoTRACE was run on these counts. Data was plotted
699 on a VlnPlot to show the variability.

700 ### Single Cell Atlas Network Analysis

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

709 We calculated the dissimilarity matrix by calculating TOM similarity of the
710 adjacency matrix and subtracting it from one. We then created a gene tree through `hclust`
711 of the dissimilarity matrix with the average method. From the gene tree, we created
712 modules using `cutreeDynamic` with deepSplit of two and the minClusterSize of 15 (setting
713 the smallest cluster size to 15). We calculated module eigengenes using `moduleEigengenes`
714 of the adjacency matrix then calculated the dissimilarity of the module eigengenes with `cor`
715 of the module eigengenes subtracted from one and clustered the module eigengenes with
716 `hclust`. We then chose a dissimilarity of 35% as the cut off for merging modules; however,
717 no modules had dissimilarity scores with each other below 40%. Since the module
718 eigengenes are calculated for each cell in the dataset, we used these values to calculate the
719 t-statistic for each cell cluster (considered each sample) in the module (all of the module
720 eigengene scores were used as the population).

721 We next tested the hypothesis that certain cell clusters were strongly associated
722 with specific gene modules through a two-way permutation test (1000 permutations) and
723 corrected p-values to control the False Discovery Rate (FDR). Additionally, we tested
724 whether specific cell clusters drove underlying gene network structure by measuring
725 network connectivity. The network connectivity was first calculated using the entire

726 dataset, then each cell cluster was progressively dropped from the dataset and the
727 connectivity was remeasured. This resulted in change in connectivity scores for each
728 module-cluster pair. To assess whether these changes were significant, we completed 1000
729 permutations whereby cells were randomly dropped (the number of cells dropped was
730 equal to the cell cluster size of the focal cluster), connectivity measured, and the change in
731 connectivity recorded. The p-value is the number of instances where the change in
732 connectivity is greater in the permutations than in the focal cell cluster run. P-values were
733 corrected to control the FDR.

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

740 *In situ* Hybridization

741 For genes chosen for further follow up analysis, we completed *in situ* hybridizations
742 in Gulf pipefish (*Syngnathus scovelli*) or bay pipefish (*Syngnathus leptorhynchus*). Primer
743 sequences were designed using NCBI Primer Blast with the Gulf Pipefish genome and
744 synthesized using Gulf pipefish embryonic cDNA. The *sfrp1a* probe was synthesized using
745 the PCR based probe preparation protocol described in the Single Cell Identification
746 section. All other probes (*bmp4*, *dlx2a*, *dlx3b*, *fgf22*, *lhx6a*, *pitx2*, and *scpp1*) were prepared
747 using TOPO cloning. Probe primer sequences as well as the species used for the *in situ*
748 experiments are described in Supplementary file 6. Pregnant male bay pipefish were
749 caught as described above. For Gulf pipefish embryos, we allowed the fish to mate in our
750 facility then harvested embryos once they reached the appropriate stages. Fish were reared
751 in 25 degrees C water with 25-28 ppt salinity. Embryos and larvae were fixed in 4% PFA,
752 dehydrated through a series of PBT/MeOH washes, and stored in MeOH at -20C. 5-12
753 embryos were used for each probe. We completed *in situ* hybridizations in keeping with
754 (Thisse & Thisse, 2007), leaving the embryos in stain until background was observed. After

755 imaging, we used the levels tool in Adobe Photoshop (v23.4.2) to white balance the
756 pictures.

757 Bone and Cartilage Staining

758 We used alcian and alizarin stains to mark cartilage and bones. Specifically, we
759 assayed cartilage and bone development in siblings of the scRNAseq samples. These
760 embryos were fixed in 4% PFA then stored at -20C in MeOH. We followed the protocol from
761 Walker and Kimmel (2007) with minor alterations. We stored samples in 50%
762 glycerol/0.1% KOH at 4C and imaged them in 100% glycerol. After imaging, we white
763 balanced the photographs using Photoshop (v23.4.2) levels tool.

764 Gene Cluster Analysis

765 To examine close syngnathid outgroups, we downloaded the 2023 mandarin
766 dragonet genome (GenBank assembly accession: GCA_027744825.1) and the 2024
767 cornetfish genome assembly (GenBank assembly accession: GCA_037954325.1) from NCBI.
768 Since these genomes were unannotated, we manually identified *scpp* genes. We searched
769 for *scpp* genes using BLASTN with medaka and additional fish sequences as the query. We
770 also searched for these genes with mVISTA plots across conserved gene synteny regions
771 (LAGAN alignment using translated anchoring) with medaka as the focal species. To
772 identify *scpp* genes in additional species, we gathered cluster information from NCBI and
773 ensembl. Additionally, we used mVISTA plots to further search for unannotated *scpp* genes.

774

775 Acknowledgements:

776 We are indebted to Emily Rose for her collection of the Gulf pipefish specimens for
777 the single cell libraries. We are grateful for everyone in the Cresko Lab for their ideas and
778 participation in early morning pipefish collection trips. We are additionally thankful for
779 Emily Beck for her assistance collecting pipefish. In particular, we are immensely
780 appreciative of Mark Currey and Tiffany Thornton's valiant efforts to culture Gulf pipefish.
781 Additionally, we are grateful for Balan Ramesh and Adam Jones for collaborating with the

782 Cresko Lab to produce new Syngnathiformes genomes. We thank Tina Arredondo from the
783 UO GC3F for her preparation of our single cell RNA sequencing libraries. This work was
784 funded by the National Science Foundation Grant OPP-2015301 (to WAC, SB, and CMS),
785 University of Oregon Research Excellence funds (WAC) and National Institute of Health
786 Fellowship F31DE032559-02 (to HMH). Additionally, HMH was supported by the Genetics
787 Training Program (NIH T32GM007413). Finally, VG received funding and mentorship from
788 the Hui Undergraduate Research Scholars and MW was supported by the Knight Campus
789 Undergraduate Scholars program at the University of Oregon.

790 Data Availability statement

791 All raw sequencing data associated with this study are published via NCBI
792 (PRJNA1168967). The integrated single cell RNA sequencing atlas is also available through
793 NCBI (GSE278814). The fasta file and updated Gulf pipefish annotation are stored on Dryad
794 (10.5061/dryad.f7m0cfz60). Code used for the analysis is available on github
795 (https://github.com/hopehealey/ssc_emb_atlas_scripts).

796 Figure captions:

797 **Figure 1: Gulf pipefish exemplify syngnathid derived traits.** Gulf pipefish have elongate
798 snouts, have lost teeth on their oral and pharyngeal jaws, possess dermal armor, and have
799 brood pouches in males (A-B). Cartilage (alcian) and bone (alizarin) stained clutch siblings
800 of embryos from the two scRNAseq samples are shown in panels C-H. Embryos have
801 cartilaginous craniofacial skeletons (panels B, C, F; E marks the (Mes)ethmoid cartilage, C
802 indicates the Ceratohyal, H shows the Hyosymplectic cartilage, M marks the Meckel's
803 cartilage, Q indicates the quadrate, B shows the Basihyal, and P marks the palatoquadrate)
804 with the onset of ossification in the jaw. They have cartilaginous fin radials in the dorsal fin
805 (D and G; F indicates fin radials and N denotes the notochord). The embryos do not have
806 signs of ossification in the trunk where the exoskeleton will form later (panels E and H; PD

807 shows the region where dermal armor primordia will arise). Panels C, D, F, and G, scale bar
808 200 μm; E, H scale bar 100 μm.

809 **Figure 2: Gulf pipefish single cell atlas contains cells from the entire embryo and**
810 **identifies genetic pathways active in different cell types.** The UMAP plot (panel A)
811 shows all of the cell clusters and their identities reduced to the first two UMAP dimensions.
812 The graph in panel B displays results of the KEGG pathway analysis in cell clusters
813 identified as connective tissue (excluding blood, pigment, digestive, and immune cells). The
814 number of Seurat identified marker genes for each cluster that was a part of each pathway
815 is displayed on the y axis. Bars are colored and labeled by cell cluster.

816 **Figure 3: Weighted Gene Network Analysis (WGCNA) identifies gene modules that**
817 **define and unite cell clusters.** A) The strength of association between the gene modules
818 and cell clusters is shown in panel A with dendrogram clustering illustrating the distance
819 between modules and cell clusters. Gene modules are represented by rows and cell clusters
820 by columns. The modules and clusters are clustered using the Pearson distance method.
821 The number of genes in each gene module are shown in the right hand bar plots. Cell
822 clusters are colored based on their identity. The asterisks indicate the module-cluster
823 relationships that have a p-value less than .05 from a two-sided permutation test after
824 correction for multiple tests (FDR). The heatmap is colored by t-statistics in a range of -10
825 to 10, with highly positive values in yellow and highly negative values in black. B) The
826 identified gene modules possess genes from KEGG pathways. The bars are labeled with the
827 gene module and the size of each bar corresponds to the number of genes from the KEGG
828 pathway in the module. Since WGCNA modules do not have p-values, only KEGG pathways

829 with more than two genes included in the gene module are shown on the plot. C) Identified
830 gene modules contain similar KEGG pathways as the cell clusters that correlated with them.
831 These relationships are shown in Upset plots where each row is a cell cluster or gene
832 module, each column represents KEGG pathways shared by the modules and clusters
833 (shared condition is shown filled in black dots connected by lines), the interaction size is
834 the number of pathways in common between the set of modules and clusters, and the set
835 size is the number of pathways that are enriched in each cluster and module. Panel C1
836 highlights that 'cytokine-cytokine receptor interaction' and 'ECM receptor interaction' are
837 present in module 6 as well as 6 and 4 connective cell clusters respectively. Panel C2 shows
838 that 'cell cycle' and 'senesence' are present in module 21 as well as clusters 8 and 16,
839 'Notch signaling' genes are present in clusters 8 and 25 as well as module 21, and 'DNA
840 replication' is present in clusters 8, 10, and 16.

841 **Figure 4: Conserved cell types and gene pathways build unique faces of syngnathids.**
842 Three main conserved signaling pathways are enriched in connective cell types, MAPK
843 signaling (including Fgf signaling), TGF-beta signaling (including Bmp signaling), and Wnt
844 signaling. Receptors and ligands expression patterns are shown in panel A heatmap from
845 all cell types with cells present in the head. This heatmap features 100 cells downsampled
846 from each cluster and illustrates that many genes from these families are expressed in
847 these cells. Yellow lines indicate high expression of a gene, while hot pink lines indicate no
848 expression. The pathways are boxed in black. Rows representing *fgf22, bmp4*, and *sfrp1a*
849 expression are marked with an asterisk and green box for each respective section of
850 signaling (Fgf, Bmp, and Wnt). Panels B, C, and D are *in situ* hybridizations of three marker
851 genes, *elnb, tnmd*, and *prdm16*. The genes *prdm16* and *elnb* mark osteochondrogenic

852 mesenchyme and *tnmd* marks tendons and ligaments. Panels E, F, and G, show expression
853 patterns of three pathway representatives *fgf22, bmp4*, and *sfrp1a*. All three genes are
854 expressed in the face: *fgf22* at the tip of the mandible and *bmp4* and *sfrp1a* above the
855 ethmoid and near the ceratohyal. Staining is circled with dashed lines. The Meckel's
856 cartilage (M), (mes)Ethmoid cartilage (E), Quadrate (Q), and Ceratohyal (C) are labeled.
857 Panel C is a dorsal view. Panels B, C, E, F, G, H, I are in lateral view. *In situ* experiments of
858 fgf22 were completed using 10dpf Gulf pipefish. In situ experiments of *bmp4, sfrp1a, tnmd*,
859 *elnb*, and *prdm16* were completed using wild caught bay pipefish at the onset craniofacial
860 elongation. Panels H and I are summary illustrations of our findings, panel H shows where
861 cells from various clusters were present in the developing head and panel I illustrates
862 where *bmp4, sfrp1a*, and *fgf22* were expressed. Panel H is colored according to cell type.

863 **Figure 5: Pipefish do not possess identifiable tooth primordium cells, but continue to**
864 **express tooth development genes in other contexts.** Panel A presents a dotplot of genes
865 classified by the tissue layer in which they are reported to be expressed during tooth
866 development in other vertebrate (Tucker and Sharpe 2004; Gilbert et al. 2019; Kawasaki
867 2009). The x-axis contains the assayed genes, with asterisks under the genes that were also
868 examined with *in situ* hybridizations. The y-axis contains all the cell clusters. The size of the
869 dot is representative of the percentage of cells from the cluster that express the gene. The
870 color of the dots is average expression of the gene in the cluster (darker purples represent
871 higher expression). Panel B includes *in situ* hybridizations of selected tooth primordia
872 markers (*bmp4, pitx2, lhx6a, dlx2a*, and *dlx3b*) and mature tooth marker (*scpp1*). *In situ*
873 experiments of *bmp4, pitx2, lhx6a, dlx2a*, and *dlx3b* were completed with wild caught bay

874 pipefish that had begun craniofacial elongation. *In situ* experiments of *scpp1* were
875 completed using 9dpf Gulf pipefish. The scale bars for all images represents 100 μm.

876 **Figure 6:** Tooth and bone development genes expressed during exoskeleton development.
877 We discovered *pitx2* (A and B) and *dlx3b* (F and E) expression during possible emergence
878 of exoskeletal primordium in wild caught bay pipefish. The embryos used for these *in situ*
879 experiments were the same stage as those from Figure 5, at the beginning of craniofacial
880 elongation. Figures A and F have 100 μm scale bars, figures B and F have 50 μm scale bars.
881 We further found *scpp1* is expressed at the mineralization front of the exoskeleton in 12dpf
882 Gulf pipefish (C) and has a 50 μm scale bar. Alizarin and alcian stained pipefish are shown
883 in panels D (12dpf Gulf pipefish), G (1dpf Gulf pipefish), and H (adult Gulf pipefish) to
884 illustrate how the exoskeleton forms. Panels D and G have 100 μm scale bars.

885 **Figure 7: Gene expression signatures suggest embryonic interactions within brood**
886 **pouch environment.** Epidermal cells (panel A), with pathways that suggest increased
887 endocytosis and metabolism pathways are in bold text. Pathways upregulated in 3dpf
888 zebrafish epidermal cells are in italics. Pipefish epidermal cells also express 15 lectin genes
889 not found in other cell types (panel B). We suggest an epidermal cell model (panel C), in
890 which we predict pipefish have an enrichment of nutrient processing genes and lectins in
891 comparison with zebrafish cells.

892 Supplementary figure captions:

893 **Figure S1: Two Gulf pipefish samples have similar contributions to cell clusters.** This
894 UMAP features cells plotted by sample (sample one in orange and sample two in cyan) in
895 the first two UMAP dimensions.

896 **Figure S2: QC metrics for both Gulf pipefish samples show similar sample quality**
897 **after initial filtering.** Sample quality was assessed using the number of genes identified
898 (nFeatureRNA), the total RNA counts (nCountRNA), the percentage of mitochondrial reads
899 (percentmt), and the likelihood of being a doublet (doubvalues). Gulf pipefish sample data
900 post-filtering is displayed for each of these metrics.

901 **Figure S3: Zebrafish muscle cell marker genes expression patterns correspond with**
902 **annotated muscle cells in Gulf pipefish atlas.** Feature plots of zebrafish muscle cell
903 marker genes identified from the Daniocell scRNAseq database demonstrate Gulf pipefish
904 expression. Annotated muscle cell clusters in pipefish are distinguished by a dashed
905 outline.

906 **Figure S4: Zebrafish muscle satellite cell marker genes expression patterns**
907 **correspond with annotated muscle primordia cells in Gulf pipefish atlas.** Feature plots
908 of zebrafish muscle satellite cell marker genes identified from the Daniocell scRNAseq
909 database demonstrate Gulf pipefish expression. Annotated muscle primordia cell cluster in
910 pipefish is distinguished by a dashed outline.

911 **Figure S5: Zebrafish cardiac muscle cell marker genes expression patterns**
912 **correspond with annotated cardiac muscle cells in Gulf pipefish atlas.** Feature plots of
913 zebrafish cardiac muscle cell marker genes identified from the Daniocell scRNAseq
914 database demonstrate Gulf pipefish expression. Annotated cardiac muscle cell cluster in
915 pipefish is distinguished by a dashed outline.

916 **Figure S6: Zebrafish retinal cell marker genes expression patterns correspond with**
917 **annotated retinal cells in Gulf pipefish atlas.** Feature plots of zebrafish retinal cell
918 marker genes identified from the Daniocell scRNAseq database demonstrate Gulf pipefish
919 expression. Annotated retinal cell cluster in pipefish is distinguished by a dashed outline.

920 **Figure S7: Zebrafish neuronal cell marker genes expression patterns correspond**
921 **with annotated neuronal cells in Gulf pipefish atlas.** Feature plots of zebrafish neuronal
922 cell marker genes identified from the Daniocell scRNAseq database demonstrate Gulf

923 pipefish expression. Annotated neuronal cell clusters in pipefish are distinguished by a
924 dashed outline.

925 **Figure S8: Zebrafish chondroblast cell marker genes expression patterns correspond**
926 **with annotated chondroblast cells in Gulf pipefish atlas.** Feature plots of zebrafish
927 chondroblast cell marker genes identified from the Daniocell scRNAseq database
928 demonstrate Gulf pipefish expression. Annotated chondroblast cell cluster in pipefish is
929 distinguished by a dashed outline.

930 **Figure S9: Zebrafish osteoblast cell marker genes expression patterns correspond**
931 **with annotated osteoblast cells in Gulf pipefish atlas.** Feature plots of zebrafish
932 osteoblast cell marker genes identified from the Daniocell scRNAseq database demonstrate
933 Gulf pipefish expression. Annotated osteoblast cell cluster in pipefish is distinguished by a
934 dashed outline.

935 **Figure S10: Zebrafish notochord cell marker genes expression patterns correspond**
936 **with annotated notochord cells in Gulf pipefish atlas.** Feature plots of zebrafish
937 notochord cell marker genes identified from the Daniocell scRNAseq database demonstrate
938 Gulf pipefish expression. Annotated notochord cell cluster in pipefish is distinguished by a
939 dashed outline.

940 **Figure S11: Zebrafish fin cell marker genes expression patterns correspond with**
941 **annotated fin cells in Gulf pipefish atlas.** Feature plots of zebrafish fin cell marker genes
942 identified from the Daniocell scRNAseq database demonstrate Gulf pipefish expression.
943 Annotated fin cell cluster in pipefish is distinguished by a dashed outline.

944 **Figure S12: Zebrafish mesenchymal cell marker genes expression patterns**
945 **correspond with annotated mesenchymal cells in Gulf pipefish atlas.** Feature plots of
946 zebrafish mesenchymal cell marker genes identified from the Daniocell scRNAseq database
947 demonstrate Gulf pipefish expression. Annotated mesenchymal cell clusters in pipefish are
948 distinguished by a dashed outline.

949 **Figure S13: Zebrafish tenocyte cell marker genes expression patterns correspond**
950 **with annotated tenocyte cells in Gulf pipefish atlas.** Feature plots of zebrafish tenocyte

951 cell marker genes identified from the Daniocell scRNAseq database demonstrate Gulf
952 pipefish expression. Annotated tenocyte and ligament cell cluster in pipefish is
953 distinguished by a dashed outline.

954 **Figure S14: Zebrafish erythrocyte cell marker genes expression patterns correspond**
955 **with annotated erythrocyte cells in Gulf pipefish atlas.** Feature plots of zebrafish
956 erythrocyte cell marker genes identified from the Daniocell scRNAseq database
957 demonstrate Gulf pipefish expression. Annotated erythrocyte cell clusters in pipefish are
958 distinguished by a dashed outline.

959 **Figure S15: Zebrafish endothelial cell marker genes expression patterns correspond**
960 **with annotated endothelial cells in Gulf pipefish atlas.** Feature plots of zebrafish
961 endothelial cell marker genes identified from the Daniocell scRNAseq database
962 demonstrate Gulf pipefish expression. Annotated endothelial cell cluster in pipefish is
963 distinguished by a dashed outline.

964 **Figure S16: Zebrafish gut and liver cell marker genes expression patterns**
965 **correspond with annotated gut and liver cells in Gulf pipefish atlas.** Feature plots of
966 zebrafish gut and liver cell marker genes identified from the Daniocell scRNAseq database
967 demonstrate Gulf pipefish expression. Annotated gut and liver cell cluster in pipefish is
968 distinguished by a dashed outline.

969 **Figure S17: Zebrafish epidermal cell marker genes expression patterns correspond**
970 **with annotated epidermal cells in Gulf pipefish atlas.** Feature plots of zebrafish
971 epidermal cell marker genes identified from the Daniocell scRNAseq database demonstrate
972 Gulf pipefish expression. Annotated epidermal cell cluster in pipefish is distinguished by a
973 dashed outline.

974 **Figure S18: Zebrafish glial cell marker genes expression patterns correspond with**
975 **annotated glial cells in Gulf pipefish atlas.** Feature plots of zebrafish glial cell marker
976 genes identified from the Daniocell scRNAseq database demonstrate Gulf pipefish
977 expression. Annotated glial cell cluster in pipefish is distinguished by a dashed outline.

978 **Figure S19: Zebrafish immune cell marker genes expression patterns correspond**
979 **with annotated immune cells in Gulf pipefish atlas.** Feature plots of zebrafish immune
980 cell marker genes identified from the Daniocell scRNAseq database demonstrate Gulf
981 pipefish expression. Annotated immune cell clusters in pipefish are distinguished by a
982 dashed outline.

983 **Figure S20: Zebrafish pigment cell marker genes expression patterns correspond**
984 **with annotated pigment cells in Gulf pipefish atlas.** Feature plots of zebrafish pigment
985 cell marker genes identified from the Daniocell scRNAseq database demonstrate Gulf
986 pipefish expression. Annotated pigment cell clusters in pipefish are distinguished by a
987 dashed outline.

988 **Figure S21: Gulf pipefish cluster 0 markers.** Feature plots of top markers for cluster 0
989 identified from the marker gene lists.

990 **Figure S22: Gulf pipefish cluster 1 markers.** Feature plots of top markers for cluster 1
991 identified from the marker gene lists.

992 **Figure S23: Gulf pipefish cluster 2 markers.** Feature plots of top markers for cluster 2
993 identified from the marker gene lists.

994 **Figure S24: Gulf pipefish cluster 3 markers.** Feature plots of top markers for cluster 3
995 identified from the marker gene lists.

996 **Figure S25: Gulf pipefish cluster 4 markers.** Feature plots of top markers for cluster 4
997 identified from the marker gene lists.

998 **Figure S26: Gulf pipefish cluster 5 markers.** Feature plots of top markers for cluster 5
999 identified from the marker gene lists.

1000 **Figure S27: Gulf pipefish cluster 6 markers.** Feature plots of top markers for cluster 6
1001 identified from the marker gene lists.

1002 **Figure S28: Gulf pipefish cluster 7 markers.** Feature plots of top markers for cluster 7
1003 identified from the marker gene lists.

Figure S29: Gulf pipefish cluster 8 markers. Feature plots of top markers for cluster 8 identified from the marker gene lists.

Figure S30: Gulf pipefish cluster 9 markers. Feature plots of top markers for cluster 9 identified from the marker gene lists.

Figure S31: Gulf pipefish cluster 10 markers. Feature plots of top markers for cluster 10 identified from the marker gene lists.

Figure S32: Gulf pipefish cluster 12 markers. Feature plots of top markers for cluster 12 identified from the marker gene lists.

Figure S33: Gulf pipefish cluster 13 markers. Feature plots of top markers for cluster 13 identified from the marker gene lists.

Figure S34: Gulf pipefish cluster 14 markers. Feature plots of top markers for cluster 14 identified from the marker gene lists.

Figure S35: Gulf pipefish cluster 15 markers. Feature plots of top markers for cluster 15 identified from the marker gene lists.

Figure S36: Gulf pipefish cluster 16 markers. Feature plots of top markers for cluster 16 identified from the marker gene lists.

Figure S37: Gulf pipefish cluster 17 markers. Feature plots of top markers for cluster 17 identified from the marker gene lists.

Figure S38: Gulf pipefish cluster 18 markers. Feature plots of top markers for cluster 18 identified from the marker gene lists.

Figure S39: Gulf pipefish cluster 19 markers. Feature plots of top markers for cluster 19 identified from the marker gene lists.

Figure S40: Gulf pipefish cluster 20 markers. Feature plots of top markers for cluster 20 identified from the marker gene lists.

Figure S41: Gulf pipefish cluster 21 markers. Feature plots of top markers for cluster 21 identified from the marker gene lists.

Figure S42: Gulf pipefish cluster 22 markers. Feature plots of top markers for cluster 22 identified from the marker gene lists.

Figure S43: Gulf pipefish cluster 23 markers. Feature plots of top markers for cluster 23 identified from the marker gene lists.

Figure S44: Gulf pipefish cluster 24 markers. Feature plots of top markers for cluster 24 identified from the marker gene lists.

Figure S45: Gulf pipefish cluster 25 markers. Feature plots of top markers for cluster 25 identified from the marker gene lists.

Figure S46: Gulf pipefish cluster 26 markers. Feature plots of top markers for cluster 26 identified from the marker gene lists.

Figure S47: Gulf pipefish cluster 27 markers. Feature plots of top markers for cluster 27 identified from the marker gene lists.

Figure S48: Gulf pipefish cluster 28 markers. Feature plots of top markers for cluster 28 identified from the marker gene lists.

Figure S49: Gulf pipefish cluster 29 markers. Feature plots of top markers for cluster 29 identified from the marker gene lists.

Figure S50: Gulf pipefish cluster 30 markers. Feature plots of top markers for cluster 30 identified from the marker gene lists.

Figure S51: Gulf pipefish cluster 31 markers. Feature plots of top markers for cluster 31 identified from the marker gene lists.

Figure S52: Gulf pipefish cluster 32 markers. Feature plots of top markers for cluster 32 identified from the marker gene lists.

1052 **Figure S53: Gulf pipefish cluster 33 markers.** Feature plots of top markers for cluster 33
1053 identified from the marker gene lists.

1054 **Figure S54: Gulf pipefish cluster 34 markers.** Feature plots of top markers for cluster 34
1055 identified from the marker gene lists.

1056 **Figure S55: Gulf pipefish cluster 35 markers.** Feature plots of top markers for cluster 35
1057 identified from the marker gene lists.

1058 **Figure S56: Gulf pipefish cluster 36 markers.** Feature plots of top markers for cluster 36
1059 identified from the marker gene lists.

1060 **Figure S57: Gulf pipefish cluster 37 markers.** Feature plots of top markers for cluster 37
1061 identified from the marker gene lists.

1062 **Figure S58: Marker genes define cell identity and are a resource for cell cluster**
1063 **exploration.** This dotplot shows the marker genes on the x-axis, and the cell cluster on the
1064 y-axis. The size of each dot indicates the percentage of cells that express the gene in each
1065 cluster. The darkness of each dot represents the average expression of the gene.

1066 **Figure S59: *slc25a4*, a marker for cluster 2, is expressed in muscle cells in bay**
1067 **pipefish.** Atlas expression is limited to cluster 2 (A). *In situ* hybridization experiments
1068 revealed muscle staining of *slc25a4* in bay pipefish (B-D). Wild caught bay pipefish at the
1069 onset of craniofacial elongation were used for the experiment. Embryos were mounted
1070 laterally to highlight expression domains.

1071 **Figure S60: *scg2a*, a marker for cluster 3, is expressed in the brain of bay pipefish.**
1072 Atlas expression is limited to cluster 3 (A). *In situ* hybridization experiments revealed brain
1073 staining of *scg2a* in bay pipefish (B-D). Wild caught bay pipefish at the onset of craniofacial
1074 elongation were used for the experiment. Embryos were mounted laterally to highlight
1075 expression domains.

1076 **Figure S61: *prdm16*, a marker for cluster 5, is expressed in the head and fins of bay**
1077 **pipefish.** Strong expression of prdm16 in cluster 5, osteochondrogenic mesenchyme (A). *In*
1078 *situ* hybridization experiments revealed mesenchymal staining of *prdm16* in the fins and

1079 face of bay pipefish (B-D). Wild caught bay pipefish at the onset of craniofacial elongation
1080 were used for the experiment. Embryos were mounted laterally (B, D, E) and with a dorsal
1081 view (C) to highlight expression domains.

1082 **Figure S62: *elnb*, a marker for cluster 6, is expressed in the head of bay pipefish.**
1083 Strong expression of *elnb* in cluster 6, osteochondrogenic mesenchyme (A). *In situ*
1084 hybridization experiments revealed mesenchymal staining of *elnb* in the face of bay
1085 pipefish (B-D). Wild caught bay pipefish at the onset of craniofacial elongation were used
1086 for the experiment. Embryos were mounted dorsally to showcase the craniofacial staining
1087 (B) and laterally (C-E) to show the lack of staining in the ceratohyal, dorsal fin, and caudal
1088 fin.

1089 **Figure S63: *fndc9*, a marker for cluster 7, is expressed in the brain of Gulf pipefish.**
1090 Expression of *fndc9* is limited to cluster 7, brain cells, in the atlas (A). *In situ* hybridization
1091 experiments revealed brain staining of *fndc9* in Gulf pipefish (B-C). One day post spawn
1092 Gulf pipefish larvae were used for this experiment. Embryos were mounted to highlight
1093 two views of the brain staining, one dorsal (B) and one lateral (C).

1094 **Figure S64: *insm1b*, a marker for cluster 8, is expressed in the brain of Gulf pipefish.**
1095 Strong expression of *insm1b* is observed in cluster 8, brain progenitor cells, in the atlas (A).
1096 *In situ* hybridization experiments revealed brain staining of *insm1b* in Gulf pipefish (B).
1097 Five day post spawn Gulf pipefish larvae were used for this experiment. Embryos were
1098 mounted laterally to highlight the brain staining.

1099 **Figure S65: *tnmd*, a marker for cluster 9, is expressed in tendons and ligaments of**
1100 **bay pipefish.** Strong expression of *tnmd* is observed in cluster 9, tenocyte and ligament
1101 cells, in the atlas (A). *In situ* hybridization experiments revealed tendon and ligament
1102 staining of *tnmd* in bay pipefish (B-E). Wild caught embryos from bay pipefish at the onset
1103 of craniofacial elongation were used for the experiment. Embryos were mounted laterally
1104 (B,E) and ventrally (C,D) to showcase tendon and ligament staining.

1105 **Figure S66: *myf5*, a marker for cluster 10, is expressed in the muscle of bay pipefish.**
1106 Strong expression of *myf5* is observed in cluster 10, muscle progenitor cells, in the atlas (A).

1107 *In situ* hybridization experiments revealed muscle staining of *myf5* in bay pipefish (B-D).
1108 Wild caught embryos from bay pipefish at the onset of craniofacial elongation were used
1109 for the experiment. Embryos were mounted laterally to highlight muscle staining.

1110 **Figure S67: *hpdb*, a marker for cluster 13, is expressed in somites of bay pipefish.**
1111 Strong expression of *hpbd* is observed in cluster 13, fibroblasts, in the atlas (A). *In situ*
1112 hybridization experiments revealed somite staining of *hpbd* in bay pipefish (B). Wild
1113 caught embryos from bay pipefish at the mid-somitogenesis stage were used for the
1114 experiment. Embryos were mounted laterally to highlight muscle staining.

1115 **Figure S68: *ifitm5*, a marker for cluster 18, is expressed in bay pipefish osteoblasts.**
1116 Specific expression of *ifitm5* is observed in cluster 18, osteoblasts, in the atlas (A). *In situ*
1117 hybridization experiments revealed bone staining of *ifitm5* in bay pipefish (B-E). Wild
1118 caught embryos from bay pipefish at the onset of craniofacial elongation were used for the
1119 experiment. Embryos were mounted laterally to highlight muscle staining.

1120 **Figure S69: *plaat4*, a marker for cluster 23, is weakly expressed in bay pipefish**
1121 **epidermal cells.** Specific expression of *plaat4* is observed in cluster 23, epidermal cells, in
1122 the atlas (A). *In situ* hybridization experiments revealed epidermal staining of *plaat4* in bay
1123 pipefish (B-D). Wild caught embryos from bay pipefish at the onset of craniofacial
1124 elongation were used for the experiment. Embryos were mounted ventrally (B) and
1125 laterally (C,D) to highlight epidermal staining.

1126 **Figure S70: *hyal6*, a marker for cluster 33, is strongly expressed in bay pipefish hair**
1127 **cells.** Specific expression of *hyal6* is observed in cluster 33, hair cells, in the atlas (A). *In situ*
1128 hybridization experiments revealed staining of *hyal6* in bay pipefish nares and ears (B-C).
1129 Wild caught embryos from bay pipefish from one day post spawn larvae were used for the
1130 experiment. Embryos were mounted dorsally (B) and laterally (C) to highlight staining in
1131 nares and ears respectively.

1132 **Figure S71: *soat2*, a marker for cluster 36, is strongly expressed in bay pipefish**
1133 **muscle cells.** Expression of *soat2* is observed in cluster 36, muscle cells, in the atlas (A). *In*
1134 *situ* hybridization experiments revealed staining of *soat2* in bay pipefish muscle and fins

1135 (B-E). Wild caught embryos from bay pipefish at the onset of craniofacial elongation were
1136 used for the experiment. Embryos were mounted laterally to highlight staining in muscle
1137 and fin.

1138 **Figure S72: KEGG analysis reveals pathways significantly enriched in cell clusters.**
1139 KEGG pathways are listed on the x-axis. For each cluster, the number of genes that belong
1140 to the specific KEGG pathway is represented by the height of each bar. Cell clusters are
1141 labeled by both color and their designated number. For clarity, only cell clusters with more
1142 than three annotated genes in each pathway are shown and certain pathways involving
1143 infection and metabolism are removed.

1144 **Figure S73: Clusters 8,10, and 16 are composed of largely undifferentiated cells.**
1145 Differentiation state of cell clusters was measured using CytoTRACE, where an output of 1
1146 indicates an undifferentiated cell and a value of 0 suggests a cell is fully differentiated. Cells
1147 are plotted on a violin plot with individual cell predictions shown as dots. Panels show
1148 results from muscle (A), neural (B), and connective (C) clusters.

1149 **Figure S74: Cell clusters drive gene module connectivity.** The heatmap shows the
1150 change in gene network connectivity when individual cell clusters are removed. Cell
1151 clusters are in the columns and gene modules are in rows. The coloration scale is based on
1152 the change in connecitvity when a cell cluster is removed, with yellow indicating a large
1153 change in connectivity. The modules and clusters are clustered using the Pearson distance
1154 method. Asterisks indicate cluster-module pairs that are significant using a one way
1155 permutation test and FDR correction. The size of the gene modules are shown to the right.

1156 **Figure S75: Gene module expression comparison with the Pigment Cluster reveals**
1157 **potential elusive cell types within the cluster.** Genes within the network are combined
1158 into one feature then plotted on a feature plot. In this case, UMAP dimensions one and two
1159 are the x and y-axis accordingly, dots represent each pigment cell, and the color of the dots
1160 represents the expression of the module. Cells that express gene modules are either
1161 colored red or green (when modules are plotted on separate plots). When plotted on the
1162 same plot, cells that express both gene modules are yellow. Panels A through J show each
1163 gene module comparison.

1164 **Figure S76: Conserved fin *prdm16* expression domains in stickleback and pipefish,**
1165 **but fishes have expression in different regions of the head.** The gene *prdm16*
1166 expression was probed using *in situ* hybridization. Embryos used in this assay were
1167 similarly staged (9dpf in threespine stickleback and mid-craniofacial elongation in bay
1168 pipefish). Stickleback (A-C) expressed *prdm16* in the lower jaw, brain, gill arches, and fins.
1169 Bay pipefish also have fin staining (F), but craniofacial expression is limited to the
1170 mesenchyme above the ethmoid region.

1171 **Figure S77: Conserved expression domains of *fgf22* in pipefish and stickleback fish.**
1172 Expression of *fgf22* was probed using *in situ* hybridization. Embryos used in this assay
1173 were similarly staged (9dpf threespine stickleback and 10dpf Gulf pipefish). Staining in the
1174 jaws (A,E), brain (B,F), gill arches (C,G), and fins (D, H) suggests similar \\textit{fgf22}
1175 expression domains in both species.

1176 **Figure S78: Broad *bmp4* craniofacial expression in stickleback, but *bmp4* is**
1177 **restricted to ethmoid and ceratohyal regions in pipefish.** Expression of *bmp4* was
1178 probed using *in situ* hybridization. Embryos used in this assay were similarly staged (9dpf
1179 threespine stickleback and mid-craniofacial elongation staged bay pipefish). Stickleback
1180 have *bmp4* staining in tooth germs (A), the jaws (A-B), and gill arches (C). However, *bmp4*
1181 expression is restricted in bay pipefish to mesenchyme above the ethmoid and ceratohyal.

1182 **Figure S79: *sfrp1a* has conserved fin expression in pipefish and stickleback, but**
1183 **craniofacial expression differs between the species.** Expression of *sfrp1a* was probed
1184 using *in situ* hybridization. Embryos used in this assay were similarly staged (9dpf
1185 threespine stickleback and mid-craniofacial elongation staged bay pipefish). Both species
1186 had *sfrp1a* staining in the fins (D,H). However, stickleback have broad craniofacial
1187 expression (in lower jaws, A-B, and gill arches, C) while expression is limited to
1188 mesenchyme above the ethmoid (E), the jaw joint (F), and around the ceratohyal (G) in bay
1189 pipefish.

1190 **Figure S80: *scpp* gene cluster analysis identifies most *scpp* losses are not unique to**
1191 **syngnathids.** A) denotes the Syngnathoidei lineage which contains syngnathids and blue
1192 spotted cornetfish (among other fish) to highlight that this clade contains various tooth

1193 alterations. Syngnathids lack teeth completely and cornetfish have small teeth limited to
1194 the lower jaw. B) denotes solely the syngnathid clade, which completely lacks teeth. The
1195 phylogenetic tree was time calibrated using Hughes et al. 2018 and Stiller et al. 2022.

1196 **Figure S81: 3dpf Zebrafish do not express lectin genes in epidermal cells.** This dotplot
1197 shows the marker genes on the x-axis and y-axis, organized by cell type. The size of the dot
1198 represents the percentage of cells that express the gene in each cluster. The darkness of the
1199 dot represents the average expression of the genes.

1200 Supplementary file captions:

1201 **SupplementaryFile1_scRNAseq_cellranger_info.csv,** a table containing quality metrics
1202 for the single cell libraries.

1203 **SupplementaryFile2_cell_cluster_identity_and_sizes. csv**, a table containing the number
1204 of cells in each cell cluster and cluster identities.

1205 **SupplementaryFile3_seurat_findallmarker_output.csv**, a file containing marker genes
1206 identified using Seurat FindAllMarkers for each cluster.

1207 **SupplementaryFile4_dotplotmethod_marker_genes.csv**, a file containing marker genes
1208 identified using the DotPlot method.

1209 **SupplementaryFile5_ select_marker_gene_detailed_information.xlsx**, a file containing
1210 additional information on the marker gene identified for every cluster.

1211 **SupplementaryFile6_insitu_probe_sequences_and_prep_information.xlsx,** a list of the
1212 *in situ* hybridization probes used in this study, the conditions used to prepare the probes,
1213 and the staging/sample information for the embryos.

1214 **SupplementaryFile7_insitu_probe_alignment.zip**, a zipped folder containing alignments
1215 of *in situ* hybridization probes with the unpublished *Syngnathus leptorhynchus* genome.

1216 **SupplementaryFile8_WGCNA_module_number_to_color_translation_table.csv**, genetic
1217 networks were initially labeled with colors, we converted these labels to numeric
1218 annotations for simplicity using this conversion table. The table also contains the number
1219 of genes in each network.

1220 **SupplementaryFile9_WGCNA_networks_with_gene_descriptions.xlsx**, a file containing
1221 all the genetic networks, the genes inside of them, and additional information for the
1222 genetic networks highlighted in this paper.

1223 **SupplementaryFile10_WGCNA_mod_association_wClus_tstat.csv**, this file contains the t-
1224 statistics derived for each module-cell cluster pair. The cell clusters are in the rows and the
1225 gene modules are in the columns.

1226 **SupplementaryFile11_WGCNA_associat_mods_w_clus_tstat_pvalues_corrected.csv,** p-
1227 values for the t-statistics of the strength of association between gene modules and cell
1228 clusters. P-values are corrected for multiple testing hypotheses using fdr. The cell clusters
1229 are in the rows and the gene modules are in the columns.

1230 **SupplementaryFile12_WGCNA_connectivity_change.csv,** the change in connectivity for
1231 gene modules when individual cell clusters are removed. The cell clusters are in the
1232 columns and the gene networks are in the rows.

1233 **SupplementaryFile13_Connectivity_Permutation_Pvalues_Corrected.csv,** the p-values
1234 for the change in connectivity are found in this csv. The cell clusters are in the columns and
1235 the gene networks are in rows.

1236 ## References

1237 Ahi, E. P., Kapralova, K. H., Pálsson, A., Maier, V. H., Gudbrandsson, J., Snorrason, S. S.,
1238 Jónsson, Z. O., & Franzdóttir, S. R. (2014). Transcriptional dynamics of a conserved
1239 gene expression network associated with craniofacial divergence in Arctic charr.
1240 *EvoDevo*, *5*(1), 1–19.

1241 Ahi, E. P., Lecaudey, L. A., Ziegelbecker, A., Steiner, O., Glabonjat, R., Goessler, W., Hois, V.,
1242 Wagner, C., Lass, A., & Sefc, K. M. (2020). Comparative transcriptomics reveals
1243 candidate carotenoid color genes in an East African cichlid fish. *BMC Genomics 2020*
1244 *21:1*, *21*(1), 1–15.

1245 Anderson, A. P., & Jones, A. G. (2019). Choosy Gulf pipefish males ignore age but prefer
1246 active females with deeply keeled bodies. *Animal Behaviour*, *155*, 37–44.

1247 Arendt, D., Musser, J. M., H Baker, C. V, Bergman, A., Cepko, C., Erwin, D. H., Pavlicev, M.,
1248 Schlosser, G., Widder, S., Laubichler, M. D., & Wagner, G. P. (2016). *The origin and*
1249 *evolution of cell types*.

1250 Bergen, D. J. M., Tong, Q., Shukla, A., Newham, E., Zethof, J., Lundberg, M., Ryan, R., Youlten,
1251 S. E., Frysz, M., Croucher, P. I., Flik, G., Richardson, R. J., Kemp, J. P., Hammond, C. L., &
1252 Metz, J. R. (2022). Regenerating zebrafish scales express a subset of evolutionary
1253 conserved genes involved in human skeletal disease. *BMC Biology 2021 20:1*, *20*(1), 1–
1254 25.

1255 Bergert, B. A., & Wainwright, P. C. (1997). Morphology and kinematics of prey capture in
1256 the syngnathid fishes Hippocampus erectus and Syngnathus floridae. *Marine Biology*,
1257 *127*(4), 563–570.

1258 Berglund, A., Rosenqvist, G., & Svensson, I. (1986). *Reversed sex roles and parental energy*
1259 *investment in zygotes of two pipefish (Syngnathidae) species*. *29*, 209–215.

1260 Bjork, B. C., Turbe-Doan, A., Prysak, M., Herron, B. J., & Beier, D. R. (2010). Prdm16 is
1261 required for normal palatogenesis in mice. *Human Molecular Genetics*, *19*(5), 774–789.

1262 Brown, G. D., Willment, J. A., & Whitehead, L. (2018). C-type lectins in immunity and
1263 homeostasis. *Nature Reviews Immunology*, *18*(6), 374–380.

1264 Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell
1265 transcriptomic data across different conditions, technologies, and species. *Nature*
1266 *Biotechnology*, *36*(5), 411–420.

1267 Cãestro, C., Yokoi, H., & Postlethwait, J. H. (2007). Evolutionary developmental biology and
1268 genomics. In *Nature Reviews Genetics* (Vol. 8, Issue 12, pp. 932–942). Nature
1269 Publishing Group.

1270 Carcupino, M. (2002). Functional significance of the male brood pouch in the reproductive
1271 strategies of pipefishes and seahorses: a morphological and ultrastructural

1272 comparative study on three anatomically different pouches. *Journal of Fish Biology*,
1273 *61*(6), 1465–1480.

1274 Chari, T., Weissbourd, B., Gehring, J., Ferraioli, A., Leclère, L., Herl, M., Gao, F., Chevalier, S.,
1275 Copley, R. R., Houliston, E., Anderson, D. J., & Pachter, L. (2021). Whole-animal
1276 multiplexed single-cell RNA-seq reveals transcriptional shifts across Clytia medusa cell
1277 types. *Science Advances*, *7*(48), 1683.

1278 Chen, J. W., & Galloway, J. L. (2014). The development of zebrafish tendon and ligament
1279 progenitors. *Development*, *141*(10), 2035–2045.

1280 Chen, Y. P., Zhang, Y., Jiang, T. X., Barlow, A. J., St. Amand, T. R., Hu, Y., Heaney, S., Francis-
1281 West, P., Chuong, C. M., & Maas, R. (2000). Conservation of early odontogenic signaling
1282 pathways in Aves. *Proceedings of the National Academy of Sciences*, *97*(18), 10044–
1283 10049.

1284 Crump, J. G., Maves, L., Lawson, N. D., Weinstein, B. M., & Kimmel, C. B. (2004). An essential
1285 role for Fgfs in endodermal pouch formation influences later craniofacial skeletal
1286 patterning. *Development*, *131*(22), 5703–5716.

1287 Denisenko, E., Guo, B. B., Jones, M., Hou, R., De Kock, L., Lassmann, T., Poppe, D., Poppe, D.,
1288 Clément, O., Simmons, R. K., Simmons, R. K., Lister, R., & Forrest, A. R. R. (2020).
1289 Systematic assessment of tissue dissociation and storage biases in single-cell and
1290 single-nucleus RNA-seq workflows. *Genome Biology*, *21*(1), 1–25.

1291 Ding, H. L., Clouthier, D. E., & Artinger, K. B. (2013). Redundant roles of PRDM family
1292 members in zebrafish craniofacial development. *Developmental Dynamics*, *242*(1), 67–
1293 79.

1294 Docheva, D., Hunziker, E. B., Fässler, R., & Brandau, O. (2005). Tenomodulin Is Necessary for
1295 Tenocyte Proliferation and Tendon Maturation. *Molecular and Cellular Biology*, *25*(2),
1296 699–705.

1297 Du, J., Miller, A. J., Widlund, H. R., Horstmann, M. A., Ramaswamy, S., & Fisher, D. E. (2003).
1298 MLANA/MART1 and SILV/PMEL17/GP100 Are Transcriptionally Regulated by MITF
1299 in Melanocytes and Melanoma. *The American Journal of Pathology*, *163*(1), 333–343.

1300 Dudley, J. S., Hannaford, P., Dowland, S. N., Lindsay, L. A., Thompson, M. B., Murphy, C. R.,
1301 Van Dyke, J. U., & Whittington, C. M. (2021). Structural changes to the brood pouch of
1302 male pregnant seahorses (Hippocampus abdominalis) facilitate exchange between
1303 father and embryos. *Placenta*, *114*, 115–123.

1304 Fabian, P., Tseng, K.-C., Thiruppathy, M., Arata, C., Chen, H.-J., Smeeton, J., Nelson, N., &
1305 Crump, J. G. (2022). Lifelong single-cell profiling of cranial neural crest diversification
1306 in zebrafish. *Nature Communications 2022 13:1*, *13*(1), 1–13.

1307 Farnsworth, D. R., Saunders, L. M., & Miller, A. C. (2020). A single-cell transcriptome atlas
1308 for zebrafish development. *Developmental Biology*, *459*(2), 100–108.

Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science*, *360*(6392).

Feregrino, C., Sacher, F., Parnas, O., & Tschopp, P. (2019). A single-cell transcriptomic atlas of the developing chicken limb. *BMC Genomics*, *20*(1), 1–15.

Flammang, B. E., Ferry-Graham, L. A., Rinewalt, C., Ardizzone, D., Davis, C., & Trejo, T. (2009). Prey capture kinematics and four-bar linkages in the bay pipefish, Syngnathus leptorhynchus. *Zoology*, *112*(2), 86–96.

Fuess, L. E., & Bolnick, D. I. (2023). GBE Single-Cell RNA Sequencing Reveals Microevolution of the Stickleback Immune System. *Genome Biology and Evolution*, *15*.

Fuiten, A. M., & Cresko, W. A. (2021). Evolutionary divergence of a Hoxa2b hindbrain enhancer in syngnathids mimics results of functional assays. *Development Genes and Evolution*, *231*(3–4), 57–71.

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

Goldsmith, M. I., Fisher, S., Waterman, R., & Johnson, S. L. (2003). Saltatory control of isometric growth in the zebrafish caudal fin is disrupted in long fin and rapunzel mutants. *Developmental Biology*, *259*(2), 303–317.

Gulati, G. S., Sikandar, S. S., Wesche, D. J., Manjunath, A., Bharadwaj, A., Berger, M. J., Ilagan, F., Kuo, A. H., Hsieh, R. W., Cai, S., Zabala, M., Scheeren, F. A., Lobo, N. A., Qian, D., Yu, F. B., Dirbas, F. M., Clarke, M. F., & Newman, A. M. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*, *367*(6476), 405–411.

Hafemeister, C., & Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology*, *20*(1), 296.

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

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

Harlin-Cognato, A., Hoffman, E. A., & Jones, A. G. (2006). Gene cooption without duplication during the evolution of a male-pregnancy gene in pipefish. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(51), 19407–19412.

Healey, H. M., Bassham, S., & Cresko, W. A. (2022). Single-cell Iso-Sequencing enables rapid genome annotation for scRNAseq analysis. *Genetics*, *220*(3).

1346 Hong, X., Wang, Y., Wang, K., Wei, C., Li, W., Yu, L., Xu, H., Zhu, J., Zhu, X., & Liu, X. (2023).
1347 Single-Cell Atlas Reveals the Hemocyte Subpopulations and Stress Responses in Asian
1348 Giant Softshell Turtle during Hibernation. *Biology*, *12*(7), 994.

1349 Howe, D. G., Bradford, Y. M., Conlin, T., Eagle, A. E., Fashena, D., Frazer, K., Knight, J., Mani, P.,
1350 Martin, R., Moxon, S. A. T., Paddock, H., Pich, C., Ramachandran, S., Ruef, B. J., Ruzicka,
1351 L., Schaper, K., Shao, X., Singer, A., Sprunger, B., … Westerfield, M. (2013). ZFIN, the
1352 Zebrafish Model Organism Database: Increased support for mutants and transgenics.
1353 *Nucleic Acids Research*, *41*(D1), D854–D860.

1354 Hughes, L. C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C. C., Thompson, A. W., Arcila, D., Betancur,
1355 R., Li, C., Becker, L., Bellora, N., Zhao, X., Li, X., Wang, M., Fang, C., Xie, B., Zhoui, Z.,
1356 Huang, H., Chen, S., … Shi, Q. (2018). Comprehensive phylogeny of ray-finned fishes
1357 (Actinopterygii) based on transcriptomic and genomic data. *Proceedings of the*
1358 *National Academy of Sciences of the United States of America*, *115*(24), 6249–6254.

1359 Jain, I., Stroka, C., Yan, J., Huang, W. M., & Iovine, M. K. (2007). Bone growth in zebrafish fins
1360 occurs via multiple pulses of cell proliferation. *Developmental Dynamics*, *236*(9), 2668–
1361 2674.

1362 Johnson, S. L., Nguyen, A. T. N., & Lister, J. A. (2011). mitfa is required at multiple stages of
1363 melanocyte differentiation but not to establish the melanocyte stem cell.
1364 *Developmental Biology*, *350*(2), 405–413.

1365 Jungerson, H. F. E. (1910). *Ichthyotomical Contributions: The structure of the Aulostomidae,*
1366 *Syngnathidae, and Solenostomidae*.

1367 Kaneda-Nakashima, K., Igawa, K., Suwanruengsri, M., Naoyuki, F., Ichikawa, T., Funamoto,
1368 T., Kurogi, S., Sekimoto, T., Yamashita, Y., Chosa, E., Yamaguchi, R., & Morishita, K.
1369 (2022). Role of Mel1/Prdm16 in bone differentiation and morphology. *Experimental*
1370 *Cell Research*, *410*(2), 112969.

1371 Kawasaki, K. (2009). The SCPP gene repertoire in bony vertebrates and graded differences
1372 in mineralized tissues. *Development Genes and Evolution*, *219*(3), 147–157.

1373 Kimura, T., Takehana, Y., & Naruse, K. (2017). Pnp4a Is the causal gene of the medaka
1374 iridophore Mutant guanineless. *G3: Genes, Genomes, Genetics*, *7*(4), 1357–1363.

1375 Koiwai, K., Kondo, H., & Hirono, I. (2023). scRNA-seq Analysis of Hemocytes of Penaeid
1376 Shrimp Under Virus Infection. *Marine Biotechnology*, *25*(3), 488–502.

1377 Kuo, R. I., Tseng, E., Eory, L., Paton, I. R., Archibald, A. L., & Burt, D. W. (2017). Normalized
1378 long read RNA sequencing in chicken reveals transcriptome complexity similar to
1379 human. *BMC Genomics*, *18*(1), 1–19.

1380 Kvarnemo, C., Mobley, K. B., Partridge, C., Jones, A. G., & Ahnesjö, I. (2011). Evidence of
1381 paternal nutrient provisioning to embryos in broad-nosed pipefish Syngnathus typhle.
1382 *Journal of Fish Biology*, *78*(6), 1725–1737.

1383 Lamason, R. L., Mohideen, M. A. P. K., Mest, J. R., Wong, A. C., Norton, H. L., Aros, M. C.,
1384 Jurynec, M. J., Mao, X., Humphreville, V. R., Humbert, J. E., Sinha, S., Moore, J. L.,
1385 Jagadeeswaran, P., Zhao, W., Ning, G., Makalowska, I., McKeigue, P. M., O'Donnell, D.,
1386 Kittles, R., … Cheng, K. C. (2005). Genetics: SLC24A5, a putative cation exchanger,
1387 affects pigmentation in zebrafish and humans. *Science*, *310*(5755), 1782–1786.

1388 Lange, M., Granados, A., VijayKumar, S., Bragantini, J., Ancheta, S., Santhosh, S., Borja, M.,
1389 Kobayashi, H., McGeever, E., Solak, A. C., Yang, B., Zhao, X., Liu, Y., Detweiler, A. M., Paul,
1390 S., Mekonen, H., Lao, T., Banks, R., Kim, Y.-J., … Royer, L. A. (2023). Zebrahub –
1391 Multimodal Zebrafish Developmental Atlas Reveals the State-Transition Dynamics of
1392 Late-Vertebrate Pluripotent Axial Progenitors. *BioRxiv*, 2023.03.06.531398.

1393 Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation
1394 network analysis. *BMC Bioinformatics*, *9*(1), 1–13.

1395 Leerberg, D. M., Hopton, R. E., & Draper, B. W. (2019). Fibroblast growth factor receptors
1396 function redundantly during zebrafish embryonic development. *Genetics*, *212*(4),
1397 1301–1319.

1398 Leysen, H., Roos, G., & Adriaens, D. (2011). Morphological variation in head shape of
1399 pipefishes and seahorses in relation to snout length and developmental growth.
1400 *Journal of Morphology*, *272*(10), 1259–1270.

1401 Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*,
1402 *34*(18), 3094–3100.

1403 Lin, Q., Fan, S., Zhang, Y., Xu, M., Zhang, H., Yang, Y., Lee, A. P., Woltering, J. M., Ravi, V.,
1404 Gunter, H. M., Luo, W., Gao, Z., Lim, Z. W., Qin, G., Schneider, R. F., Wang, X., Xiong, P., Li,
1405 G., Wang, K., … Venkatesh, B. (2016). The seahorse genome and the evolution of its
1406 specialized morphology. *Nature*, *540*(7633), 395–399.

1407 Liu, Z., Liu, S., Yao, J., Bao, L., Zhang, J., Li, Y., Jiang, C., Sun, L., Wang, R., Zhang, Y., Zhou, T.,
1408 Zeng, Q., Fu, Q., Gao, S., Li, N., Koren, S., Jiang, Y., Zimin, A., Xu, P., … Waldbieser, G. C.
1409 (2016). The channel catfish genome sequence provides insights into the evolution of
1410 scale formation in teleosts. *Nature Communications*, *7*.

1411 Manning, C. G., Foster, S. J., & Vincent, A. C. J. (2019). A review of the diets and feeding
1412 behaviours of a family of biologically diverse marine fishes (Family Syngnathidae).
1413 *Reviews in Fish Biology and Fisheries 2019 29:2*, *29*(2), 197–221.

1414 McCarthy, N., Sidik, A., Bertrand, J. Y., & Eberhart, J. K. (2016). An Fgf-Shh signaling
1415 hierarchy regulates early specification of the zebrafish skull. *Developmental Biology*,
1416 *415*(2), 261–277.

1417 Melamed, P., Xue, Y., Poon, J. F. D., Wu, Q., Xie, H., Yeo, J., Foo, T. W. J., & Chua, H. K. (2005).
1418 The male seahorse synthesizes and secretes a novel C-type lectin into the brood pouch
1419 during early pregnancy. *The FEBS Journal*, *272*(5), 1221–1235.

1420 Miao, M., Bruce, A. E. E., Bhanji, T., Davis, E. C., & Keeley, F. W. (2007). Differential
1421 expression of two tropoelastin genes in zebrafish. *Matrix Biology*, *26*(2), 115–124.

1422 Miyake, A., & Itoh, N. (2013). Fgf22 regulated by Fgf3/Fgf8 signaling is required for
1423 zebrafish midbrain development. *Biology Open*, *2*(5), 515–524.

1424 Mori, S., & Nakamura, T. (2022). Redeployment of odontode gene regulatory network
1425 underlies dermal denticle formation and evolution in suckermouth armored catfish.
1426 *Scientific Reports 2022 12:1*, *12*(1), 1–16.

1427 Moriyama, Y., Ito, F., Takeda, H., Yano, T., Okabe, M., Kuraku, S., Keeley, F. W., & Koshiba-
1428 Takeuchi, K. (2016). Evolution of the fish heart by sub/neofunctionalization of an
1429 elastin gene. *Nature Communications 2016 7:1*, *7*(1), 1–10.

1430 Morrison, J. A., McLennan, R., Wolfe, L. A., Gogol, M. M., Meier, S., McKinney, M. C., Teddy, J.
1431 M., Holmes, L., Semerad, C. L., Box, A. C., Li, H., Hall, K. E., Perera, A. G., & Kulesa, P. M.
1432 (2017). Single-cell transcriptome analysis of avian neural crest migration reveals
1433 signatures of invasion and molecular transitions. *ELife*, *6*.

1434 Parker, J., Guslund, N. C., Jentoft, S., & Roth, O. (2022). Characterization of Pipefish Immune
1435 Cell Populations Through Single-Cell Transcriptomics. *Frontiers in Immunology*, *13*,
1436 820152.

1437 Partridge, C., Shardo, J., & Boettcher, A. (2007). Osmoregulatory role of the brood pouch in
1438 the euryhaline Gulf pipefish, Syngnathus scovelli. *Comparative Biochemistry and*
1439 *Physiology - A Molecular and Integrative Physiology*, *147*(2 SPEC. ISS.), 556–561.

1440 Potts, H. G., Lemieux, M. E., Rice, E. S., Warren, W., Choudhury, R. P., & Mommersteeg, M. T.
1441 M. (2022). Discordant Genome Assemblies Drastically Alter the Interpretation of
1442 Single-Cell RNA Sequencing Data Which Can Be Mitigated by a Novel Integration
1443 Method. *Cells*, *11*(4), 608.

1444 Qu, M., Liu, Y., Zhang, Y., Wan, S., Ravi, V., Qin, G., Jiang, H., Wang, X., Zhang, H., Zhang, B.,
1445 Gao, Z., Huysseune, A., Zhang, Z., Zhang, H., Chen, Z., Yu, H., Wu, Y., Tang, L., Li, C., … Lin,
1446 Q. (2021). Seadragon genome analysis provides insights into its phenotype and sex
1447 determination locus. *Science Advances*, *7*(34), eabg5196.

1448 Ramesh, B., Small, C. M., Healey, H., Johnson, B., Barker, E., Currey, M., Bassham, S., Myers,
1449 M., Cresko, W. A., & Jones, A. G. (2023). Improvements to the Gulf pipefish Syngnathus
1450 scovelli genome. *GigaByte*, *2023*.

1451 Ripley, J. L., & Foran, C. M. (2006). Differential parental nutrient allocation in two
1452 congeneric pipefish species (Syngnathidae: Syngnathus spp.). *Journal of Experimental*
1453 *Biology*, *209*(6), 1112–1121.

1454 Ripley, J. L., Williams, P. S., & Foran, C. M. (2010). Morphological and quantitative changes in
1455 paternal brood-pouch vasculature during embryonic development in two syngnathus
1456 pipefishes. *Journal of Fish Biology*, *77*, 67–79.

1457 Rose, E., Johnson, B. D., Armitage, V., & Jones, A. G. (2023). The response of the brood pouch
1458 transcriptome to synthetic estrogen exposure in the Gulf pipefish (Syngnathus
1459 scovelli). *Frontiers in Marine Science*, *10*(June), 1–12.

1460 Roth, O., Klein, V., Beemelmanns, A., Scharsack, J. P., & Reusch, T. B. H. (2012). Male
1461 pregnancy and biparental immune priming. *American Naturalist*, *180*(6), 802–814.

1462 Roth, O., Solbakken, M. H., Tørresen, O. K., Bayer, T., Matschiner, M., Baalsrud, H. T., Hoff, S.
1463 N. K., Brieuc, M. S. O., Haase, D., Hanel, R., Reusch, T. B. H., & Jentoft, S. (2020).
1464 Evolution of male pregnancy associated with remodeling of canonical vertebrate
1465 immunity in seahorses and pipefishes. *Proceedings of the National Academy of Sciences*,
1466 201916251.

1467 Royan, M. R., Siddique, K., Csucs, G., Puchades, M. A., Nourizadeh-Lillabadi, R., Bjaalie, J. G.,
1468 Henkel, C. V., Weltzien, F. A., & Fontaine, R. (2021). 3D Atlas of the Pituitary Gland of
1469 the Model Fish Medaka (Oryzias latipes). *Frontiers in Endocrinology*, *12*, 719843.

1470 Salamanca-Díaz, D. A., Schulreich, S. M., Cole, A. G., & Wanninger, A. (2022). Single-Cell RNA
1471 Sequencing Atlas From a Bivalve Larva Enhances Classical Cell Lineage Studies.
1472 *Frontiers in Ecology and Evolution*, *9*, 783984.

1473 Schneider, R. F., Gunter, H. M., Salewski, I., Woltering, J. M., & Meyer, A. (2023). Growth
1474 dynamics and molecular bases of evolutionary novel jaw extensions in halfbeaks and
1475 needlefishes (Beloniformes). *Molecular Ecology*, *32*(21), 5798–5811.

1476 Schneider, R. F., Woltering, J. M., Adriaens, D., & Roth, O. (2023). A comparative analysis of
1477 the ontogeny of syngnathids (pipefishes and seahorses) reveals how heterochrony
1478 contributed to their diversification. *Developmental Dynamics*, *252*(5), 553–588.

1479 Shema, E., Bernstein, B. E., & Buenrostro, J. D. (2019). Single-cell and single-molecule
1480 epigenomics to uncover genome regulation at unprecedented resolution. *Nature*
1481 *Genetics*, *51*(1), 19–25.

1482 Shull, L. C., Lencer, E. S., Kim, H. M., Goyama, S., Kurokawa, M., Costello, J. C., Jones, K., &
1483 Artinger, K. B. (2022). PRDM paralogs antagonistically balance Wnt/β-catenin activity
1484 during craniofacial chondrocyte differentiation. *Development (Cambridge, England)*,
1485 *149*(4).

1486 Shull, L. C., Sen, R., Menzel, J., Goyama, S., Kurokawa, M., & Artinger, K. B. (2020). The
1487 conserved and divergent roles of Prdm3 and Prdm16 in zebrafish and mouse
1488 craniofacial development. *Developmental Biology*, *461*(2), 132–144.

1489 Sire, J. Y., Delgado, S. C., & Girondot, M. (2008). Hen's teeth with enamel cap: From dream to
1490 impossibility. *BMC Evolutionary Biology*, *8*(1), 1–11.

1491 Small, C. M., Bassham, S., Catchen, J., Amores, A., Fuiten, A. M., Brown, R. S., Jones, A. G., &
1492 Cresko, W. A. (2016). The genome of the Gulf pipefish enables understanding of
1493 evolutionary innovations. *Genome Biology*, *17*(1), 258.

1494 Small, C. M., Harlin-Cognato, A. D., & Jones, A. G. (2013). Functional similarity and molecular
1495 divergence of a novel reproductive transcriptome in two male-pregnant Syngnathus
1496 pipefish species. *Ecology and Evolution*, *3*(12), 4092–4108.

1497 Small, C. M., Healey, H. M., Currey, M. C., Beck, E. A., Catchen, J., Lin, A. S. P., Cresko, W. A., &
1498 Bassham, S. (2022). Leafy and weedy seadragon genomes connect genic and repetitive
1499 DNA features to the extravagant biology of syngnathid fishes. *Proceedings of the*
1500 *National Academy of Sciences of the United States of America*, *119*(26), e2119602119.

1501 Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L.,
1502 Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino, M. L.,
1503 Boismoreau, F., Hilscher, M. M., Yokota, C., Qian, X., Nilsson, M., Bronner, M. E., …
1504 Adameyko, I. (2019). Spatiotemporal structure of cell fate decisions in murine neural
1505 crest. *Science*, *364*(6444).

1506 Sommer, S., Whittington, C. M., & Wilson, A. B. (2012). Standardised classification of pre-
1507 release development in male-brooding pipefish, seahorses, and seadragons (Family
1508 Syngnathidae). *BMC Developmental Biology*, *12*, 12–15.

1509 Songco-Casey, J. O., Coffing, G. C., Piscopo, D. M., Pungor, J. R., Kern, A. D., Miller, A. C., & Niell,
1510 C. M. (2022). Cell types and molecular architecture of the Octopus bimaculoides visual
1511 system. *Current Biology : CB*, *32*(23), 5031.

1512 Steger, J., Cole, A. G., Denner, A., Lebedeva, T., Genikhovich, G., Ries, A., Reischl, R., Taudes, E.,
1513 Lassnig, M., & Technau, U. (2022). Single-cell transcriptomics identifies conserved
1514 regulators of neuroglandular lineages. *Cell Reports*, *40*(12).

1515 Stiller, J., Short, G., Hamilton, H., Saarman, N., Longo, S., Wainwright, P., Rouse, G. W., &
1516 Simison, W. B. (2022). Phylogenomic analysis of Syngnathidae reveals novel
1517 relationships, origins of endemic diversity and variable diversification rates. *BMC*
1518 *Biology 2022 20:1*, *20*(1), 1–21.

1519 Stock, D. W., Jackman, W. R., & Trapani, J. (2006). Developmental genetic mechanisms of
1520 evolutionary tooth loss in cypriniform fishes. *Development*, *133*(16), 3127–3137.

1521 Stölting, K. N., & Wilson, A. B. (2007). Male pregnancy in seahorses and pipefish: Beyond
1522 the mammalian model. *BioEssays*, *29*(9), 884–896.

1523 Szabo-Rogers, H. L., Geetha-Loganathan, P., Nimmagadda, S., Fu, K. K., & Richman, J. M.
1524 (2008). FGF signals from the nasal pit are necessary for normal facial morphogenesis.
1525 *Developmental Biology*, *318*(2), 289–302.

1526 Tardaguila, M., Fuente, L. de la, Marti, C., Pereira, C., Pardo-Palacios, F. J., Risco, H. del,
1527 Ferrell, M., Mellado, M., Macchietto, M., Verheggen, K., Edelmann, M., Ezkurdia, I.,
1528 Vazquez, J., Tress, M., Mortazavi, A., Martens, L., Rodriguez-Navarro, S., Moreno-
1529 Manzano, V., & Conesa, A. (2018). SQANTI: Extensive characterization of long-read
1530 transcript sequences for quality control in full-length transcriptome identification and
1531 quantification. *Genome Research*, *28*(3), 396.

Teichmann, S. A., & Babu, M. M. (2004). Gene regulatory network growth by duplication. *Nature Genetics*, *36*(5), 492–496.

Tengfei, D., Yongshuang, X., Haixia, Z., Li, Z., Qinghua, L., Xueying, W., Jun, L., Shihong, X., Yanfeng, W., Jiachen, Y., Lele, W., Yunong, W., & Guang, G. (2021). Multiple Fetal Nutritional Patterns Before Parturition in Viviparous Fish Sebastes schlegelii (Hilgendorf, 1880). *Frontiers in Marine Science*, *7*, 571946.

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

Thisse, C., & Thisse, B. (2007). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nature Protocols 2008 3:1*, *3*(1), 59–69.

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.

Ton, M. L. 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).

Ton, M. L. N., 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.

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.

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.

Uniken Venema, W. T. C., Ramírez-Sánchez, A. D., Bigaeva, E., Withoff, S., Jonkers, I., McIntyre, R. E., Ghouraba, M., Raine, T., Weersma, R. K., Franke, L., Festen, E. A. M., & van der Wijst, M. G. P. (2022). Gut mucosa dissociation protocols influence cell type proportions and single-cell gene expression levels. *Scientific Reports 2022 12:1*, *12*(1), 1–14.

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

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

1569 Verreijdt, L., Debiais-Thibaud, M., Borday-Birraux, V., Van Der Heyden, C., Sire, J. Y., &
1570 Huysseune, A. (2006). Expression of the dlx gene family during formation of the
1571 cranial bones in the zebrafish (Danio rerio): Differential involvement in the visceral
1572 skeleton and braincase. *Developmental Dynamics*, *235*(5), 1371–1389.

1573 Vonk, A. C., Zhao, X., Pan, Z., Hudnall, M. L., Oakes, C. G., Lopez, G. A., Hasel-Kolossa, S. C.,
1574 Kuncz, A. W. C., Sengelmann, S. B., Gamble, D. J., & Lozito, T. P. (2023). Single-cell
1575 analysis of lizard blastema fibroblasts reveals phagocyte-dependent activation of
1576 Hedgehog-responsive chondrogenesis. *Nature Communications 2023 14:1*, *14*(1), 1–20.

1577 Wagner, A. (2011). The molecular origins of evolutionary innovations. *Trends in Genetics*,
1578 *27*(10), 397–410.

1579 Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., & Klein, A. M. (2018).
1580 Single-cell mapping of gene expression landscapes and lineage in the zebrafish
1581 embryo. *Science*, *360*(6392), 981–987.

1582 Wagner, G. P., & Lynch, V. J. (2010). Evolutionary novelties. In *Current Biology* (Vol. 20,
1583 Issue 2, pp. R48–R52). Cell Press.

1584 Walshe, J., & Mason, I. (2003). Fgf signalling is required for formation of cartilage in the
1585 head. *Developmental Biology*, *264*(2), 522–536.

1586 Wang, W., Yang, N., Wang, L., Zhu, Y., Chu, X., Xu, W., Li, Y., Xu, Y., Gao, L., Zhang, B., Zhang, G.,
1587 Sun, Q., Wang, W., Wang, Q., Zhang, W., & Chen, D. (2024). The TET-Sall4-BMP
1588 regulatory axis controls craniofacial cartilage development. *Cell Reports*, *43*(3),
1589 113873.

1590 Wetzel, J. T., & Wourms, J. P. (2004). Embryogenesis in the Dwarf Seahorse, Hippocampus
1591 zosterae (Syngnathidae). *Gulf and Caribbean Research*, *16*(1), 27–35.

1592 Whittington, C. M., & Friesen, C. R. (2020). The evolution and physiology of male pregnancy
1593 in syngnathid fishes. *Biological Reviews*, *95*(5), 1252–1272.

1594 Whittington, C. M., Griffith, O. W., Qi, W., Thompson, M. B., & Wilson, A. B. (2015a). Seahorse
1595 Brood Pouch Transcriptome Reveals Common Genes Associated with Vertebrate
1596 Pregnancy. *Molecular Biology and Evolution*, *32*(12), 3114–3131.

1597 Whittington, C. M., Griffith, O. W., Qi, W., Thompson, M. B., & Wilson, A. B. (2015b). Seahorse
1598 Brood Pouch Transcriptome Reveals Common Genes Associated with Vertebrate
1599 Pregnancy. *Molecular Biology and Evolution*, *32*(12), msv177.

1600 Williams, R. M., Candido-Ferreira, I., Repapi, E., Gavriouchkina, D., Senanayake, U., Ling, I. T.
1601 C., Telenius, J., Taylor, S., Hughes, J., & Sauka-Spengler, T. (2019a). Reconstruction of
1602 the Global Neural Crest Gene Regulatory Network In Vivo. *Developmental Cell*, *51*(2),
1603 255-276.e7.

1604 Williams, R. M., Candido-Ferreira, I., Repapi, E., Gavriouchkina, D., Senanayake, U., Ling, I. T.
1605 C., Telenius, J., Taylor, S., Hughes, J., & Sauka-Spengler, T. (2019b). Reconstruction of

the Global Neural Crest Gene Regulatory Network In Vivo. *Developmental Cell*, *51*(2), 255-276.e7.

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.

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).

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

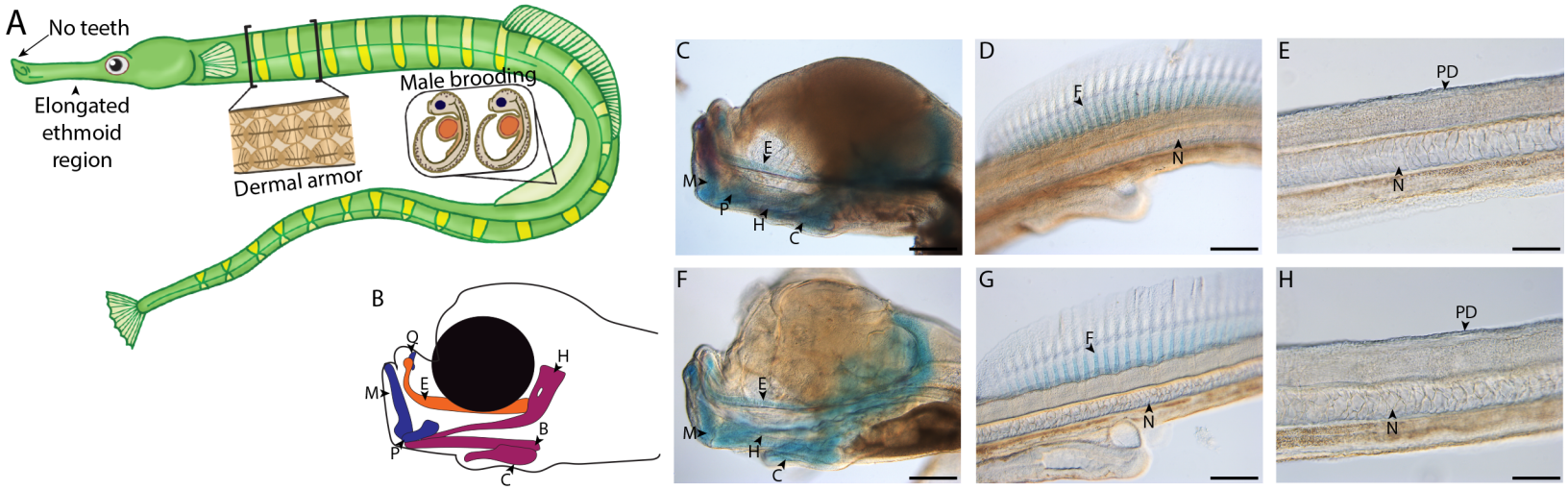
Woych, J., Gurrola, A. O., Deryckere, A., Jaeger, E. C. B., Gumnit, E., Merello, G., Gu, J., Araus, A. J., Leigh, N. D., Yun, M., Simon, A., & Tosches, M. A. (2022). Cell-type profiling in salamanders identifies innovations in vertebrate forebrain evolution. *Science*, *377*(6610).

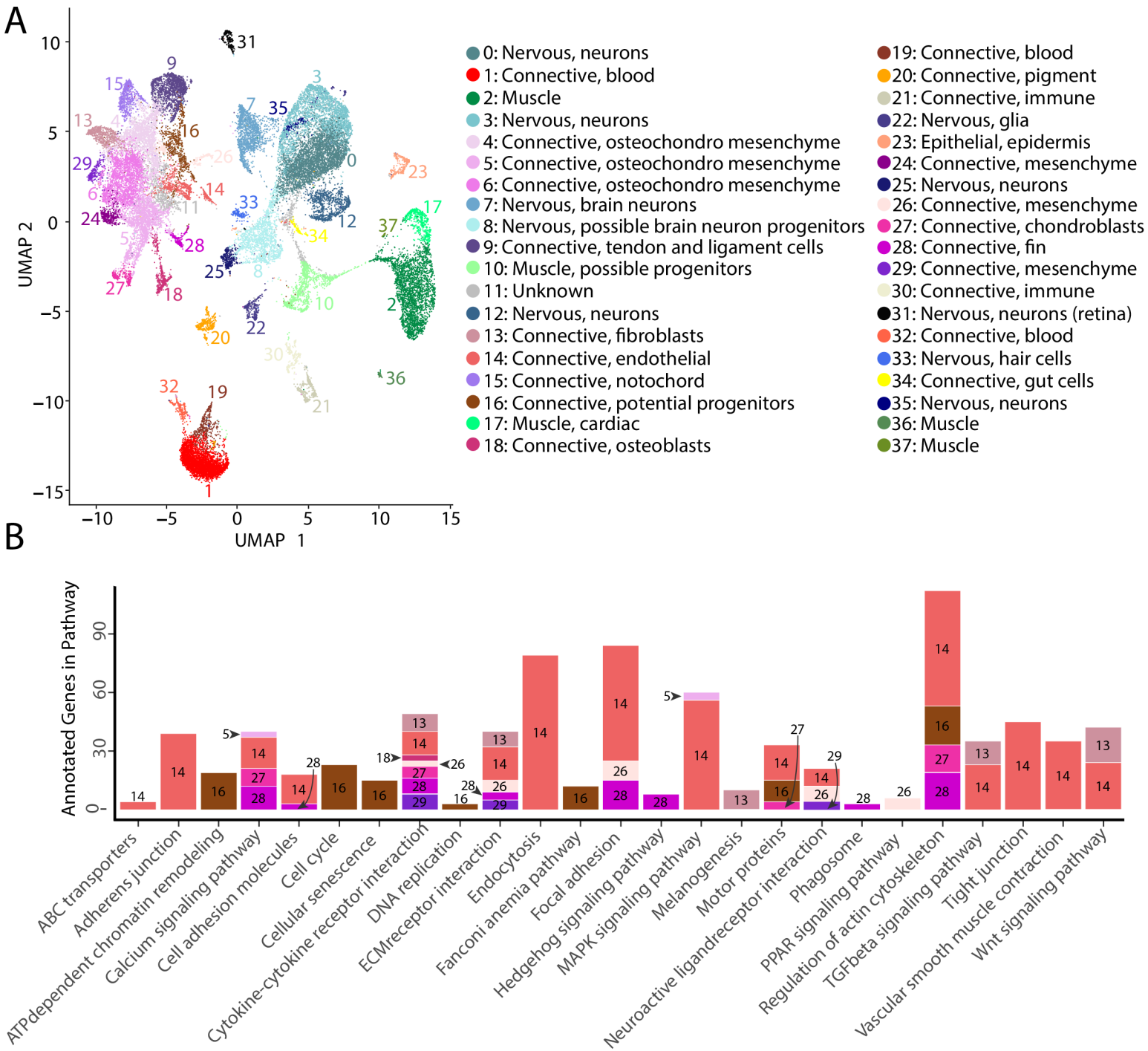
Young, M. D., & Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *GigaScience*, *9*(12), 1–10.

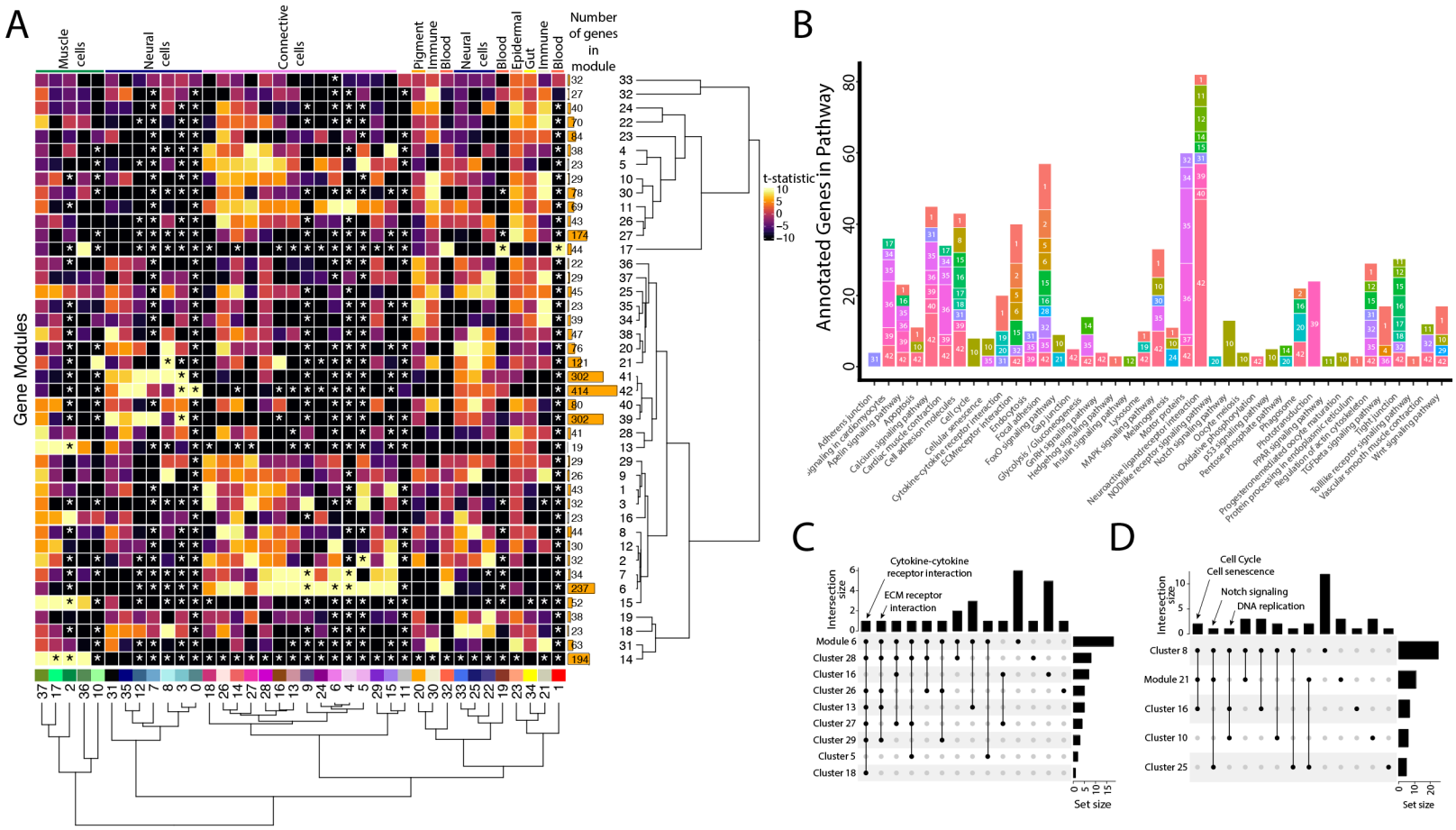
Zhang, Y., Ravi, V., Qin, G., Dai, H., Zhang, H., Han, F., Wang, X., & Liu, Y. (2020a). Comparative genomics reveal shared genomic changes in syngnathid fishes and signatures of genetic convergence with placental mammals ! Phylogenomics and evolutionary rate Transposable elements in syngnathids Expansion and contraction of gene families in t. *National Science Review*, *7*(6), 964–977.

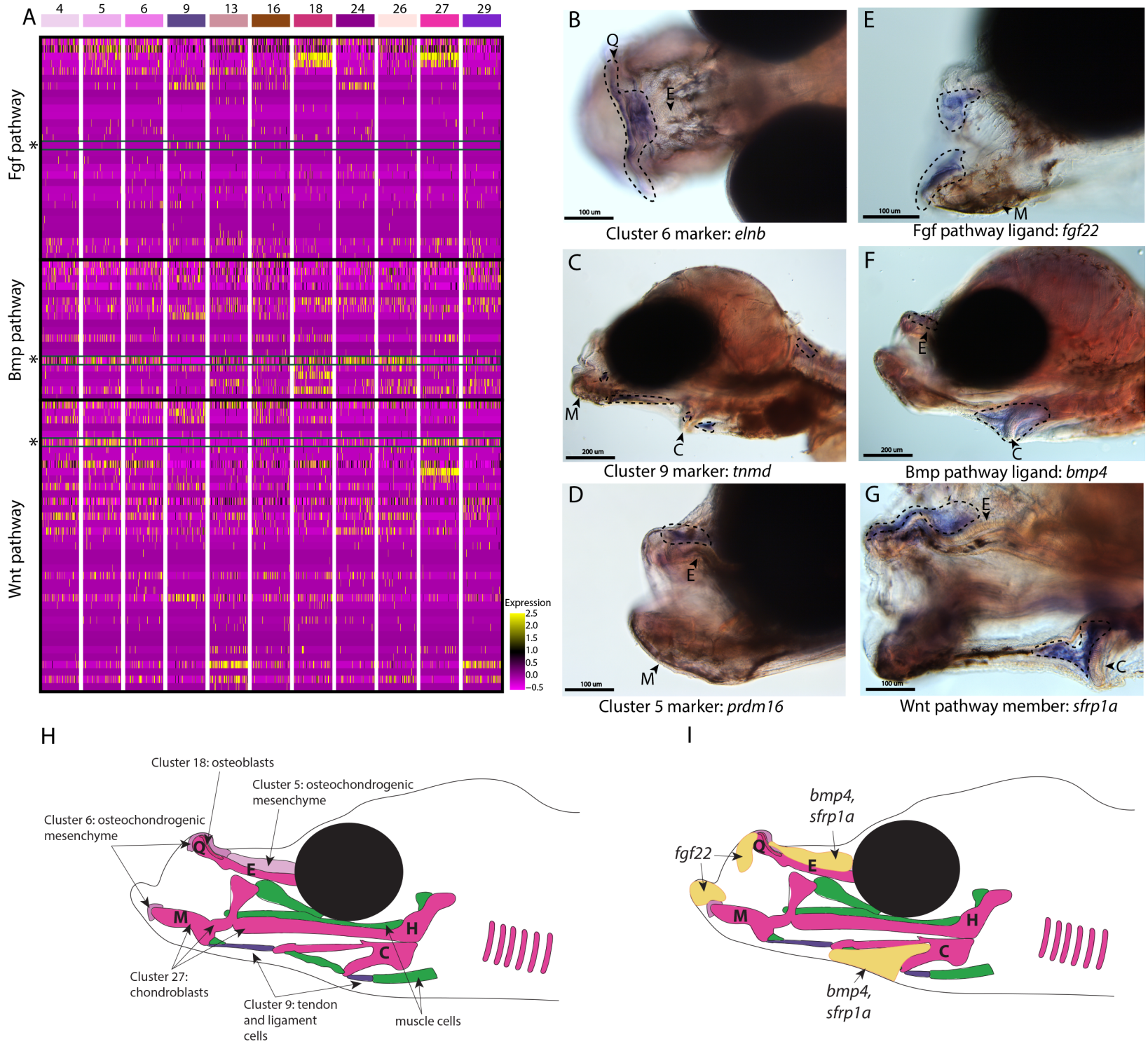
Zhang, Y., Ravi, V., Qin, G., Dai, H., Zhang, H., Han, F., Wang, X., & Liu, Y. (2020b). *Comparative genomics reveal shared genomic changes in syngnathid fishes and signatures of genetic convergence with placental mammals ! Phylogenomics and evolutionary rate Transposable elements in syngnathids Expansion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

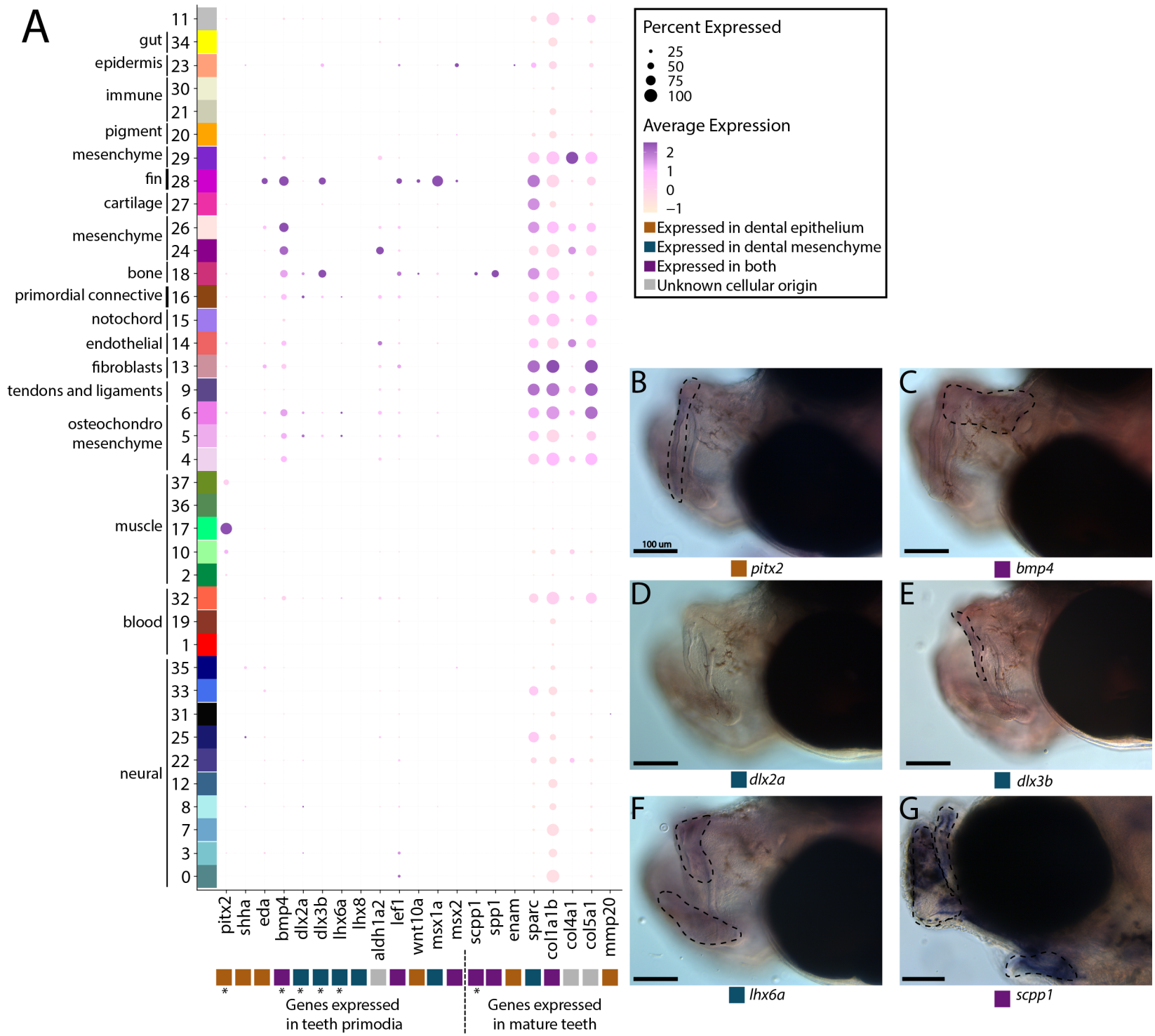
Zhou, J., Gao, Y., Lan, Y., Jia, S., & Jiang, R. (2013). Pax9 regulates a molecular network involving Bmp4, Fgf10, Shh signaling and the Osr2 transcription factor to control palate morphogenesis. *Development*, *140*(23), 4709–4718.

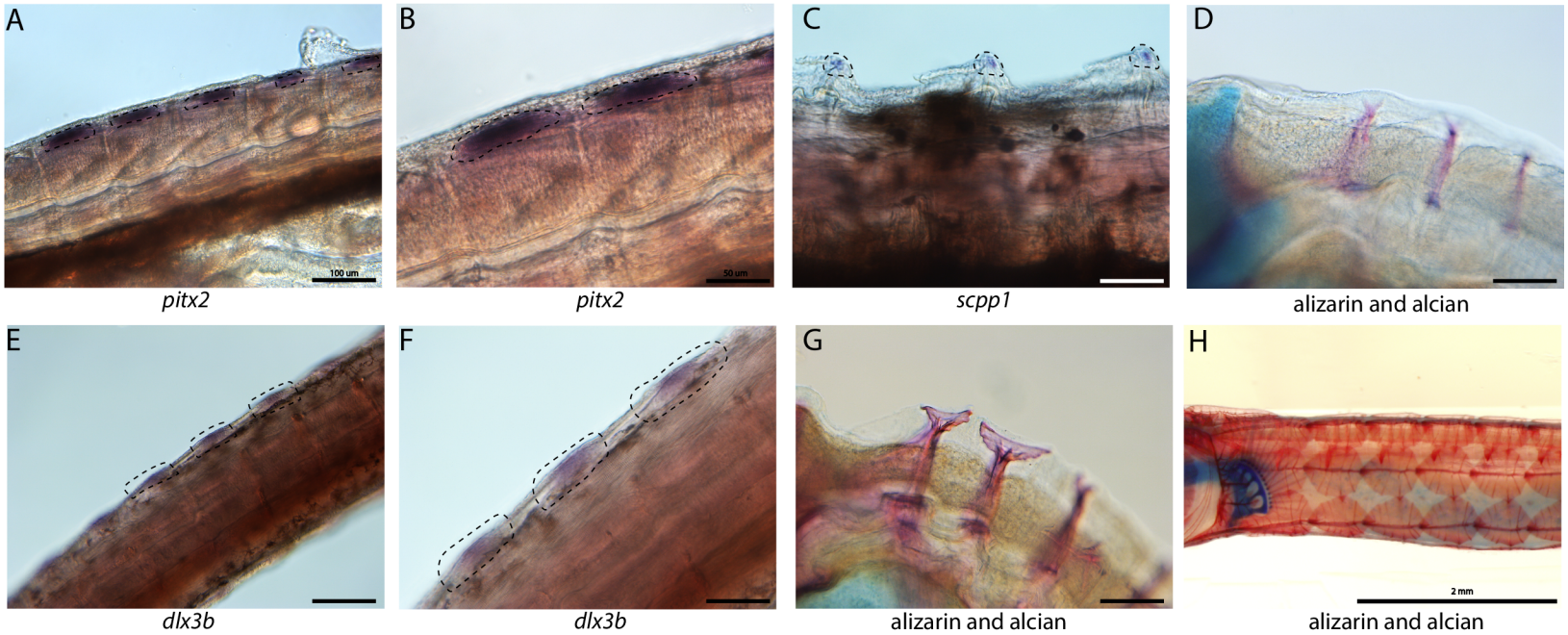

A
No teeth
Elongated ethmoid region
Dermal armor
Male brooding
B
Q
E
M
P
H
B
C
C
E
M
P
H
C
D
F
N
E
PD
N
F
E
M
H
C
G
F
N
H
PD
N

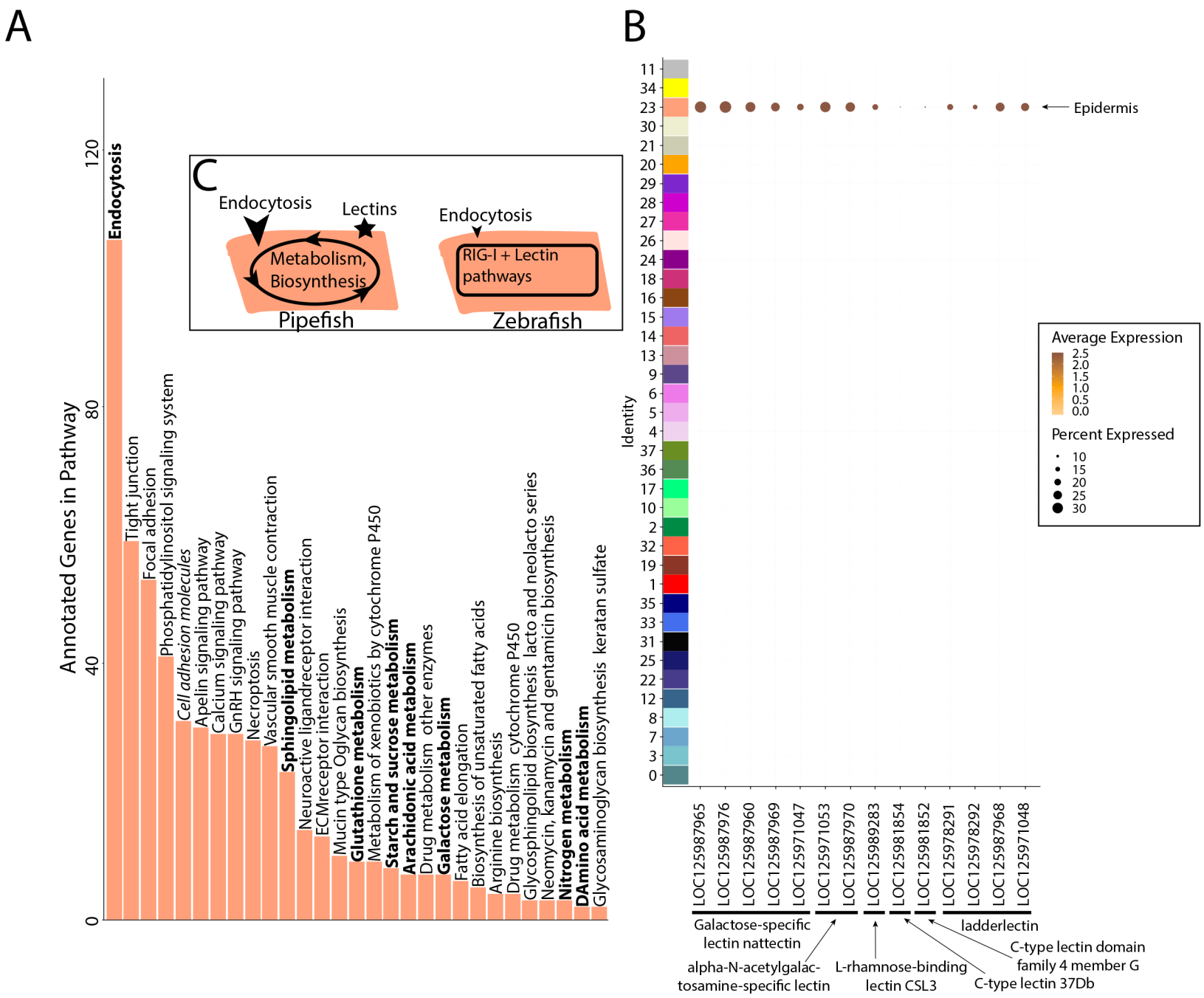

A
UMAP 2
UMAP 1
0: Nervous, neurons
1: Connective, blood
2: Muscle
3: Nervous, neurons
4: Connective, osteochondro mesenchyme
5: Connective, osteochondro mesenchyme
6: Connective, osteochondro mesenchyme
7: Nervous, brain neurons
8: Nervous, possible brain neuron progenitors
9: Connective, tendon and ligament cells
10: Muscle, possible progenitors
11: Unknown
12: Nervous, neurons
13: Connective, fibroblasts
14: Connective, endothelial
15: Connective, notochord
16: Connective, potential progenitors
17: Muscle, cardiac
18: Connective, osteoblasts
19: Connective, blood
20: Connective, pigment
21: Connective, immune
22: Nervous, glia
23: Epithelial, epidermis
24: Connective, mesenchyme
25: Nervous, neurons
26: Connective, mesenchyme
27: Connective, chondroblasts
28: Connective, fin
29: Connective, mesenchyme
30: Connective, immune
31: Nervous, neurons (retina)
32: Connective, blood
33: Nervous, hair cells
34: Connective, gut cells
35: Nervous, neurons
36: Muscle
37: Muscle
B
Annotated Genes in Pathway
ABC transporters
Adherens junction
ATPdependent chromatin remodeling
Calcium signaling pathway
Cell adhesion molecules
Cell cycle
Cellular senescence
Cytokine-cytokine receptor interaction
DNA replication
ECMreceptor interaction
Endocytosis
Fanconi anemia pathway
Focal adhesion
Hedgehog signaling pathway
MAPK signaling pathway
Melanogenesis
Motor proteins
Neuroactive ligandreceptor interaction
Phagosome
PPAR signaling pathway
Regulation of actin cytoskeleton
TGFbeta signaling pathway
Tight junction
Vascular smooth muscle contraction
Wnt signaling pathway


A
Muscle cells
Neural cells
Connective cells
Pigment
Immune
Blood
Neural cells
Blood
Epidermal
Gut
Immune
Blood
Number of genes in module
Gene Modules
t-statistic
B
Annotated Genes in Pathway
Adherens junction
Apelin signaling pathway
Apoptosis
Calcium signaling pathway
Cell adhesion molecules
Cell cycle
Cellular senescence
Cytokine-cytokine receptor interaction
ECM receptor interaction
Endocytosis
Focal adhesion
FoxO signaling pathway
Glycolysis / Gluconeogenesis
Gap junction
Hedgehog signaling pathway
Insulin signaling pathway
MAPK signaling pathway
Lysosome
Melanogenesis
Motor proteins
Neuroactive ligand-receptor interaction
NOD-like receptor signaling pathway
Notch signaling pathway
Oocyte meiosis
Oxidative phosphorylation
Pentose phosphate pathway
Phagosome
Phototransduction
PPAR signaling pathway
Progesterone-mediated oocyte maturation
Protein processing in endoplasmic reticulum
Regulation of actin cytoskeleton
TGF-beta signaling pathway
Tight junction
Toll-like receptor signaling pathway
Vascular smooth muscle contraction
Wnt signaling pathway
C
Intersection size
Cytokine-cytokine receptor interaction
ECM receptor interaction
Module 6
Cluster 28
Cluster 16
Cluster 26
Cluster 13
Cluster 27
Cluster 29
Cluster 5
Cluster 18
Set size
D
Intersection size
Cell Cycle
Cell senescence
Notch signaling
DNA replication
Cluster 8
Module 21
Cluster 16
Cluster 10
Cluster 25
Set size


A
4
5
6
9
13
16
18
24
26
27
29
Fgf pathway
Bmp pathway
Wnt pathway
Expression
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
B
Q
E
100 um
Cluster 6 marker: elnb
E
M
100 um
Fgf pathway ligand: fgf22
C
M
C
200 um
Cluster 9 marker: tnmd
F
E
C
200 um
Bmp pathway ligand: bmp4
D
E
M
100 um
Cluster 5 marker: prdm16
G
E
C
100 um
Wnt pathway member: sfrp1a
H
Cluster 18: osteoblasts
Cluster 5: osteochondrogenic mesenchyme
Cluster 6: osteochondrogenic mesenchyme
Q
E
M
H
C
Cluster 27: chondroblasts
Cluster 9: tendon and ligament cells
muscle cells
I
bmp4, sfrp1a
fgf22
Q
E
M
H
C
bmp4, sfrp1a


A
gut
epidermis
immune
pigment
mesenchyme
fin
cartilage
mesenchyme
bone
primordial connective
notochord
endothelial
fibroblasts
tendons and ligaments
osteochondro mesenchyme
muscle
blood
neural
11
34
23
30
21
20
29
28
27
26
24
18
16
15
14
13
9
6
5
4
37
36
17
10
2
32
19
1
35
33
31
25
22
12
8
7
3
0
pitx2
shha
eda
bmp4
dlx2a
dlx3b
lhx6a
lhx8
aldh1a2
lef1
wnt10a
msx1a
msx2
scpp1
spp1
enam
sparc
col1a1b
col4a1
col5a1
mmp20
Genes expressed in teeth primodia
Genes expressed in mature teeth
Percent Expressed
25
50
75
100
Average Expression
2
1
0
−1
Expressed in dental epithelium
Expressed in dental mesenchyme
Expressed in both
Unknown cellular origin
B
100 um
pitx2
C
bmp4
D
dlx2a
E
dlx3b
F
lhx6a
G
scpp1


A
100 um
pitx2
B
50 um
pitx2
C
scpp1
D
alizarin and alcian
E
dlx3b
F
dlx3b
G
alizarin and alcian
H
2 mm
alizarin and alcian


A
Annotated Genes in Pathway
0
40
80
120
Endocytosis
Tight junction
Focal adhesion
Phosphatidylinositol signaling system
Cell adhesion molecules
Apelin signaling pathway
Calcium signaling pathway
GnRH signaling pathway
Necroptosis
Vascular smooth muscle contraction
Sphingolipid metabolism
Neuroactive ligandreceptor interaction
ECMreceptor interaction
Mucin type Oglycan biosynthesis
Glutathione metabolism
Metabolism of xenobiotics by cytochrome P450
Starch and sucrose metabolism
Arachidonic acid metabolism
Drug metabolism other enzymes
Galactose metabolism
Fatty acid elongation
Biosynthesis of unsaturated fatty acids
Arginine biosynthesis
Drug metabolism cytochrome P450
Glycosphingolipid biosynthesis lacto and neolacto series
Neomycin, kanamycin and gentamicin biosynthesis
Nitrogen metabolism
DAmino acid metabolism
Glycosaminoglycan biosynthesis keratan sulfate
C
Endocytosis
Lectins
Metabolism, Biosynthesis
Pipefish
Endocytosis
RIG-I + Lectin pathways
Zebrafish
B
Identity
11
34
23
30
21
20
29
28
27
26
24
18
16
15
14
13
9
6
5
4
37
36
17
10
2
32
19
1
35
33
31
25
22
12
8
7
3
0
Epidermis
Average Expression
2.5
2.0
1.5
1.0
0.5
0.0
Percent Expressed
10
15
20
25
30
LOC125987965
LOC125987976
LOC125987960
LOC125987969
LOC125971047
LOC125971053
LOC125987970
LOC125989283
LOC125981854
LOC125981852
LOC125978291
LOC125978292
LOC125987968
LOC125971048
Galactose-specific lectin nattectin
alpha-N-acetylgalac-tosamine-specific lectin
L-rhamnose-binding lectin CSL3
ladderlectin
C-type lectin 37Db
C-type lectin domain family 4 member G