# A Molecular, Spatial, and Regulatory Atlas of the *Hydra vulgaris* Nervous System

Hannah Morris Little[a*], Abby S. Primack[a*], Jennifer Tsverov[a*], Susanne Mühlbauer[b], Ben D. Cox[a], Christina Busse[b], Sandra Schneid[b], Amber Louwagie[b], Jack F. Cazet[a], Charles N. David[b], Jeffrey A. Farrell[b], Celina E. Juliano[a**]

[a] Department of Molecular and Cellular Biology, University of California, Davis, CA 95616

[b] Department of Biology, Ludwig-Maximilians-University Munich, 82152 Martinsried, Germany

[c] Division of Developmental Biology, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Bethesda, MD 20814, USA

*These authors made equal contributions

**Corresponding Author

Email: cejuliano@ucdavis.edu

## Summary Statement

*Hydra vulgaris* is a model for regenerative and functional neuroscience. This study identifies fifteen neuron subtypes using scRNA-seq, maps spatial distributions, explores regulatory mechanisms, and provides an accessible web portal.

## Abstract

*Hydra vulgaris*, a cnidarian with a simple nerve net, is an emerging model for developmental, regenerative, and functional neuroscience. Its genetic tractability and capacity for whole-system imaging make it well suited for studying neuron replacement, regeneration, and neural circuit function. Here, we present the most comprehensive molecular and spatial characterization of the *H. vulgaris* nervous system to date. Using single-cell RNA sequencing, we identified eight neuron types, each defined by distinct neuropeptide expression, and further resolved these into fifteen transcriptionally distinct subtypes with unique spatial distributions and morphologies. To investigate the gene regulatory networks underlying neuronal differentiation, we applied trajectory inference, identified key transcription factors, and performed ATAC-seq on sorted neurons to map chromatin accessibility. All datasets are available through an interactive, user-friendly web portal to support broad use by the research community. Together, these resources provide a foundation for uncovering molecular mechanisms that govern nervous system development, homeostasis, and regeneration in *H. vulgaris*.

## Introduction

The small freshwater cnidarian *Hydra vulgaris* (**Fig. 1A**) is a classic model for studying regeneration due to its ability to regenerate its entire body, including the nervous system (Bode et al., 1988; Rentzsch et al., 2019; Trembley et al., 1744; Vogg et al., 2019b). Even in uninjured animals, neurogenesis pathways remain continuously active, enabling constant neuron replacement and providing easy experimental access to study the molecular mechanisms of neurogenesis (David and Gierer, 1974; David and Murphy, 1977; Siebert et al., 2019). Recently, *H. vulgaris* has also gained recognition as a valuable model in functional neuroscience, with the identification of neural circuits that drive specific behaviors (Dupre and Yuste, 2017; Giez et al., 2023; Yamamoto and Yuste, 2023). Building on this foundational research, *H. vulgaris* now serves as a robust model for regenerative neuroscience, offering tools to investigate the complete regeneration of an adult nervous system, from molecules to behavior.

The *H. vulgaris* body is formed by two epithelial monolayers, an outer ectodermal layer and an inner endodermal layer (**Fig. 1A**). Its nervous system consists of approximately 3,000–5,000 neurons (around 3-5% of total body cells), organized into two distinct nerve nets, one within

each epithelial layer (Bode et al., 1973; Keramidioti et al., 2024). *H. vulgaris* neurons are part of the interstitial cell lineage, and their replacement during both homeostasis and regeneration is supported by multipotent adult interstitial stem cells (ISCs) located among the ectodermal epithelial cells (**Fig. 1A**) (David and Gierer, 1974; David and Murphy, 1977). Due to passive tissue displacement toward the extremities, which leads to ongoing cell loss, ISCs continuously replenish neurons even in the absence of injury (Campbell, 1967; David and Gierer, 1974; Hager and David, 1997). The constant activity of neurogenesis pathways in uninjured adult polyps allows for the detailed profiling of differentiation states using single-cell RNA sequencing (scRNA-seq) (Siebert et al., 2019).

In addition to exhibiting continual, widespread neuronal regeneration, *H. vulgaris* are amenable to tests of gene function, such as gene knock-down through the expression of RNA hairpins or the electroporation of siRNAs (Juliano et al., 2014a; Lommel et al., 2017; Vogg et al., 2019a). Additionally, creating stable transgenic *H. vulgaris* lines is straightforward using cell-type-specific promoters (Siebert et al., 2008; Siebert et al., 2019; Wittlieb et al., 2006). This has enabled the expression of fluorescent calcium reporters like GCaMP to monitor neuronal activity and the use of nitroreductase to selectively ablate specific neurons (Dupre and Yuste, 2017; Giez et al., 2023; Noro et al., 2021). Together, these attributes make *H. vulgaris* an accessible model for investigating the molecular mechanisms underlying nervous system development and regeneration, from cellular differentiation to the formation of neural circuits driving behavior. Additionally, as a member of Cnidaria, the sister group to bilaterians, *H. vulgaris* provides a unique opportunity to explore the evolution of the molecular processes involved in nervous system development, regeneration, and function.

Decades of research have established a foundational framework for studying the *H. vulgaris* nervous system. Our basic understanding includes the source of new neurons, the relative distribution of neurons along the oral-aboral axis, and neurogenesis rates in an uninjured *H. vulgaris* (Bode et al., 1973; Bode et al., 1988; David, 2012; David and Gierer, 1974). In our previous work, where we created a whole-animal single-cell expression map, we transcriptionally profiled *H. vulgaris* neurons and largely analyzed in situ hybridization patterns of marker genes from existing literature to estimate their spatial distribution (Siebert et al.,

82 2019). However, due to the relatively small number of neurons sequenced (~3,500) in that study,
83 there was a possibility of missing some neuron populations. Additionally, we did not capture
84 enough intermediate states of neuronal differentiation to fully resolve the transcriptional changes
85 occurring as neurons develop from ISCs. Another study used single-cell SMART-seq2 to
86 sequence 1,152 *H. vulgaris* interstitial cells, including 384 neurons (Klimovich et al., 2020).
87 Although this method provided deeper sequencing, it did not substantially increase the number of
88 neurons available for analysis.
89
90 We addressed previous limitations by sequencing 10-fold as many neurons, generating ~35,000
91 single-cell transcriptomes of neurons and neural progenitors. This enhanced dataset allowed us to
92 confirm the twelve neuron subtypes identified in our original study (Siebert et al., 2019) and to
93 identify three previously unrecognized subtypes. Using fluorescent in situ hybridization (FISH),
94 in situ hybridization chain reaction (HCR), and immunostaining, we refined the spatial map of
95 the *H. vulgaris* nervous system, revealing how neuron subtypes are organized into established
96 circuits characterized by shared neuropeptide expression. To further dissect the regulatory
97 underpinnings of neural differentiation, we performed URD trajectory inference and ATAC-seq
98 (Assay for Transposase-Accessible Chromatin using Sequencing) on sorted neurons to
99 characterize the regulatory landscape during nervous system differentiation (Buenrostro et al.,
100 2015; Farrell et al., 2018) and we demonstrate how these data can be used to formulate
101 hypothetical gene regulatory networks (GRNs) for future testing. To support broad use and
102 accelerate discovery in *H. vulgaris* neurobiology and regeneration, all data and tools are freely
103 available through an interactive, user-friendly web portal.
104

105 ## Results

106 **Expansion of the neuronal single cell transcriptome dataset for *Hydra vulgaris***

107 In our previous study, we used Drop-seq to create a single cell atlas of the adult *H. vulgaris*
108 polyp, which included approximately 3,500 single-cell transcriptomes of differentiated neurons
109 and cells undergoing neurogenesis (Macosko et al., 2015; Siebert et al., 2019). To build on this
110 initial study, we aimed to increase the number of neural single-cell transcriptomes to uncover
111 any additional molecular diversity in the *H. vulgaris* neuron repertoire that may have been
112 missed due to the relatively small number of cells sequenced. To accomplish this, we performed

113 scRNA-seq using the 10x Genomics platform to expand our dataset. To enrich for neurons and
114 neural progenitors, we used Fluorescent Activated Cell Sorting (FACS) to collect cells from two
115 different transgenic lines (**Fig. S1, Table S1**): (1) *Tg(actin1:GFP)*$^{rs3\text{-}in}$, a previously published
116 line in which GFP is expressed in all differentiated neurons, neural progenitors, and ISCs
117 (Keramidioti et al., 2024), and (2) *Tg(tba1c:mNeonGreen)*$^{cj1\text{-}gt}$, a line created for this study in
118 which mNeonGreen is expressed under the control of the pan-neuronal gene *tba1c* and is
119 therefore predicted to be expressed in all differentiated neurons (see Material and Methods for
120 explanation of transgenic line naming scheme). However, we observed that certain neuron
121 subtypes were underrepresented in the scRNA-seq data collected from this line, suggesting that
122 our cloned promoter conferred some neuronal subtype specificity, possibly due to incomplete
123 coverage of the *tba1c* regulatory region in our transgenic construct. (**Figs. S1, S2**).
124
125 We combined our new data with previously collected neuronal single-cell transcriptomes from
126 our earlier study for downstream analysis (**Table S2**) (Siebert et al., 2019). Sequencing reads
127 from all libraries were mapped to the *H. vulgaris* strain AEP gene models and processed
128 following standard protocols (**Fig. S3**) (Cazet et al., 2023; Siebert et al., 2019). The expression of
129 the pan-neuronal marker *elav2* was used to determine neuron identity and therefore remove non-
130 neurons from the data set (Siebert et al., 2019). After filtering, we recovered 35,071 single-cell
131 neural transcriptomes, expanding our dataset tenfold. We detected a median of 1371.5 genes and
132 2887.5 UMIs per cell (**Table S2**), a substantial improvement compared to the median of 563.5
133 genes and 1,082 UMIs per cell detected in the neuron-enriched Drop-seq libraries from our
134 previous study (Siebert et al., 2019). To identify distinct neuron subtypes in our dataset, we used
135 Seurat to perform Louvain clustering and visualized the results using Uniform Manifold
136 Approximation and Projection (UMAP) (**Fig. 1B, Fig. S4, Table S3**) (Hao et al., 2021; McInnes
137 et al., 2018; Satija et al., 2015; Stuart et al., 2019). Fifteen clusters of putative differentiated
138 neurons were identified based on their high expression of *tba1c* and the absence of stem and
139 progenitor gene expression (**Fig. S5**). As detailed below, all fifteen of these clusters were
140 confirmed to consist of differentiated neurons. We performed non-negative matrix factorization
141 (NMF) (Kotliar et al., 2019) to identify groups of co-expressed genes ("metagenes"), and we
142 recovered metagenes for the fifteen neuron clusters, and stem and progenitor cells (**Fig. S6**).
143

144 **An integrated cell-type atlas of the *Hydra vulgaris* nervous system**
145 This expanded scRNA-seq dataset provides a valuable resource for studying nervous system
146 development, regeneration, and function in *H. vulgaris*. In this section, our goal is to present a
147 comprehensive description of the *H. vulgaris* nervous system by integrating our new data with
148 previously published work, including recent work by Keramidioti et al. (2024) using a novel pan-
149 neural antibody (PNab) to provide a detailed anatomical description of the *H. vulgaris* nerve net.
150 We begin by revisiting and expanding the neuron naming scheme originally proposed in Siebert
151 et al., 2019, which defined eight neuron types further subdivided into subtypes. For each
152 type/subtype, we describe roles in known neural circuits where applicable, determine
153 neuropeptide expression patterns, and use FISH or HCR to characterize the spatial localization of
154 subtypes, and in some cases identify novel subtypes (**Fig. 2, Fig. 3**). These integrated data
155 provide a detailed classification of neuron subtypes within the *H. vulgaris* nervous system.
156
157 ***Classification scheme for H. vulgaris neurons*:** To systematically classify *H. vulgaris* neurons,
158 we developed a hierarchical and expandable naming scheme that can accommodate future
159 insights into neuronal function, morphology, and spatial organization. The *H. vulgaris* nervous
160 system is comprised of two nerve nets, one embedded in the ectoderm and one in the endoderm
161 (Keramidioti et al., 2024). The ectodermal nerve net extends from the basal disk to the tentacle
162 tips, with higher neuron density at the oral (tentacle and hypostome) and aboral (peduncle and
163 basal disk) regions as compared to the body column (Keramidioti et al., 2024). Along the length
164 of the animal, ectodermal neurons project neurites that form bundles and communicate laterally
165 through gap junctions (Keramidioti et al., 2024). Two major non-overlapping circuits,
166 contraction burst (CB) and rhythmic potential 1 (RP1), run along the axis in the ectoderm and
167 regulate behaviors such as longitudinal contractions and somersaulting (Dupre and Yuste, 2017;
168 Yamamoto and Yuste, 2023). The endodermal nerve net is distributed along the body in the
169 endodermal layer but is absent in the tentacles (Keramidioti et al., 2024). The rhythmic potential
170 2 (RP2) neural circuit, which is involved in feeding behavior, is part of the endodermal nerve net
171 (Dupre and Yuste, 2017; Giez et al., 2023; Keramidioti et al., 2024; Siebert et al., 2019).
172
173 Building on our previous work, we first assigned neurons to either the ectodermal ("ec") or
174 endodermal ("en") nerve nets, using expression of the endodermal marker gene *budhead* to guide

175 this distinction (**Fig. 1D, Fig. S5**) (Siebert et al., 2019). Our prior classification grouped
176 ectodermal neurons into five types (ec1-ec5); here, we confirm and extend this framework by
177 showing that each type has a unique neuropeptide expression profile (**Fig. 1C, Table S4**). This
178 neuropeptide-based strategy is consistent with approaches used to classify neuron types in *C.*
179 *elegans*, mice, and humans (Das Gupta et al., 2021; Taylor et al., 2021; Zhong et al., 2022). As
180 described in detail below, we further subdivided the five ectodermal types into twelve
181 transcriptionally and spatially distinct subtypes (**Fig. 1C, Fig. 2, Fig. 3A,B**). In contrast, the
182 endodermal nerve net is less complex, consisting of three transcriptionally distinct clusters that
183 show moderate variation in neuropeptide expression (**Fig. 1C**). As described in detail below,
184 these clusters correspond to three neuron types (en1-en3) (**Fig. 1C**). Overall, this classification
185 extends the original ec1-ec5 and en1-en3 scheme described in Siebert et al. (2019) and provides
186 a robust framework for future refinement as additional data emerge.
187
188 ***ec3 neurons:*** The RP1 circuit is composed of ec3 neurons, which express a suite of
189 neuropeptides, most prominently *Hym-355* and products of the *GLWamide* gene (**Fig. 1C**)
190 (Keramidioti et al., 2024; Leviev et al., 1997; Yamamoto and Yuste, 2023). To visualize the
191 architecture of this circuit, we performed co-immunostaining using an antibody recognizing
192 GLWamide peptides (Koizumi et al., 2015) and PNab (stains all neurons), revealing ec3 neurons
193 extending throughout the length of the body column (**Fig. 3A-D**). ec3 neurons also uniquely
194 express *innexin 6* and *innexin 14*, which likely contribute to gap junction formation and electrical
195 coupling within the RP1 circuit (**Fig. S9B,C**) (Keramidioti et al., 2024). Consistent with our
196 previously published data set, the ec3 neuron population forms a transcriptional continuum that
197 can be divided into three clusters, each enriched for different marker genes **(Fig. 1B, Fig. S5**)
198 (Siebert et al., 2019). We used fluorescent in situ hybridization (FISH) to map spatial
199 distributions of these neuronal clusters and observed distinct localization along the oral-aboral
200 axis, supporting classification as three ec3 subtypes. ec3A neurons are found in the basal disk,
201 ec3B neurons are found in the body column and hypostome, and ec3C neurons are found in the
202 tentacles (**Fig. 1B,D**, **Fig. 2G, I, K, Fig. 3A,B**).
203
204 ***ec1 and ec5 neurons:*** Together, ec1 and ec5 neurons comprise the CB circuit, which drives
205 longitudinal contractions (Dupre and Yuste, 2017; Escobar et al., 2024; Noro et al., 2019; Noro

206 et al., 2021). Both ec1 and ec5 neurons express *innexin-2*, a gap junction gene that facilitates
207 communication between cells within the circuit (**Fig. S9A**) (Keramidioti et al., 2024; Siebert et
208 al., 2019; Takaku et al., 2014). ec1 neurons express *FRamide* and members of the *Hym176*
209 neuropeptide family, including *Hym176B* **(Fig. 1C**). We visualized ec1 neurons using a
210 transgenic reporter line expressing GFP under the *Hym176B* promoter (Noro et al., 2019),
211 combined with GFP and PNab co-immunostaining. This revealed the distribution of the ec1
212 neurons in the CB circuit, extending from just above the peduncle to the hypostome and tentacle
213 base (**Fig. 3A,B,G,H)**. Co-immunostaining the *Hym176B* promoter line with PNab and
214 GLWamide, a marker for ec3 neurons, confirmed that the body column ectodermal nerve net
215 contains only ec1A and ec3B neurons (**Fig. 3A,B, Fig. S8**). ec5 neurons complete the CB circuit
216 and share a similar neuropeptide expression profile with ec1, with additional expression of
217 *Hym176C* and *RFamideA* (**Fig. 1C**). We confirmed their location in the peduncle using FISH for
218 *Hym176C* (**Fig. 2J**) and RFamide immunostaining (**Fig. 3A,B,I,J**), noting that RFamide is also
219 expressed in ec4 neurons (Koizumi et al., 2015).
220
221 In our previous study, we identified two distinct ec1 subtypes, ec1A and ec1B, based on distinct
222 molecular markers (Siebert et al., 2019) (**Fig. S5**). Using FISH, we now confirm that ec1A
223 neurons localize to the body column, whereas ec1B neurons are restricted to the tentacles,
224 supporting their classification as separate subtypes (**Fig. 2E,H**). In our current dataset, we
225 identified three additional clusters expressing *FRamide* and *Hym176* family neuropeptides,
226 consistent with ec1 identity (**Fig. 1B, C**). FISH or HCR with cluster-specific marker genes
227 confirmed three new neuron subtypes: ec1C, ec1D, and ec1E (**Fig. 2A,B,L, Fig. S5**). ec1C
228 neurons form a ring around the mouth, potentially resembling the nerve ring previously
229 described in *Hydra oligactis* (**Fig. 2A, Fig. S7C**) (Grimmelikhuijzen, 1985). ec1E neurons are
230 located just above ec1C in the lower hypostome, while ec1D neurons localize to the basal disk
231 (**Fig. 2B,L**).
232
233 Previous studies reported only a single neuronal population in the basal disk (Dupre and Yuste,
234 2017; Siebert et al., 2019). Therefore, to distinguish ec1D from ec3A neurons, we performed
235 double FISH using subtype-specific markers, confirming their separation (**Fig. 4A-C**). Co-
236 staining with GLWamide and PNab further supported the presence of a second, non-ec3A

237 neuronal population in the basal disk, as we observed PNab-positive cells lacking GLWamide
238 expression (**Fig. 4D-F**). In summary, the CB circuit is composed of six neuronal subtypes, ec5
239 and ec1A-E, which span the entire body axis from the basal disk (ec1D) to the tentacles (ec1B).
240
241 For both the RP1 and CB circuits, the potential functional roles of the distinct neuronal subtypes
242 remain unclear. However, differential gene expression within our dataset may provide insights
243 into their specialization. Given the spatial restriction of these subtypes, observed gene expression
244 differences may also reflect positional cues mediated by the Wnt signaling gradient originating
245 from the oral end.
246
247 ***ec2 and ec4 neurons:*** In our previous work, we found that ec2 and ec4 neurons are localized to
248 the oral end and likely do not participate in the RP1 or CB circuits (Keramidioti et al., 2024;
249 Siebert et al., 2019). These neurons express distinct, non-overlapping sets of neuropeptide genes
250 (**Fig. 1C)**. The ec4 cluster can be further subdivided into two transcriptionally distinct subtypes:
251 ec4A and ec4B (**Fig. 1B, Fig. S5**). FISH analysis revealed that ec4A neurons are in the tentacles,
252 while ec4B neurons are in the hypostome (**Fig. 2C,D, Fig. S8A,B**). Thus, four ectodermal
253 neuron subtypes are in the hypostome (ec1C, ec1E, ec3B, and ec4B) and four are in the tentacles
254 (ec1B, ec2, ec3C, and ec4A) (**Fig. 2**). The increased neuronal complexity at the oral end likely
255 reflects the demands of coordinating tentacle and hypostome activity during behaviors such as
256 feeding and locomotion.
257
258 We visualized the architecture of the hypostomal nerve net using immunostaining for RFamide
259 (ec4), LWamide (ec3), and GFP in the *Hym176B* reporter line (ec1) (**Fig. 4G-L**). The tentacular
260 nerve net remains less well resolved; however, ec2 neurons are likely sensory neurons embedded
261 in the battery cell complex, an ectodermal epithelial cell that contains embedded nematocytes
262 and the sensory neuron responsible for triggering nematocyte discharge (Hobmayer et al., 1990).
263 In our whole-animal scRNA-seq dataset, ectodermal battery cells were sequenced along with
264 their associated neurons, and we detected strong expression of the ec2 marker *RFamideC*, further
265 supporting this interpretation (Siebert et al., 2019).
266

267 ***en1, en2, and en3 neurons:*** Compared to ectodermal neurons, endodermal neurons express
268 fewer neuropeptide genes. en1 neurons express *GLWamide*, while en2 and en3 neurons have no
269 strong neuropeptide expression (**Fig. 1C**). Immunostaining with the GLWamide antibody shows
270 that en1 neurons form a ganglion-like network which constitutes the RP2 circuit (**Fig. 3A,E,F**).
271 en1 neurons can communicate through gap junctions formed by *innexin 2* (**Fig. 9A**) (Dupre and
272 Yuste, 2017; Keramidioti et al., 2024; Siebert et al., 2019). In our previous study, we identified
273 en2 neurons as sensory cells using a transgenic reporter line; these neurons are characterized by a
274 short apical sensory cilium projecting into the gastric cavity and a longer basal neurite that
275 connects with the en1 neuron network (Keramidioti et al., 2024; Siebert et al., 2019). en2 cells
276 express *innexin 15* (**Fig. S9 D**) and could form gap junctions but do not appear to form a neural
277 circuit. To investigate the identity of en3 neurons, we generated a transgenic line expressing
278 mNeonGreen under the control of the en3-specific gene *G022927* regulatory region. Confocal
279 imaging of this reporter line revealed that en3 neurons are an additional type of sensory neuron
280 (**Fig. S10**). Notably, a scanning electron microscopy study from four decades ago described two
281 types of endodermal sensory neurons with unique morphologies, likely corresponding to the en2
282 and en3 neuron types in our data set (Westfall and Epp, 1985). Taken together with their
283 different transcriptional profiles, these morphological differences support the classification of
284 en2 and en3 as separate neuron types.

285 **Developmental trajectory inference of the homeostatic *H. vulgaris* nervous system**

286 The *H. vulgaris* nervous system undergoes continuous renewal, with complete neuronal turnover
287 approximately every three weeks (Bode et al., 1988). As a result, scRNA-seq of homeostatic
288 adult *H. vulgaris* captures a dynamic mixture of stem cells, differentiated neurons, and cells in
289 various stages of differentiation (Siebert et al., 2019). Given the increased number of neuronal
290 progenitors captured in our new data set which provide information about the developmental
291 relationships between neuronal cell types, we reconstructed developmental trajectories under
292 steady-state conditions to describe the molecular events as these diverse neuron types are
293 specified and differentiate. To achieve this, we applied URD, a diffusion-based algorithm that
294 infers developmental trajectories as branching trees (**Fig. 5A, Fig. S11**) (Farrell et al., 2018),
295 using ISCs as the root, defined by the marker *Hy-icell1* (G002332) (Siebert et al., 2019), and
296 terminal tips defined by neuron subtype-specific markers and absence of progenitor markers

297 (**Fig. S5**). Our analysis predicted the relationship between ISCs and eleven terminal branches of
298 neurons. Limitations of the analysis are that ec4A and ec4B were collapsed into a single group,
299 as trajectory connections to ISCs could not be resolved separately, and the newly described
300 ec1C-E subtypes were excluded because progenitors that link them to ISCs were seemingly not
301 captured.

302 We previously showed that ISCs transitioning to neuronal fates express the transcription factors
303 *myc3* and *myb*, which are downregulated upon terminal differentiation (Siebert et al., 2019).
304 Here, our trajectory analysis identified three early-branching populations expressing both *myc3*
305 and *myb* (**Fig. S12**). Two populations are multi-potent progenitors, one of which gives rise to
306 endodermal neurons (en1, en2, en3), and the other of which gives rise to most ectodermal neuron
307 types (ec1–ec4) (**Fig. 5A**). Notably, ec5 neurons appear to differentiate directly from ISCs,
308 though this may reflect incomplete capture of intermediate states (**Fig. 5A**). Together, these
309 findings suggest that endodermal and ectodermal neurons each arise from distinct progenitor
310 pools that become progressively fate-restricted during differentiation, ultimately generating the
311 full diversity of *H. vulgaris* neuron subtypes.

312 The progressive specification and differentiation of neuronal subtypes predicted here will be
313 accompanied by a cascade of transcriptional events that regulate each step. As proof of principle,
314 we used our trajectory analysis to visualize the expression dynamics of transcription factors and
315 terminal marker genes that belong to one such cascade during ec3A differentiation. Plotting the
316 expression of *myb*, *bhlha15*, *gata3, hym355*, and ec3A marker *G021930* shows their sequential
317 expression accompanied by transition states with temporary co-expression (**Fig. 5B**). To validate
318 these predicted transition states, we performed double fluorescent *in situ* hybridization (FISH)
319 using pairs of genes with overlapping expression domains across three predicted transition states
320 (**Fig. 5C-U**). For gene pair, *bhlha15* and *gata3* (**Fig. 5C-G**), *gata3* and *hym355* (**Fig. 5H-L**), and
321 *hym355* and *G021930* **(Fig. 5M-Q**), we identified cells expressing each gene individually as well
322 as cells co-expressing both genes, consistent with our predicted differentiation trajectory.
323 Notably, most neurons co-express the terminal marker genes *hym355* and *G021930* in the basal
324 disk (**Fig. 5M, R-U**), which aligns with the expected localization of terminally differentiated
325 ec3A neurons (**Fig. 2K**). These findings validate our trajectory analysis and demonstrate its
326 ability to accurately identify gene expression dynamics during the differentiation of neuronal

327 subtypes in *H. vulgaris*. This trajectory analysis is available as part of our online portal to enable
328 visualization of similar gene expression cascades during the specification of each *H. vulgaris*
329 neuronal subtype.
330
331 **Characterizing the chromatin landscape in the *H. vulgaris* nervous system**
332
333 Our work establishes *H. vulgaris* as one of the few adult organisms in which the nervous system
334 has been transcriptionally defined, including a trajectory analysis describing the molecular
335 changes that occur during neuronal differentiation. This provides a unique opportunity to identify
336 the GRNs that describe the differentiation of all neurons in an adult nervous system. To advance
337 this goal, we analyzed the expression of transcription factors (TFs) in our data set. We previously
338 used phylogenetic analysis to identify 811 putative TFs, defined as genes with predicted DNA
339 binding domains, with detectable expression in *H. vulgaris* (Cazet et al., 2023). We identified the
340 expression patterns of these TFs across our neuronal cell types, finding that 665 of those 811 TFs
341 are expressed within the nervous system, and we used weighted gene co-expression network
342 analysis (WGCNA) to cluster them based on average expression in each cell type (**Fig. S13**)
343 (Langfelder and Horvath, 2008). Most neuronal subtypes are characterized by the unique
344 expression of at least a dozen TFs, and most broader classes of neurons (e.g. ec1A/B/C/D/E) are
345 also characterized by a few TFs with shared expression across the class.
346
347 To build GRNs underlying nervous system development in *H. vulgaris*, it is essential to identify
348 the cis-regulatory elements (CREs) bound by neuronal transcription factors. As a foundational
349 step towards this goal, we performed ATAC-seq on FACS-isolated neurons from the
350 *Tg(actin1:GFP)*$^{rs3\text{-}in}$ and *Tg(tba1c:mNeonGreen)*$^{cj1\text{-}gt}$ transgenic lines (**Table S5**). These neuron-
351 specific datasets were of high quality and showed strong replicate concordance, meeting
352 ENCODE standards (**Fig. S14, Table S6**) (Landt et al., 2012). Importantly, the data quality
353 matched that of our previously published whole-animal ATAC-seq from *H. vulgaris*, indicating
354 that FACS isolation did not compromise chromatin integrity (Cazet et al., 2023). We observed
355 significantly increased accessibility near neuronal genes in the neuron-specific libraries
356 compared to whole-animal data ($p < 0.001$ for each line; **Fig. 6A, B**). Moreover, this approach
357 uncovered accessible regions at neuronal loci that were not detected in whole-animal, while

358 confirming a lack of accessibility at non-neuronal genes such as *wnt3* (**Fig. 6C-H**) (Cazet et al.,
359 2023). These findings provide an essential resource for defining the regulatory architecture of *H.*
360 *vulgaris* neurogenesis.
361
362 **Gene regulatory network inference demonstrated with ec4 neuron specification**
363
364 To illustrate the utility of the data presented in this study for hypothesis generation, we
365 constructed a putative GRN describing the differentiation of the oral neuron type ec4 (**Fig. 7**).
366 Reasoning that oral neurons might share core regulatory features, we first identified transcription
367 factors expressed during the differentiation of ec1B, ec2, ec3C, and ec4 neurons, and visualized
368 their expression dynamics along pseudotime using spline plots (**Fig. 7A, Fig. S15A**). As
369 observed previously, *Myb* and *Myc3* are activated early during neurogenesis across all neuronal
370 subtypes, marking the onset of neuronal specification (Siebert et al., 2019). Here we show that
371 *Sox2* is expressed broadly across oral neuron subtypes (**Fig. 7B, Fig. S15B**) and that open
372 chromatin near the *Sox2* locus has both *Myb* and *Tcf* binding sites, consistent with the possibility
373 that Sox2 expression within oral neurons is directly regulated by *Myb* and receives inputs from
374 the Wnt signals emanating from the oral organizer (**Fig. 7H, Fig. S15C**).
375
376 As differentiation progresses, the oral neurons separate into two groups: ec1B/ec2/ec4, which
377 express *Zic4*, and ec3C, which express *Gata3* (**Fig. 7C,E, Fig. S15D,F**). Previous studies have
378 described antagonism between *Zic4* and *Gata3* in the ectodermal epithelium, a similar mutual
379 antagonism may regulate this cell fate decision during neurogenesis (Ferenc et al., 2024; Vogg et
380 al., 2022). Following *Zic4* expression, the ec4 neuronal identity becomes increasingly refined,
381 first when *Arx* is jointly expressed in both ec2 and ec4, which is subsequently followed by the
382 differential expression of *Litaf* in ec2 and *Gsx1/2* in ec4 (**Fig. 7D,F, Fig. S15E,G**).
383
384 Although the mechanism driving *Gsx1/2* expression specifically in ec4 neurons is not yet clear,
385 our data suggest a regulatory interaction between *Gsx1/2* and the transcription factor *Noto* that
386 may be key to stabilizing ec4 fate. *Noto* is expressed relatively early in the differentiation of oral
387 neurons, and accessible binding sites for both *Myc3* and *Tcf* suggest possible upstream inputs
388 (**Fig. 7A,G,I**). Interestingly, *Noto* expression declines in ec4 neurons as *Gsx1/2* expression

389 increases (**Fig. 7A,G, Fig. S15H**). Given that the *Noto* locus contains an accessible *Gsx2*-binding
390 site (**Fig. 7I**), and that *Gsx1/2* TFs can function as repressors when acting as monomers
391 (Salomone et al., 2021), we propose that *Gsx1/2* represses *Noto* in ec4 neurons to reinforce ec4
392 identity. In contrast, *Noto* remains expressed in differentiated ec2 and ec3C neurons (**Fig. 7F,G,**
393 **Fig. S15I**). Based on the observations described in this section, we constructed a testable GRN
394 model for ec4 neuron specification (**Fig. 7J**).
395
396 **A broad community resource**
397 Our results for ec4 neurons support the value of the data presented here for generating testable
398 hypotheses to understand neuronal development in *Hydra vulgaris*. To support on-going and
399 future research in *Hydra* neuroscience, we have made our datasets accessible through interactive
400 web portals. Our scRNA-seq data can be explored via a user-friendly browser
401 (https://research.nhgri.nih.gov/HydraAEP/SingleCellBrowser/), which allows researchers to
402 visualize gene expression across neuron types using UMAP projections, dot plots, and trajectory
403 branching trees and spine plots; examples of plots derived from our portal are shown in figures 7
404 and S15. The portal also includes functional annotations and conversion tools for matching gene
405 IDs across *H. vulgaris* genomic and transcriptomic references. In parallel, we have integrated our
406 neuron-specific ATAC-seq data into our existing *H. vulgaris* AEP genome browser
407 (https://research.nhgri.nih.gov/HydraAEP/), enabling visualization of chromatin accessibility
408 near neural genes. Easy access to our portals is provided through the community Resource Hub
409 (openhydra.org), which offers centralized links to these high-resolution datasets. These tools are
410 designed to facilitate broad data accessibility and accelerate discovery in *Hydra* neuroscience.
411
412 **Discussion**
413 The number of research organisms used in neuroscience has grown significantly in recent years,
414 driven by advances in single cell transcriptomics that enable the exploration of neuronal diversity
415 across a broad range of species, including less conventional research organisms (Chari et al.,
416 2021; Cole et al., 2024; Fincher et al., 2018; Hulett et al., 2024; Sebé-Pedrós et al., 2018a; Sebé-
417 Pedrós et al., 2018b). Despite this progress, comprehensive transcriptional characterizations of
418 adult nervous systems remain limited. Here, we present a detailed molecular map of the adult *H.*
419 *vulgaris* nervous system based on ~35,000 single-cell transcriptomes, including differentiated

420 neurons and cells undergoing neurogenesis. In addition to confirming the twelve neuron subtypes
421 described previously (Siebert et al., 2019), we identified three novel subtypes in the ectodermal
422 nerve net. For each ectodermal neuron subtype, we defined unique cell markers and determined
423 their spatial distribution, providing the first high-resolution map of the *H. vulgaris* ectodermal
424 nerve net. Our study also confirmed the presence of the three previously described endodermal
425 neuron subtypes (Siebert et al., 2019) and we identify the en3 subtype as having sensory neuron
426 morphology. We propose a nomenclature scheme for *H. vulgaris* neuronal subtypes, designed to
427 evolve with future discoveries.
428
429 *H. vulgaris* offers numerous experimental advantages, including small size, optical transparency,
430 and accessible transgenic tools, which have accelerated efforts to link neural circuits to behavior.
431 In a landmark study, the first use of GCaMP in *H. vulgaris* revealed three distinct longitudinal
432 circuits (CB, RP1, RP2) that coordinate basic behaviors (Dupre and Yuste, 2017). Building on
433 this work, others have linked specific neuron types and circuits to behaviors such as
434 somersaulting and feeding (Giez et al., 2023; Yamamoto and Yuste, 2023), while innovative
435 tools now allow precise sensory stimulation and automated behavior tracking (Badhiwala et al.,
436 2018; Badhiwala et al., 2021; Escobar et al., 2024; Gonzales et al., 2017; Kim et al., 2024;
437 Tzouanas et al., 2021). The transcriptional and spatial map presented here complements and
438 expands this functional toolkit. By identifying subtype-specific markers and chromatin
439 accessibility patterns, our dataset will facilitate targeted genetic manipulation and help to
440 uncover the molecular mechanisms underlying neuronal function.
441
442 *H. vulgaris* is also a unique model for nervous system development and regeneration. In
443 uninjured adults, ISCs continuously replace neurons, supporting a dynamic state of homeostasis.
444 ISCs also enable complete nervous system regeneration following catastrophic injury, with a
445 single ISC capable of regenerating the entire nervous system (David and Murphy, 1977). These
446 features provide an opportunity to compare homeostatic and regenerative neurogenesis within a
447 single organism and uncover regulatory mechanisms specific to either process. Using trajectory
448 inference, we mapped the gene expression programs driving ISC differentiation into neuron
449 subtypes under steady-state conditions, establishing a foundation for future comparisons with
450 regeneration. Future research should aim to connect the molecular events driving neuronal

differentiation with the acquisition of neural circuit synchronization and the recovery of behaviors during regeneration. In recent work, we correlated the timing of ec5 neuron differentiation with re-synchronization of the CB circuit during behavioral recovery, highlighting the value of integrating molecular and functional readouts (Escobar et al., 2024). This approach will be essential for understanding how circuits are reassembled and behaviors restored after injury.

In summary, we provide an unparalleled molecular characterization of the *H. vulgaris* nervous system, including newly identified neuron subtypes, a spatial map of ectodermal neurons, and a framework for building GRNs to study neuronal differentiation. These resources advance *H. vulgaris* as a powerful model for nervous system biology and are publicly available through user-friendly web portals to support broad use by the research community and inspire new investigators to explore this regenerative system.

## Materials and Methods

### *Hydra vulgaris* Strains and Culturing Conditions

All *Hydra vulgaris* strains used in this study were cultured in *Hydra* medium (0.38mM $CaCl_2$, 0.32 mM $MgSO_4$ X 7H2O, 0.5 mM $NaHCO_3$, 0.08 mM $K_2CO_3$) at 18°C with a 12 hr light-dark cycle. *Hydra* were fed freshly hatched, premium grade *Artemia nauplii* (https://www.brineshrimpdirect.com) 2-4 times a week and were transferred into fresh *Hydra* medium 4-8 hours after feeding. Animals were starved for at least 48 hours prior to experimentation.

### Generation of transgenic strains

Transgenic *Hydra* lines *Tg(tba1c:mNeonGreen)*$^{cj1\text{-}gt}$ and *Tg(G022927:mNeonGreen)*$^{cj1\text{-}i}$ were generated using established protocols (Juliano et al., 2014b; Klimovich et al., 2019; Wittlieb et al., 2006). Promoter regions of 1901 bp (*tba1c*) and 1576 bp (*G022927*) upstream of their respective transcription start sites were PCR-amplified from *Hydra vulgaris* (strain 105) genomic DNA using Phusion™ High-Fidelity DNA Polymerase (ThermoFisher Scientific, #F530S), with a 55°C annealing and 60°C extension temperature. BamH1 and Xba1 restriction sites were added

482 via the forward and reverse primers for cloning into the reporter plasmid upstream of
483 mNeonGreen (**Table S7**).
484
485 PCR products were gel-purified, digested at 37°C for 10 minutes using BamH1 and Xba1
486 FastDigest enzymes (ThermoFisher Scientific, #FD0054 and #FD0684), and purified with the
487 Zymo DNA Clean & Concentrator Kit (#D4003). Digested fragments were ligated into the
488 BsAmp_mNG_tdT_backbone plasmid using Promega 10X Fast Ligase buffer (#M1801) for 4
489 hours at room temperature. Ligation products were transformed into DH5α-competent E. coli,
490 followed by plasmid miniprep (Qiagen, #27106) and Sanger sequencing for insert verification.
491 Final plasmids were amplified via maxiprep (Qiagen, #12162) and eluted in autoclaved Milli-Q
492 water for microinjection.
493
494 Microinjections were performed into 1-cell stage *Hydra vulgaris* AEP embryos as previously
495 described (Juliano et al., 2014b), with the following modifications: (1) the injection solution
496 consisted of 1 µL of 0.5% phenol red (Sigma #P0290-100ML) mixed with 6 µL of plasmid
497 DNA, followed by centrifugation; (2) embryos were fertilized for 1–2 hours before injection.
498 Embryos were injected using an Eppendorf FemtoJet 4x and InjectMan NI 2 microinjector under
499 a Leica M165 C stereo microscope. Transgenic animals were propagated asexually. Germline
500 transmission was confirmed by crossing male and female Tg(tba1c:mNeonGreen)cj1-in animals
501 to produce an F1 population.
502
503 **Naming scheme for transgenic lines**
504 We have established a standardized naming convention modeled after the zebrafish community
505 guidelines to ensure that each transgenic Hydra line has a unique and informative identifier. The
506 format is: **Tg(promoter:gene_or_reporter)^lab#-lineage**, where **Tg** designates the line as
507 transgenic. The **promoter** indicates the regulatory element driving expression, and
508 **gene_or_reporter** specifies the transgene product (e.g., GFP or a specific gene). The **lab** field
509 uses the initials of the principal investigator to denote the lab of origin (e.g., "cj" for the Juliano
510 lab, "rs" for the Steele lab), followed by a unique allele number to distinguish independent
511 integration events. The **lineage** suffix indicates the cell lineage in which the transgene is
512 integrated—**ec** for ectoderm, **en** for endoderm, **in** for interstitial—or **gt** if the transgene has been

513 germline transmitted. Using this naming scheme, we have renamed the transgenic line previously
514 described in both Siebert et al., 2019 and Keramidioti et al., 2024 as *Tg(actin1:GFP)*$^{rs3\text{-}in}$. In this
515 line, GFP is driven by the actin1 promoter and is integrated into the interstitial lineage. However,
516 likely due to positional effects of the integration site, GFP expression is largely restricted to
517 neurons and neuronal progenitors.
518
519 **Dissociation of *Hydra* into Single Cells for Fluorescence-Activated Cell Sorting (FACS)**
520 Forty to forty-five asexual, starved *Hydra* polyps were washed three times in sterile *Hydra*
521 medium and transferred to 1.5 mL Eppendorf tubes. Medium was replaced with ~75 U/mL
522 Pronase E (VWR, E629-1G) in 1 mL of room temperature Dissociation Medium (DM; 5 mM
523 $CaCl_2 \cdot 2H_2O$, 1 mM $MgSO_4 \cdot 7H_2O$, 2.8 mM KCl, 11 mM HEPES, 0.67 mM $Na_2HPO_4$, 0.44 mM
524 $KH_2PO_4$, 5 mM sodium pyruvate, and 5 mM sodium citrate·$2H_2O$; pH 6.9–7.0). Polyps were
525 dissociated for 90 minutes at 22–24°C with gentle agitation. Dissociated cells were transferred to
526 a small Petri dish in 2 mL DM, gently pipetted up and down 5–10 times using a 1000 µL pipette,
527 and filtered through a pre-soaked 70 µm Corning® strainer (Sigma #CLS431751) into a 50 mL
528 conical tube containing 1 mL DM. The Petri dish and strainer were rinsed with an additional 1
529 mL DM. Cells were centrifuged at 300 g for 5 minutes (brake off), resuspended in 2 mL DM,
530 and re-spun under the same conditions. The pellet was resuspended in 1 mL DM and filtered
531 through a pre-soaked 40 µm Corning® strainer (Sigma #CLS431750) into a fresh conical tube.
532
533 **Fluorescence-Activated Cell Sorting (FACS)**
534 FACS was used to collect fluorescent cells from transgenic *Hydra* lines
535 *Tg*(*tba1c:mNeonGreen*)$^{(cj1-f1)}$ and Tg(*actin1:GFP*)$^{(rs3-in)}$ for both scRNA-seq and ATAC-seq. For
536 scRNA-seq, one library was prepared from bud-free *Tg*(*tba1c:mNeonGreen*) $^{(cj1-f1)}$ animals, and
537 three libraries from *Tg*(*actin1:GFP*) animals, one from bud-free and two from budding animals.
538 For ATAC-seq, two libraries were generated from bud-free *Tg*(*tba1c:mNeonGreen*) $^{(cj1-f1)}$
539 animals and four from non-budding *Tg*(*actin1:GFP*)$^{(rs3-in)}$ animals. In all cases, fluorescent cells
540 were sorted using a 100 µm nozzle on a MoFlo Astrios EQ Cell Sorter (Beckman Coulter) into
541 400 µL dissociation medium (DM) supplemented with 0.01% BSA. For ATAC-seq, sorting was
542 performed into DNA LoBind® tubes (Eppendorf #0030122348), and 1 mL of additional DM
543 was added prior to sorting to bring the final volume to 2 mL. Cells were pelleted at 300 g for 5

544 minutes and resuspended in 50 µL DM + 0.01% BSA for scRNA-seq, or counted using Hoechst
545 staining and a Fuchs-Rosenthal hemocytometer (InCyto DHC-F01-5) for ATAC-seq. FACS
546 gating used *Tg*(*actin1:GFP*)$^{(rs3-in)}$ and *Hydra vulgaris* AEP as positive and negative controls,
547 respectively, and followed Siebert et al. (2019).
548
549 **10x Genomics Single Cell RNA Library Preparation, Sequencing, and Analysis**
550 A step-by-step description of the scRNA-seq cluster analysis, including all relevant code, is
551 provided in the markdown document *01_hydraNeuronAtlas* available at
552 https://github.com/cejuliano/neuron_single_cell.
553
554 scRNA-seq libraries were built using the Chromium Next GEM Single Cell 3' Kit v3.1 (10x
555 Genomics) at the UC Davis Genome Center with a target recovery of 10,000 cells. Final cell
556 counts were: 11,183 for *Tg*(*tba1c:mNeonGreen*)$^{(cj1-f1)}$, and 7,988, 7,374, and 6,124 for the three
557 *Tg*(*actin1:GFP*)$^{(rs3-in)}$ libraries. scRNA-seq libraries generated with the 10x Genomics Chromium
558 platform were processed using Cell Ranger v4.0.0 following the manufacturer's instructions and
559 aligned to the *Hydra vulgaris* AEP transcriptome (Cazet et al., 2023). The average alignment rate
560 across libraries was 59.9%, with individual alignment rates of 62.5% (*Tg(tba1c:mNeonGreen)*$^{cj1-}$
561 $^{gt}$), 58.6% (*Tg(actin1:GFP)*$^{rs3}$ non-budding), 60.4% (budding1), and 58.1% (budding2).
562 Previously published scRNA-seq data from Siebert et al. (2019), mapped against the same AEP
563 reference, were subsetted to include interstitial stem cells, neuronal progenitors, and
564 differentiated neurons based on cluster identity. All downstream analyses were conducted using
565 Seurat v4.0.5 (Butler et al., 2018; Hao et al., 2021; Satija et al., 2015; Stuart et al., 2019).
566
567 Individual Seurat objects were generated per library and filtered to retain cells with 300–7,000
568 detected genes, 500–50,000 transcripts, and <5% mitochondrial reads (Satija et al., 2015; Siebert
569 et al., 2019). Data quality was assessed with violin plots (**Fig. S3**). For Tg(actin1:GFP)$^{rs3-in}$
570 libraries, which contain a minor population of gland cells and nematocytes, clustering was
571 performed and clusters expressing canonical markers for these cell types were identified and
572 excluded (Siebert et al., 2019). To account for technical variation, module scores were calculated
573 for a curated set of "stress markers" and regressed during dataset integration.
574

575 An initial dataset integration was performed using the SCTransform workflow (Hafemeister and
576 Satija, 2019), incorporating twelve previously published whole-animal Drop-seq libraries
577 (subsetted for interstitial stem cells, neurons, and neural progenitors), two published Drop-seq
578 libraries of sorted neurons from the Tg(actin1:GFP)$^{(rs3)}$ line, and the four newly generated 10X
579 Genomics libraries in this study. SCTransform was first used to normalize the datasets, with
580 regression against stress marker module scores and the percentage of mitochondrial RNA reads.
581 Integration was performed with Seurat::SelectIntegrationFeatures (nfeatures = 3000),
582 PrepSCTIntegration, FindIntegrationAnchors, and IntegrateData. Dimensionality reduction was
583 carried out with Seurat::RunPCA. The number of principal components (PCs) retained was based
584 on the point where the percent variance change between PCs dropped below 0.1%. A preliminary
585 clustering and UMAP was generated using the Louvain approach [Seurat::*FindClusters*,
586 (*dims.use* = 1:25), Seurat::*FindNeighbors* (*resolution* = 0.8), Seurat::*RunUMAP*, (*dims.use* =
587 1:25)]. One cluster expressing stress markers, mitochondrial genes, or non-neuronal lineage
588 markers was excluded.
589
590 Following removal of the non-neuronal and high-stress cluster, datasets were reintegrated using
591 the same SCTransform workflow. PCA was repeated, and clustering was performed with 21 PCs
592 (resolution = 0.8). A new UMAP was generated using RunUMAP (min.dist = 0.25, spread =
593 0.7). Clusters were annotated based on expression of known markers for neuron subtypes,
594 progenitors, and ISCs. Of the resulting clusters, four could not be definitively identified. Marker
595 expression from these clusters was cross-referenced with whole-animal scRNA-seq datasets
596 (Cazet et al., 2023; Siebert et al., 2019). One unknown cluster was enriched for stress genes and
597 excluded. The remaining three expressed markers associated with ec1A, ec1B, and ec5 neurons
598 and were retained. We ran Seurat::FindAllMarkers() to identify molecular markers for the 3
599 remaining novel clusters and validated these novel subtype markers by FISH (see below).
600
601 **ATAC-seq Library Preparation, Sequencing, and Analysis**
602 Neuron-enriched ATAC-seq libraries were generated using a modified OMNI-ATAC protocol
603 (Cazet et al., 2023; Corces et al., 2017; Siebert et al., 2019). FACS-isolated cells (described
604 above) were pelleted at 1,000 g for 5 min at 4°C and resuspended in 50 µL freshly prepared,
605 chilled resuspension buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$) supplemented

606 with 0.1% Tween-20, 0.1% NP-40, and 0.01% digitonin. After 3 minutes on ice, lysis was
607 quenched with 1 mL chilled RSB + 0.1% Tween-20, inverting tubes 3x to mix. Nuclei were
608 pelleted at 500 g for 10 min at 4°C and resuspended in tagmentation mix [1x TD buffer (Illumina
609 #20034197), 33% PBS, 0.01% digitonin, 0.1% Tween-20, 5 µL TDE1 (Illumina)]. Samples were
610 tagmented for 30 min at 37°C in a ThermoShaker (1,000 RPM), quenched with 250 µL Qiagen
611 PB buffer (MinElute Kit #28004), and stored at -20°C.
612
613 Tagmented DNA was purified using the Qiagen MinElute Kit, eluted in 21 µL EB buffer, and
614 amplified with NEBNext 2X Master Mix (NEB M0541S). PCR cycle numbers were determined
615 by qPCR (Buenrostro et al., 2015) (**Table S5**). Libraries were subjected to two rounds of size-
616 selected (100–700 bp) using AMPure XP beads[HM1] (Beckman Coulter #A63881), quantified
617 with Qubit dsDNA HS Assay (ThermoFisher #Q32851), and assessed with a Bioanalyzer High-
618 Sensitivity DNA Kit (Agilent #5067-4626) at the UC Davis DNA Core. Based on concentration
619 and trace quality, two of five Tg(tba1c:mNeonGreen)cj1-gt replicates and four of six
620 Tg(actin1:GFP)rs3-in replicates were selected for sequencing. Libraries were pooled and
621 sequenced using 2×150 bp reads on an Illumina HiSeq 4000 (*Tg(tba1c:mNeonGreen)cj1-gt*) or
622 HiSeq X Ten (*Tg(actin1:GFP)rs3-in*).
623
624 A step-by-step description of the ATAC-seq analysis, including all relevant code, is provided in
625 the markdown document *03_hydraNeuronRegulation* available at
626 https://github.com/cejuliano/neuron_single_cell.
627
628 To analyze the ATAC-seq data collected from whole strain AEP *H. vulgaris* polyps, we first
629 filtered the raw reads using Trimmomatic (Bolger et al., 2014) to remove stretches of low-
630 quality base-calls and contaminating adapter sequence. The filtered reads were then aligned to
631 the AEP assembly using Bowtie2 (Langmead and Salzberg, 2012). To remove mitochondrial
632 reads, we also aligned the ATAC-seq data to the *H. vulgaris* mitochondrial genome (Voigt et al.,
633 2008) and subsequently discarded any reads that aligned to the mitochondrial and nuclear
634 genome references using Picard Tools (broadinstitute.github.io/picard/). We next identified and
635 removed PCR duplicates from the aligned data using Samtools (Li et al., 2009) and Picard Tools.
636 Peaks were called using a pipeline adapted from Cazet et al. (2023), originally based on the

637 ENCODE ATAC-seq workflow (Landt et al., 2012). We called peaks for each ATAC-seq
638 biological replicate using MACS2 (Zhang et al., 2008). To generate a consensus peakset of
639 biologically reproducible ATAC-seq peaks, we first calculated irreproducible discovery rate
640 (IDR) (Li et al., 2011) peak scores for each pairwise combination of biological replicates. We
641 defined a reproducible peak as one that received an IDR score ≤ 0.1 for at least two pairwise
642 comparisons between biological replicates. TSS enrichment, self-consistency, and rescue ratios
643 were computed for quality control (Cazet et al., 2023; Siebert et al., 2019). Neuron-enriched
644 libraries were of comparable quality to whole-animal ATAC-seq libraries (Cazet et al., 2023)
645 (**Table S5**).
646
647 To identify differentially accessible regions, read counts across all peaks were quantified using
648 DiffBind and annotated to the nearest transcription start site (TSS) with UROPA (Kondili et al.,
649 2017). Differential accessibility between *Tg(tba1c:mNeonGreen)cj1-gt*, *Tg(actin1:GFP)rs3-in*,
650 and AEP whole animal datasets (Cazet et al., 2023) was assessed using edgeR (Robinson et al.,
651 2010). Results were saved as RData objects and summary tables, and genomic tracks were
652 visualized using Gviz (Hahne and Ivanek, 2016). To assess whether neuron-enriched ATAC-seq
653 libraries show greater accessibility near the TSS of neuronal genes, we first identified genes
654 enriched (NeuroG) or depleted (nonNeuroG) in differentiated neurons using Seurat analysis of
655 *H. vulgaris* whole-animal scRNA-seq data (Cazet et al., 2023). ATAC peaks associated with
656 these genes were extracted from edgeR comparisons between *Tg(tba1c:mNeonGreen)cj1-gt*
657 /whole animal AEP and *Tg(actin1:GFP)rs3-in* /whole animal AEP datasets. Log fold changes
658 for NeuroG- and nonNeuroG-associated peaks were visualized using ggplot2 and compared
659 using a two-sided t-test to evaluate statistical significance.
660
661 **Generation of Developmental Trajectories**
662 A step-by-step description of the developmental trajectory analysis, including all relevant code,
663 is provided in the markdown document *02_hydraNeuronTree* available at
664 https://github.com/cejuliano/neuron_single_cell.
665
666 To reconstruct neuronal differentiation trajectories in Hydra, we used URD (Farrell et al., 2018)
667 to trace the emergence of 11 neuron subtypes from interstitial stem cells (ISCs). Initial analyses

revealed strong batch effects between 10x Genomics and Drop-seq datasets that impaired proper cell-to-cell transition mapping. Given that the Drop-seq data represented a minority of cells (~5,400), we excluded it from trajectory inference and focused solely on the higher-resolution 10x scRNA-seq data. We used the SeuratToURDv3 function to generate an URD object, and manually added in the integrated data from our Seurat object (urd.obj@logupx.data <- seurat.obj@assays$integrated@data). Next, we used the URD::urdSubset function to eliminate the Drop-seq data and cell clusters that pilot analyses demonstrated lacked clear intermediate progenitors linking them back to the ISCs (ec1C, ec1D, and ec1E clusters).

Outliers, cells with poor connectivity, were removed based on k-nearest neighbor distances using URD::knnOutliers with parameters: x.max = 86, slope.r = 0.12, int.r = 79, slope.b = 1.1, and int.b = 7.75 (209 cells removed). Doublets were identified based on high co-expression of mutually exclusive gene modules (Siebert et al., 2019). NMF module scores were scaled between 0-1, and visualized on the UMAP to confirm cell-type specificity (**Fig. S6**). Doublets were called using URD::NMFDoubletsDetermineCells with parameters: module.expressed.thresh = 0.2, frac.overlap.max = 0.07, frac.overlap.diff.max = 0.15. This process flagged 1,257 potential cell doublets, which were subsequently removed.

Pseudotime trajectories were computed using a diffusion map (URD::CalcDM, with knn = 100, sigma.use = local, distance = cosine), optimized for high progenitor–differentiated cell connectivity and low noise across unrelated terminal branches. The ISC population (identified by Hy-icell1 marker G002332) was designated as the root (Siebert et al., 2019). Terminal tips were defined using Infomap-Jaccard clustering (URD::graphClustering, num.nn = 120), and selected based on late pseudotime and differential gene expression.

Pseudotime was calculated using URD::floodPseudotime (n = 50, minimum.cells.flooded = 2), and transition matrix bias was tuned with optimal.cells.forward = 0, max.cells.back = 200 to allow transitions through densely populated differentiated states. Biased random walks were performed (n.per.tip = 50,000, root.visits = 1), and visitation patterns were checked on UMAPs. In cases of multiple tips per subtype, they were merged using URD::combineTipVisitation. Final tree construction was carried out with URD::buildTree using: divergence.method = "preference",

699 save.all.breakpoint.info = TRUE, cells.per.pseudotime.bin = 25, bins.per.pseudotime.window =
700 8, p.thresh = 1e-6, and min.cells.per.segment = 10.
701
702 To examine the temporal dynamics of gene expression during differentiation, we used spline-
703 based modeling. Genes expressed in at least 1% of cells were considered. Expression profiles
704 across pseudotime were fit and visualized with URD::geneSmoothFit (spar = 0.875) within
705 lineage-specific segments.
706
707 **Identification of gene modules using non-negative matrix factorization (NMF)**
708 To identify gene co-expression modules (i.e., metagenes) in our neuronal scRNA-seq dataset, we
709 performed non-negative matrix factorization (NMF) using the cNMF package (Kotliar et al.,
710 2019) (Kotliar et al., 2019). We used raw read counts from all neuronal and interstitial stem cell
711 transcriptomes across both 10x Genomics and Drop-seq libraries. Prior to factorization, data
712 were normalized and filtered using the 'prepare' function to exclude genes with low variability
713 across cells. Because the optimal number of modules (*k*) must be empirically determined, we
714 used the 'factorize' function to perform independent NMF analyses for *k* values ranging from 5
715 to 100 in increments of 5. For each *k*, 200 independent runs were performed to assess metagene
716 reproducibility. Results were compiled using the 'combine' function and evaluated for
717 robustness and accuracy with the '*k*_selection_plot' function. This initial screen suggested that
718 the most stable solutions fell between *k* = 25 and 35. To refine our selection, we repeated the
719 analysis for all individual k values within that range. Based on this finer-grained analysis, we
720 selected *k* = 27 as the optimal number of modules. We then used the 'consensus' function to
721 generate a final set of 27 consensus metagenes by filtering out runs that deviated most from the
722 majority of independent solutions.
723
724 **Probe Generation for Fluorescent in Situ Hybridization**
725 To generate labeled RNA probes, PCR products (**Table S7**) were amplified from oligo-dT-
726 primed cDNA derived from *Hydra vulgaris* AEP. Reverse primers included T7 or SP6 promoters
727 (Qiagen #28506). Amplicons were cloned into the Invitrogen Zero Blunt vector (#K2700-20),
728 and correct inserts were confirmed by colony PCR and Sanger sequencing. Plasmids were
729 amplified via miniprep (Qiagen #27106) and used as templates for a second round of PCR to

730 generate transcription templates. Approximately 250 ng of gel-purified product was used as a
731 transcription template for the Roche DIG RNA Labeling Kit (Sigma #11175025910) with DIG-
732 U-11 or FITC-U-12 labeling mix (Sigma #11685619910). RNAse-free conditions were
733 maintained using Ambion RNAse-in, and resulting RNA probes were purified with the Zymogen
734 RNA Clean & Concentrator-25 kit (Zymo Research #R1017), diluted to 750 ng in 7.5 µL $H_2O$,
735 and stored at –80°C.
736
737 **Fluorescent in Situ Hybridization**
738 Fixation and Bleaching (Day 1): Starved *Hydra vulgaris* AEP polyps were washed in *Hydra*
739 medium (HM), relaxed in 2% urethane in HM for one minute, then fixed in 4%
740 paraformaldehyde (PFA) in HM for 1 hour at room temperature (RT) with rocking. After
741 fixation, samples were washed 3 × 10 min in PBT (0.1% Tween-20 in PBS) and dehydrated in a
742 stepwise methanol series to 100% MeOH, followed by bleaching overnight at –20°C in fresh
743 100% MeOH. Rehydration, Permeabilization, and Probe Hybridization (Day2): Samples were
744 rehydrated through 66% and 33% MeOH/PBT washes, followed by 3 × 5 min washes in PBT.
745 Tissue was permeabilized with 10 µg/mL proteinase K in PBT for 5 minutes, quenched in 4
746 mg/mL glycine, and washed 3 × 5 min in PBT. Samples were then washed sequentially in 0.1 M
747 triethanolamine in PBT ± acetic anhydride (3 and 6 µL/mL), refixed in 4% PFA (1 hour), washed
748 in PBT, and equilibrated in 2× SSC. Samples were prehybridized in 50% 2× SSC/50%
749 hybridization solution (HS: 50% formamide, 5× SSC, 1× Denhardt's, 100 µg/mL heparin, 0.1%
750 Tween-20, 0.1% CHAPS), first at RT then at 56°C. They were incubated for 2 hours in HS with
751 10 µL/mL sheared salmon sperm DNA. For sample, ~180 ng of DIG- and/or FITC-labeled RNA
752 was denatured at 85°C for 5–10 min, cooled to 56°C, and added to 400 µL HS. Samples were
753 hybridized in 450 µL total probe mix at 56°C for ~65 hours, with gentle mixing once every 24
754 hours. Probe Washes and Antibody Staining (Day3): Samples were washed at 56°C with
755 decreasing HS/SSC concentrations (100%, 75%, 50%, 25%) followed by 2 × 30 min washes in
756 2× SSC + 0.1% CHAPS (first at 56°C, then RT). Samples were washed 4 × 10 min in MABT
757 (100 mM maleic acid, 150 mM NaCl, 0.1% Tween-20, pH 7.5), then blocked in MABT + 1%
758 BSA for 1 hour and in blocking buffer (80% MABT + 1% BSA / 20% sheep serum) for 2 hours
759 at 4°C. Anti-FITC-POD (Sigma #11426346910) or Anti-DIG-POD (Sigma #11207733910) was
760 added at 1:2000 in blocking buffer and incubated overnight at 4°C. Tyramide Detection and

761 Second Antibody Application (Day 4): After returning to RT, excess antibody was removed with
762 2 × 20 min MABT-BSA and 5 × 20 min MABT washes. Samples were washed 2 × 5 min in 100
763 mM borate buffer + 0.1% Tween, then stained with 75 µL tyramide solution (100 mM borate
764 buffer, 2% dextran sulfate, 0.1% Tween-20, 0.003% $H_2O_2$, 0.15 mg/mL 4-iodophenol, 1:100
765 Alexa Fluor 488 or 594 tyramide) for 25 minutes. The reaction was stopped with 4 × PBT
766 washes and quenched in 100 mM glycine (pH 2.0) for 10 minutes, then washed 5 × 5 min in
767 PBT. At this point, samples could either be mounted for imaging (see below) or processed for
768 double FISH. For double labeling, samples were blocked for 2 hours at 4°C in blocking buffer
769 (80% MABT + 1% BSA / 20% sheep serum), then incubated overnight at 4°C in 1:2000 Anti-
770 DIG-POD (Sigma #11207733910) in blocking buffer, protected from light. Second Tyramide
771 Detection and Mounting (Day 5): Following antibody incubation, samples were washed and
772 processed for the second round of tyramide detection as described above, using the alternate
773 fluor. After staining, samples were again quenched in 100 mM glycine (pH 2.0) and washed 5 ×
774 5 minutes in PBT. To prepare for imaging, tissues were stained with 1:1000 Hoechst for 30
775 minutes, then dehydrated through a graded series of glycerol (30%, 50%, 80% in PBT) for at
776 least 1 hour in each step. Specimens were mounted in 80% glycerol and stored at 4°C until
777 imaging. Confocal images were acquired using the Zeiss 980 LSM with Airyscan 2 confocal
778 microscope. For whole animal images, max projections were stitched together using ImageJ.
779

780 **In Situ Hybridization Chain Reaction**

781 The ec1E cluster could not be identified using traditional FISH and instead required detection via
782 In Situ Hybridization Chain Reaction (HCR) (Choi et al., 2018). Probe Design and Synthesis:
783 probes were designed using the in situ_probe_generator tool
784 (https://github.com/guttalab/in_situ_probe_generator) (Kuehn et al., 2022). Probes were
785 designed to target transcripts of interest and synthesized by IDT. Sample Preparation, Fixation,
786 and Permeabilization: *Hydra* were washed twice in sterile *Hydra* Medium (HM) and relaxed in
787 2% urethane in HM for 1 minute. Samples were fixed in 4% paraformaldehyde (PFA) in HM for
788 1 hour with gentle rocking. Fixed samples were washed three times for 10 minutes each in
789 phosphate-buffered saline with Tween (PBT), followed by gradual dehydration through 5-minute
790 washes in increasing concentrations of methanol (MeOH) in PBT: 33%, 66%, and 100%.
791 Samples were then washed twice for 30 minutes in 100% MeOH and stored in fresh 100%

792 MeOH at -20°C overnight. Probe Hybridization: Samples were rehydrated by sequential 5-
793 minute washes in 66% MeOH in water and 33% MeOH in PBT, followed by three 5-minute
794 washes in PBT. Samples were then incubated in 1ml of 50% probe hybridization buffer in PBT
795 for 5 minutes, followed by incubation in 300µl of 100% probe hybridization buffer for 1 hour at
796 37°C. Probes (1–2 pmol each) were diluted into pre-warmed hybridization buffer (250µl total
797 volume), and samples were incubated in this hybridization mix overnight at 37°C. Probe
798 Detection and Signal Amplification: Following hybridization, samples were washed four times
799 for 15 minutes each with 500µl of probe wash buffer pre-warmed to 37°C. Samples were then
800 washed three times for 5 minutes each in 5x SSC with 0.1% Tween-20 (SSCT) at room
801 temperature, followed by a 5-minute wash in PBT. Samples were equilibrated in 300µl of
802 amplification buffer for 30 minutes. Meanwhile, hairpin amplifiers were prepared by heating
803 each pair of hairpins separately to 95°C for 90 seconds, followed by cooling to room temperature
804 in the dark for 30 minutes. The hairpins (15 pmol each) were then combined in amplification
805 buffer (250µl total volume), and samples were incubated in the hairpin mix overnight in the dark.
806 Post-Amplification Washes and Mounting: Following amplification, samples were washed twice
807 for 5 minutes and twice for 30 minutes in 5x SSCT. DAPI was optionally added during the first
808 30-minute wash. Samples were stored in 5x SSCT at 4°C until imaging. Prior to imaging,
809 samples were washed twice for 5 minutes in PBS and mounted in ProLong Glass mounting
810 medium. Confocal images were acquired using the Zeiss 980 LSM with Airyscan 2 confocal
811 microscope.
812
813 **Immunofluorescence**
814 *Hydra* polyps were anesthetized in 2% urethane in *Hydra* medium for 4–5 minutes, then fixed in
815 4% paraformaldehyde (PFA) in 1× *Hydra* PBS (120 mM NaCl, 40 mM K2HPO4, 12 mM
816 NaH2PO4; pH 7.2-7.4) for 1 hour at room temperature. Following fixation, animals were washed
817 three times with 1× *Hydra* PBS for 10 minutes each on a shaker. Fixed animals were then cut as
818 needed and permeabilized in 0.5% Triton X-100 in 1× *Hydra* PBS for 15 minutes in the dark.
819 Blocking solution was prepared by dissolving 0.05 g BSA and 50 µL 10% Triton X-100 in 4950
820 µL 1× *Hydra* PBS, and animals were incubated in this solution for 20 minutes in the dark.
821 Primary antibodies were diluted in blocking solution and applied at the following concentrations:
822 pan-neural antibody (1:1000, raised in rabbit) (Keramidioti et al., 2024), anti-GFP (1:500, raised

823 in chicken, Aves Labs, GFP-1020), anti-mNeonGreen (1:50, raised in mouse, Chromotek, 32f6),
824 anti-GLWa (1:10, raised in mouse) (Koizumi et al., 2015), and anti-RFa (1:50, raised in mouse)
825 (Koizumi et al., 2015). 100 µL of the appropriate antibody dilution was added per well (one
826 animal per well), and samples were incubated overnight at 4 °C. The following day, animals
827 were washed three times in 1× *Hydra* PBS for 10 minutes each on a shaker in the dark.
828 Secondary antibodies were diluted in blocking solution and included Alexa Fluor 488 donkey
829 anti-chicken (1:400, Dianova, 703-546-155), Cy3 donkey anti-rabbit (1:400, Dianova, 711-165-
830 152), Alexa Fluor 488 donkey anti-mouse (1:300, Dianova, 715-545-150), Cy3 donkey anti-
831 mouse (1:200, Dianova, 715-166-151), and Alexa Fluor 647 donkey anti-mouse (1:300, Dianova,
832 715-605-150). 100 µL of secondary antibody solution was added per well and samples were
833 incubated for 2 hours at room temperature in the dark, followed by three 10-minute washes in 1×
834 *Hydra* PBS. DAPI (Sigma) was used to stain nuclei at a dilution of 1:1000 in 1× *Hydra* PBS,
835 incubating 100 µL per well for 7 minutes. After a final 5-minute wash in 1× *Hydra* PBS, the
836 stained specimens were mounted on microscope slides using VECTASHIELD HardSet Antifade
837 Mounting Medium to preserve fluorescence. Bee wax was applied around the edges of the
838 coverslip to prevent compression of the specimens, and the coverslip was sealed with transparent
839 nail polish. Preparations were stored at 4°C until imaging. Fluorescent confocal microscopy was
840 performed on a Leica SP5 scanning confocal microscope (Leica Microsystems, Heidelberg,
841 Germany). Z-shift correction and RGB stack generation were performed using the StackGroom
842 plugin in ImageJ with the "3channels" command. The resulting RGB stacks were further
843 processed in ImageJ, and all images were generated as maximum intensity projections using the
844 "Image > Stack > Z Project" function.
845

846 **Acknowledgments**

847 We would like to thank Rob Steele for providing the *Tg(actin1:GFP)*$^{rs3\text{-}in}$ line, which was
848 foundational to this study; the Juliano lab, Farrell lab, and David lab for invaluable discussions
849 and feedback; Bruce Draper, Rob Steele, and Stefan Siebert for manuscript review and feedback;
850 Clara Stuligross for analysis help and moral support; Shun Hamada and Osam Koizumi for
851 providing GLWa and RFa antibodies; Yohana Berhe for help with confocal images used in Fig.
852 S2; Diana Burkart-Waco, Siranoosh Ashtari, and Emily Kumimoto from the DNA Technologies
853 and Expression Analysis Core at the UC Davis Genome Center (supported by NIH Shared

Instrumentation Grant 1S10OD010786-01) for technical advice and assistance with library preparation and sequencing; Bridget McLaughlin and Jonathan Van Dyke from the UC Davis Flow Cytometry Shared Resource Laboratory for technical advice and assistance with FACS [(supported by NCI P30 CA093373 (Comprehensive Cancer Center), S10 OD018223 (Astrios Cell Sorter), S10 RR 026825 (Fortessa Cytometer), and NIH NCRR C06-RR12088 (Tupper Hall room 3425 sorting suite renovations)].

**Competing Interests**

No competing interests declared.

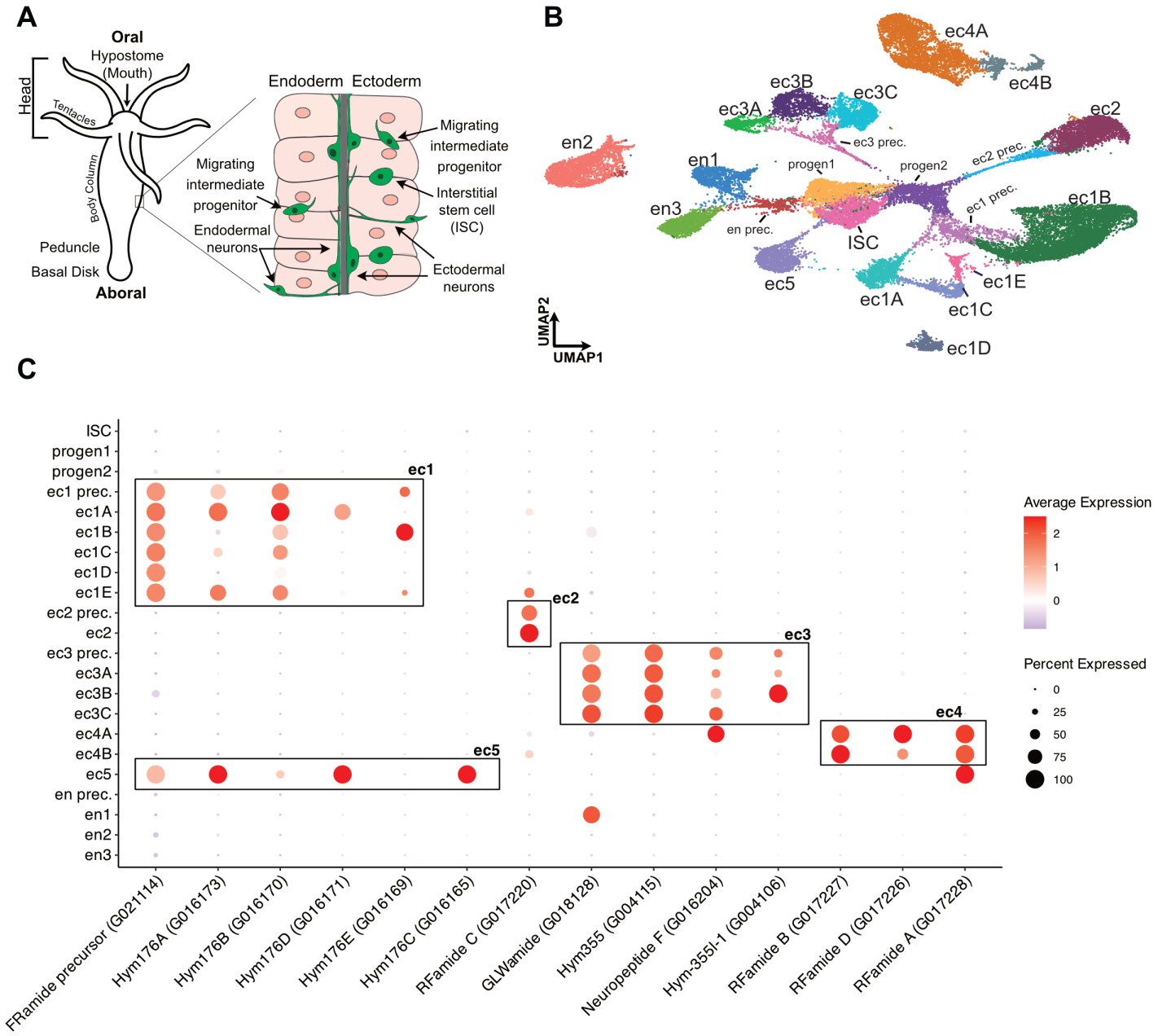
**Funding**

This work was supported by a National Institutes of Health (NIH) grant (R35 GM133689) and a W. M. Keck Foundation grant to C.E.J., and by the NICHD Intramural Research Program to J.A.F. (ZIAHD008997). H.M.L., A.S.P., and J.T. were supported by an NIH T32 predoctoral training program (T32 GM007377). H.M.L. was also supported by an NIH F31 fellowship (F31 GM149176), and J.T. was also supported by a second NIH T32 predoctoral training program (T32 GM153586).

**Data and resource availability**

The raw sequencing data in this study have been submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA1261600. Step-by-step descriptions of all computational analyses conducted as part of this study, including all relevant code, formatted both as markdown and HTML documents, are available at GitHub (https://github.com/cejuliano/neuron_single_cell).

## Figures


A
Oral
Hypostome
(Mouth)
Head
Tentacles
Body Column
Peduncle
Basal Disk
Aboral
Endoderm Ectoderm
Migrating
intermediate
progenitor
Interstitial
stem cell
(ISC)
Ectodermal
neurons
Endodermal
neurons
B
ec4A
ec4B
ec3B
ec3C
ec3A
ec2
en2
en1
ec3 prec.
progen1
progen2
ec2 prec.
ec1 prec.
ec1B
en3
en prec.
ISC
ec5
ec1A
ec1E
ec1C
ec1D
UMAP2
UMAP1
C
ISC
progen1
progen2
ec1 prec.
ec1A
ec1B
ec1C
ec1D
ec1E
ec2 prec.
ec2
ec3 prec.
ec3A
ec3B
ec3C
ec4A
ec4B
ec5
en prec.
en1
en2
en3
ec1
ec2
ec3
ec4
ec5
FRamide precursor (G021114)
Hym176A (G016173)
Hym176B (G016170)
Hym176D (G016171)
Hym176E (G016169)
Hym176C (G016165)
RFamide C (G017220)
GLWamide (G018128)
Hym355 (G004115)
Neuropeptide F (G016204)
Hym-355I-1 (G004106)
RFamide B (G017227)
RFamide D (G017226)
RFamide A (G017228)
Average Expression
2
1
0
Percent Expressed
0
25
50
75
100
D
Neurons
(elav positive)
Non-neurons
(elav negative)
Ectodermal Neurons
(budhead negative)
Endodermal Neuron Types
(budhead positive)
Ectodermal Types
(neuropeptide expression)
ec1
ec2
ec3
ec4
ec5
en1
en2
en3
Ectodermal Subtypes
(spatial segregation)
ec1A
ec1B
ec1C
ec1D
ec1E
ec3A
ec3B
ec3C
ec4A
ec4B

885 **Figure 1. Single cell RNA sequencing reveals fifteen transcriptionally distinct neuron**
886 **subtypes comprising the *Hydra vulgaris* nervous system. (A)** The *Hydra* body is a radially
887 symmetric, hollow tubular animal arranged around an oral-aboral axis. The oral end consists of
888 the hypostome and tentacles (collectively referred to as the "head"), while the aboral end
889 comprises the peduncle and basal disk. The body is formed by two epithelial monolayers, the
890 endoderm and ectoderm, separated by an extracellular matrix. Neurons reside within the
891 interstitial spaces between epithelial cells, forming two distinct nerve nets: one in the ectodermal
892 layer and another in the endodermal layer (Keramidioti et al., 2024). Interstitial stem cells
893 (ISCs), which give rise to neurons, are confined to the ectoderm, but intermediate neural
894 progenitors migrate to generate endodermal neuron subtypes. **(B)** Single-cell RNA sequencing
895 (scRNA-seq) was performed using two methods: Chromium Single-Cell Gene Expression (10x
896 Genomics) (29,671 cells from this study) and Drop-Seq (5,400 cells from Siebert et al., 2019). A
897 UMAP representation displays cell clusters annotated by their respective cell states: ISC
898 (interstitial stem cell), progen (progenitor) prec (precursors), ec (ectodermal neurons), and en
899 (endodermal neurons). **(C)** Dot plot showing the expression of neuropeptide genes across *Hydra*
900 *vulgaris* neuron clusters. Dot color represents the average expression level of each gene within a
901 cluster, while dot size indicates the percentage of cells in the cluster expressing the gene. The
902 ec1, ec2, ec3, ec4, and ec5 neuron types have distinct neuropeptide expression patterns.
903 **(D)** Schematic outlining the classification process for *H. vulgaris* neurons. First, neurons were
904 identified in the scRNA-seq dataset by the expression of *elav*. Cells were further categorized into
905 either the ectodermal or endodermal nerve net based on the expression of the endodermal marker
906 *budhead*. Ectodermal neuron types were defined by distinct neuropeptide expression profiles,
907 and ectodermal neuron subtypes were classified based on spatially restricted domains (see main
908 text for detailed information on the classification scheme).
909
910
911
912
913
914

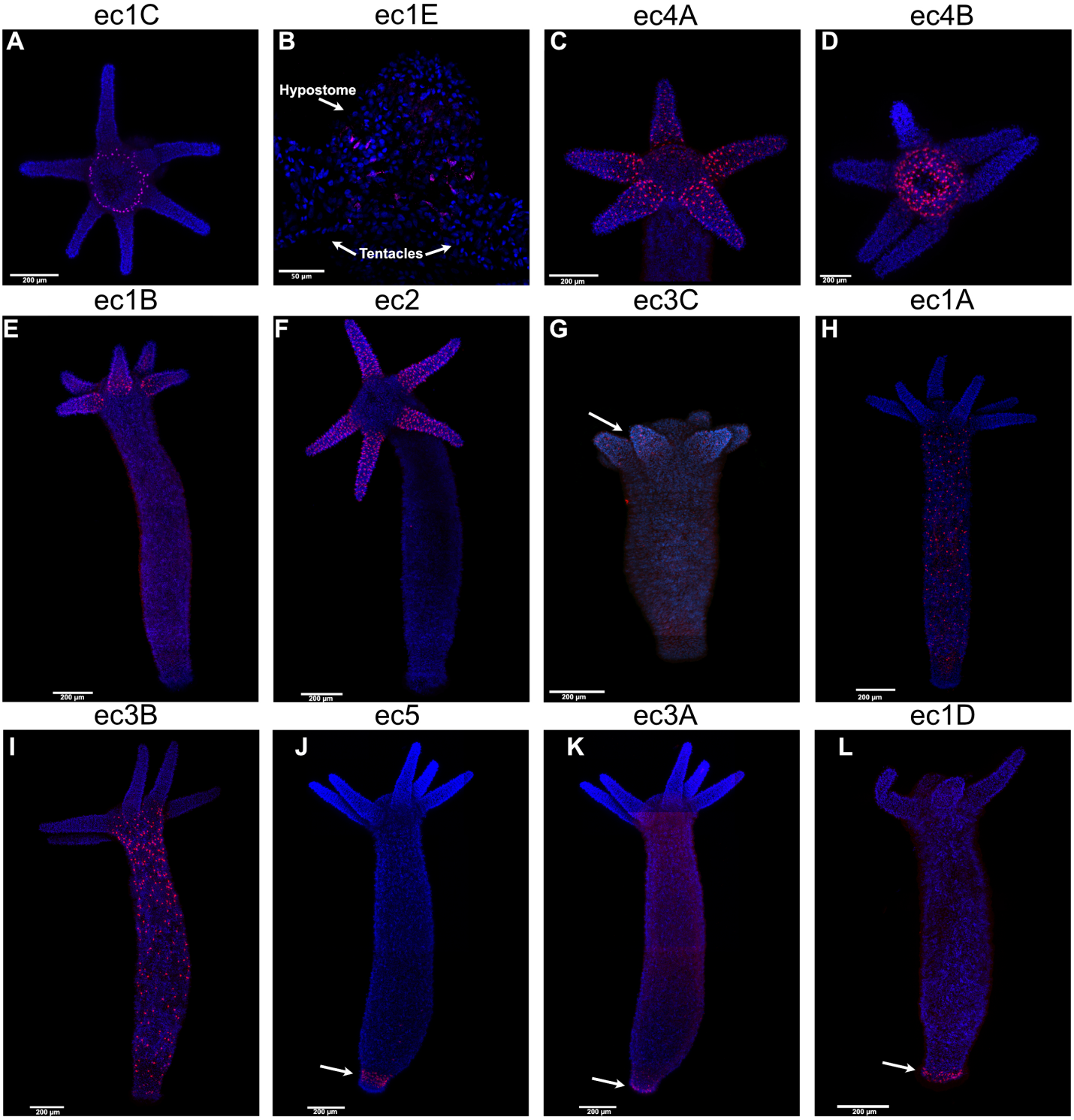

ec1C
A
200 µm
ec1E
B
Hypostome
Tentacles
50 µm
ec4A
C
200 µm
ec4B
D
200 µm
ec1B
E
200 µm
ec2
F
200 µm
ec3C
G
200 µm
ec1A
H
200 µm
ec3B
I
200 µm
ec5
J
200 µm
ec3A
K
200 µm
ec1D
L
200 µm

**Figure 2. Spatial organization of ectodermal neuron subtypes in revealed by fluorescence in situ hybridization (FISH).** Confocal images show FISH labeling of twelve distinct ectodermal neuron subtypes using subtype-specific molecular markers. **(A–G)** Oral neuron subtypes: ec1C (*G007791*), ec1E (*G003886*), ec4A (*RFamide C*, *G017226*), ec4B (*RFamide B*, *G017227*), ec1B (*G016169*), ec2 (*G010335*), and ec3C (*G026484*). Arrow in panel G indicates expression in the tentacles. **(H–I)** Body column neuron subtypes: ec1A (*G026993*) and ec3B (G004106). **(J–L)** Aboral neuron subtypes: ec5 (Hym176C, G016165), localized to the peduncle (arrow), and ec3A (G021930) and ec1D (G004200), localized to the basal disk (arrows). See Fig. S7 for whole animal images of ec4A, ec4B, and ec1C FISH experiments. Nuclei are counterstained with Hoechst (blue). All panels show standard FISH labeling except panel B, in which ec1E neurons required localization using hybridization chain reaction (HCR).

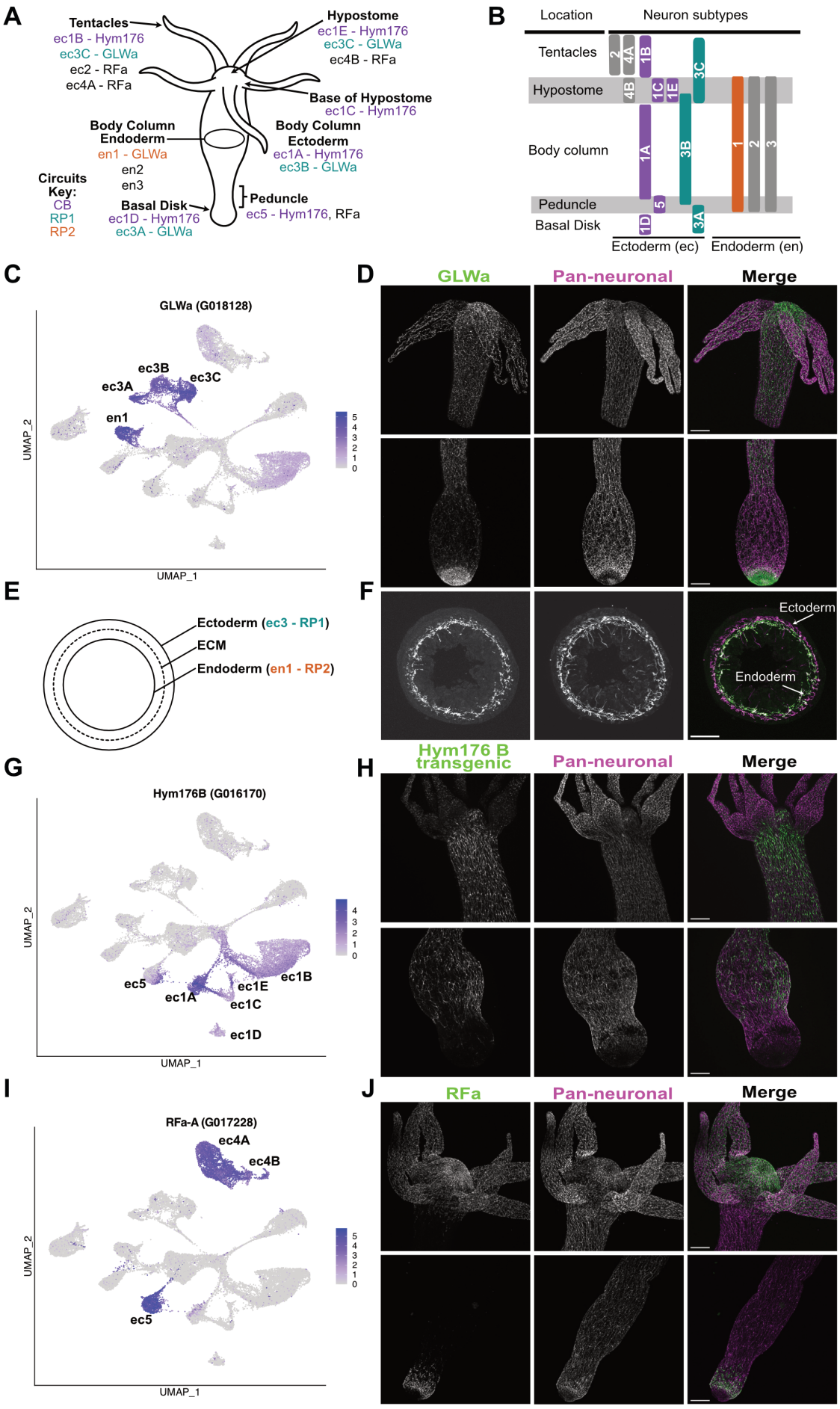
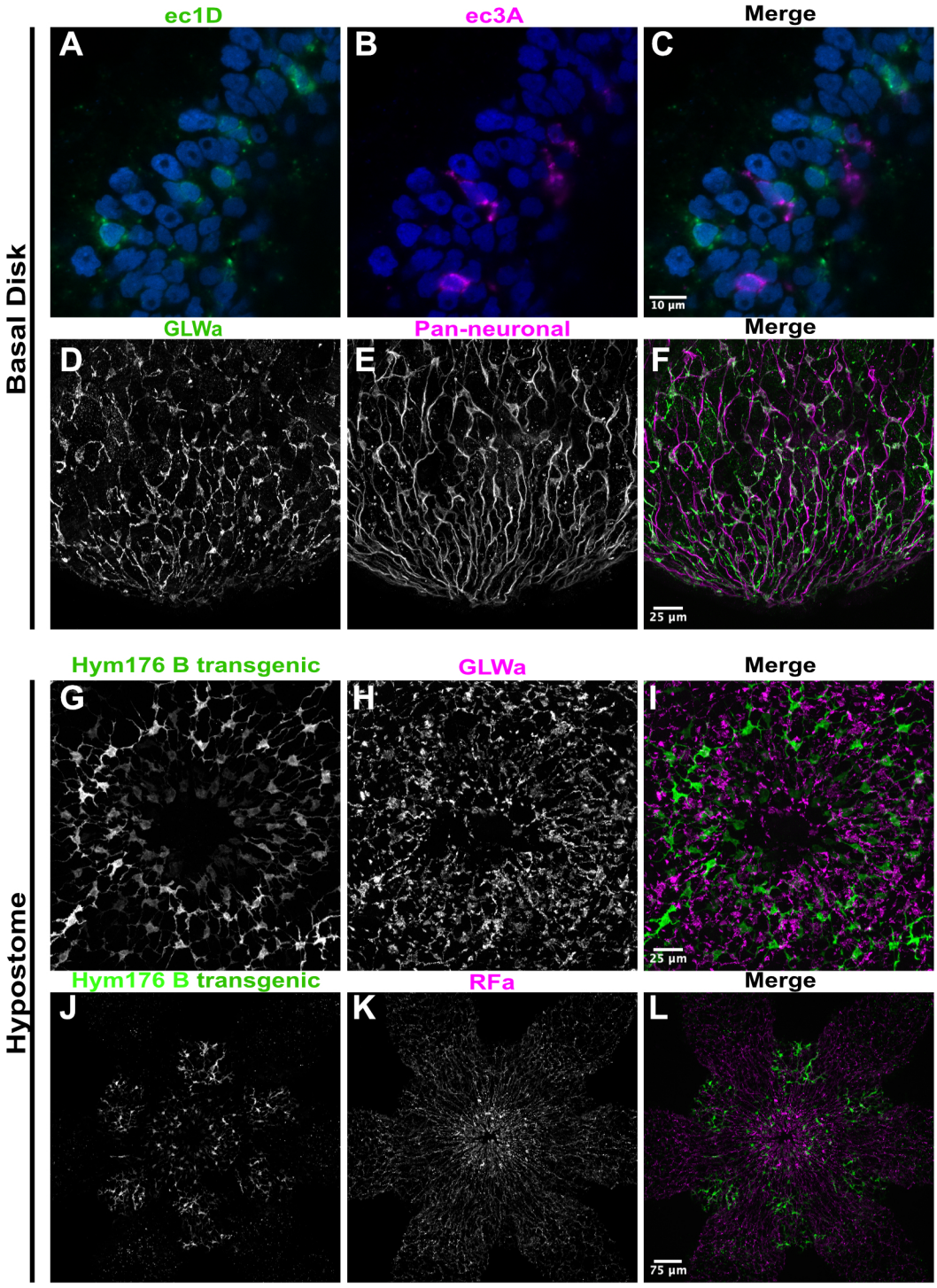
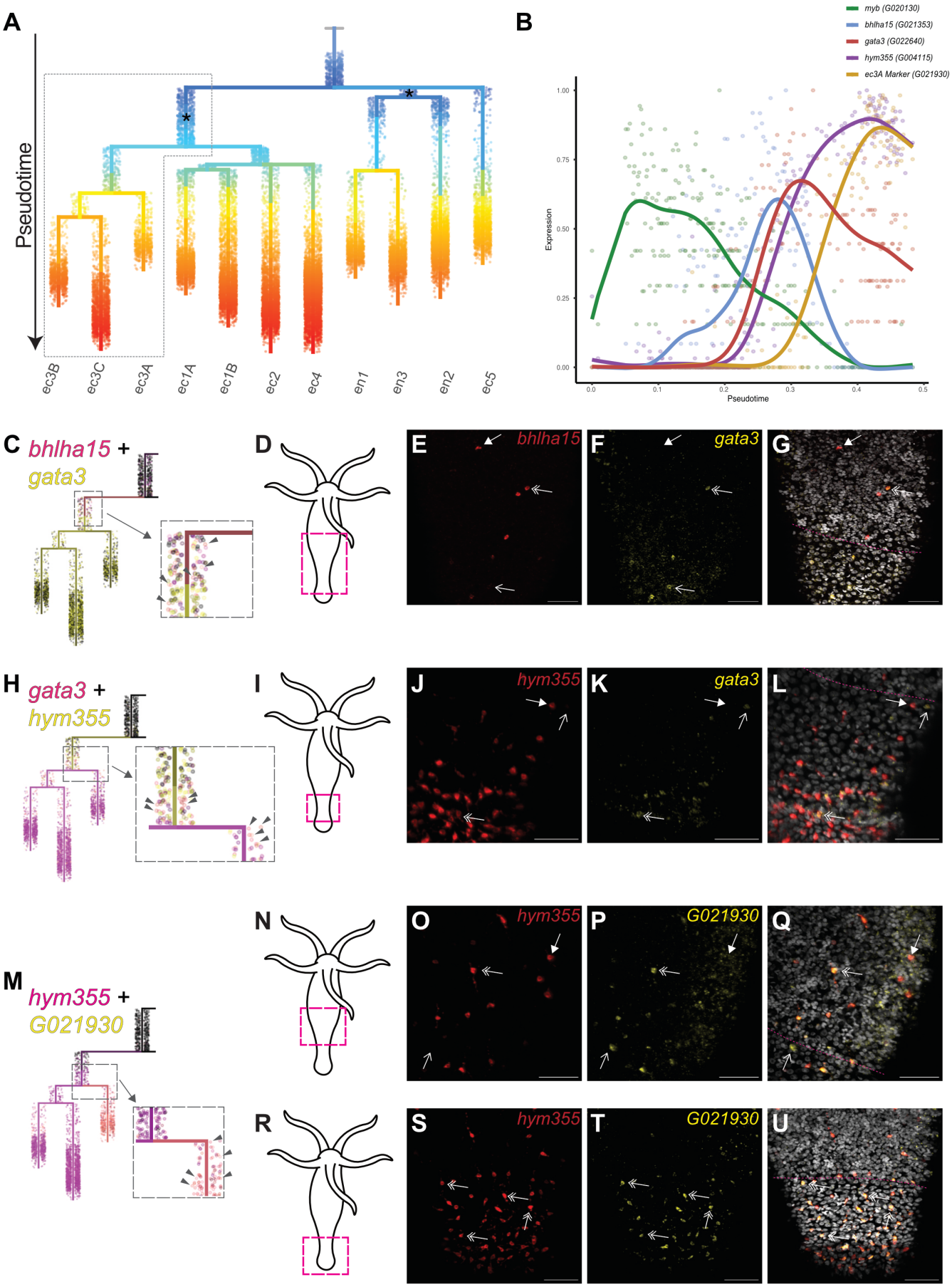

A
Tentacles
ec1B - Hym176
ec3C - GLWa
ec2 - RFa
ec4A - RFa
Hypostome
ec1E - Hym176
ec3C - GLWa
ec4B - RFa
Base of Hypostome
ec1C - Hym176
Body Column Endoderm
en1 - GLWa
en2
en3
Body Column Ectoderm
ec1A - Hym176
ec3B - GLWa
Circuits Key:
CB
RP1
RP2
Basal Disk
ec1D - Hym176
ec3A - GLWa
Peduncle
ec5 - Hym176, RFa
B
Location
Neuron subtypes
Tentacles
Hypostome
Body column
Peduncle
Basal Disk
Ectoderm (ec)
Endoderm (en)
C
GLWa (G018128)
ec3B
ec3C
ec3A
en1
UMAP_2
UMAP_1
D
GLWa
Pan-neuronal
Merge
E
Ectoderm (ec3 - RP1)
ECM
Endoderm (en1 - RP2)
F
Ectoderm
Endoderm
G
Hym176B (G016170)
ec5
ec1A
ec1E
ec1C
ec1B
ec1D
UMAP_2
UMAP_1
H
Hym176 B transgenic
Pan-neuronal
Merge
I
RFa-A (G017228)
ec4A
ec4B
ec5
UMAP_2
UMAP_1
J
RFa
Pan-neuronal
Merge

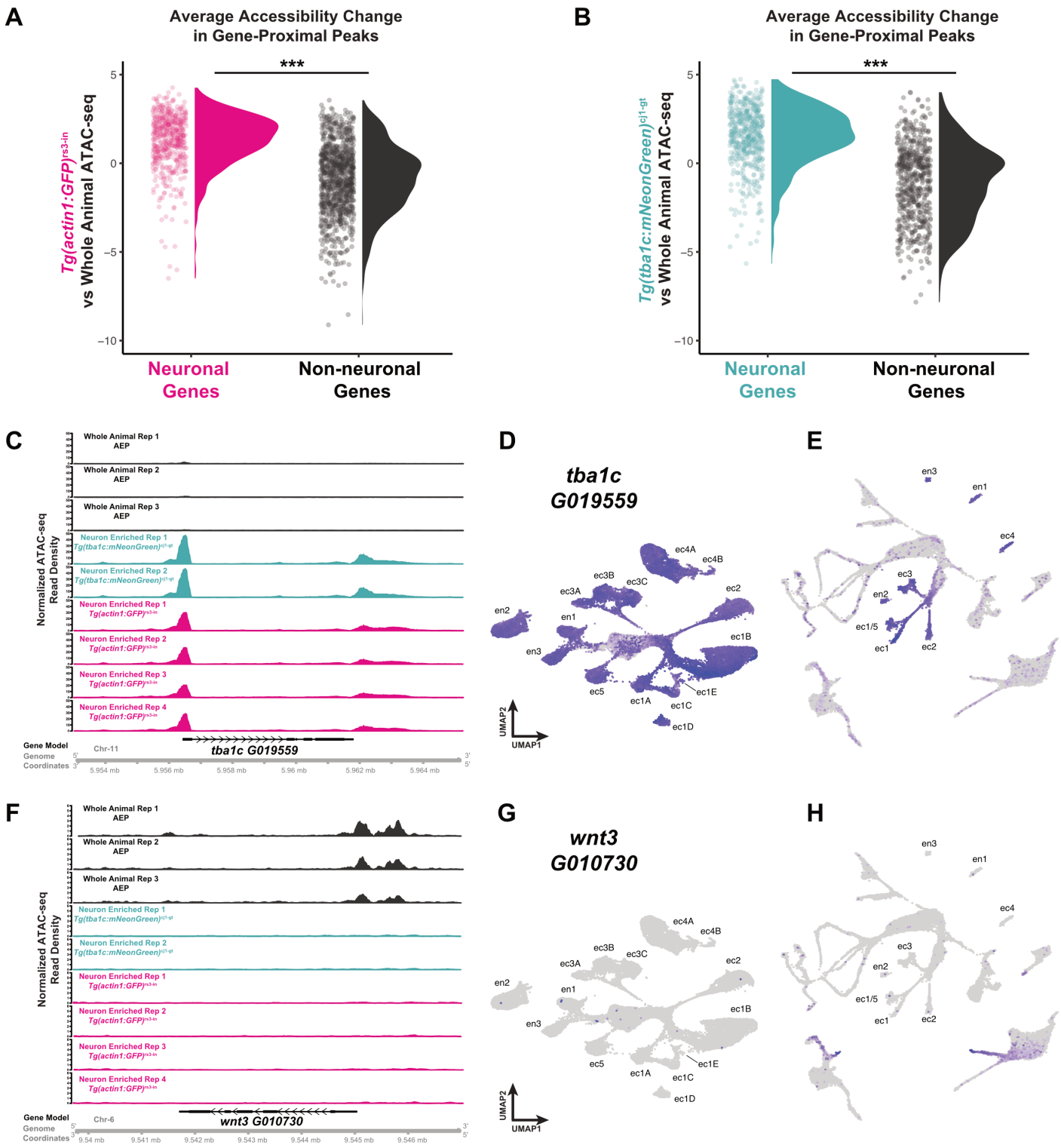
**Figure 3. Neuropeptide expression delineates distinct neural circuits in *Hydra vulgaris*. (A, B)** Schematics illustrate the spatial distribution of neuron subtypes along the *Hydra* body column. **(C)** The RP1 circuit comprises ec3A–C neurons, which express neuropeptides from the GLWa family. The RP2 circuit includes endodermal en1 neurons, which also express GLWa. **(D)** Immunostaining using an antibody against GLWa (green) reveals the RP1 circuit extending from the basal disk to the tentacles. **(E, F)** GLWa immunostaining in body column rings shows endodermal en1 neurons of the RP2 circuit. **(G)** The CB circuit includes ectodermal neuron subtypes ec1A–D and ec5, which express Hym176 neuropeptides. **(H)** GFP immunostaining (green) in the *Hym176B* reporter line marks CB circuit neurons along the body column. **(I)** The RFa-A neuropeptide is expressed in ec4 and ec5 neurons, the latter being part of the CB circuit. **(J)** Immunostaining with an antibody against RFa (green) reveals ec5 neurons in the peduncle and ec4 neurons near the oral end. In all immunostaining panels (D, F, H, J), co-labeling with PNab (magenta) highlights the full structure of the nervous system. Scale bars: 200 µm. The lower images in panel D and the upper images in panel H are reproduced from Keramidioti et al., 2024.


Basal Disk
ec1D
ec3A
Merge
A
B
C
10 µm
GLWa
Pan-neuronal
Merge
D
E
F
25 µm
Hypostome
Hym176 B transgenic
GLWa
Merge
G
H
I
25 µm
Hym176 B transgenic
RFa
Merge
J
K
L
75 µm

975 **Figure 4. Localization of neuron subtypes in the basal disk and hypostome.**
976 **(A–C)** Double FISH using probes for *G021930* (magenta, ec3A) and *G004200* (green, ec1D)
977 confirms that these neuron subtypes are distinct populations in the basal disk. Nuclei are
978 counterstained with Hoechst (blue) **(D–F)** Immunofluorescence with antibodies against GLWa
979 (green, ec3) and PNab (magenta, pan-neuronal) reveals GLWa-negative/PNab-positive neurons
980 in the basal disk, consistent with ec1D identity. **(G–I)** Co-immunostaining in the *Hym17B-GFP*
981 reporter line using anti-GFP (green) and anti-GLWa (magenta) shows the spatial relationship of
982 ec1 and ec3 neurons in the hypostome. **(J–L)** Co-immunostaining in the Hym17B-GFP reporter
983 line with anti-GFP (green) and anti-RFa (magenta) illustrates the positioning of ec1 and ec4B
984 neurons in the hypostome.
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005


A
Pseudotime
ec3B
ec3C
ec3A
ec1A
ec1B
ec2
ec4
en1
en3
en2
ec5
B
myb (G020130)
bhlha15 (G021353)
gata3 (G022640)
hym355 (G004115)
ec3A Marker (G021930)
Expression
Pseudotime
C bhlha15 +
gata3
D
E bhlha15
F gata3
G
H gata3 +
hym355
I
J hym355
K gata3
L
M hym355 +
G021930
N
O hym355
P G021930
Q
R
S hym355
T G021930
U

1006 **Figure 5. Trajectory reconstruction of neurogenesis reveals distinct developmental**
1007 **pathways for ectodermal and endodermal neurons. (A)** Differentiation trajectories for eleven
1008 *Hydra* neuron subtypes were reconstructed using URD from single-cell transcriptomics data
1009 (Farrell et al., 2018). Interstitial stem cells (ISCs) were set as the "root" or starting point, and
1010 each neuron subtype served as a "tip" or endpoint. The tree is colored by pseudotime
1011 (developmental progression), with earlier pseudotime at the top (blue) and later pseudotime at
1012 the bottom (red). The boxed region highlights areas detailed in panels C, H, and M. Asterisks
1013 indicate segments corresponding to early differentiation events from ISCs into ectodermal or
1014 endodermal neuronal progenitors. **(B)** Spline plot showing expression trends for five genes
1015 during ec3A differentiation. The x-axis represents pseudotime (earlier on the left, later on the
1016 right), and the y-axis represents scaled gene expression levels. Each point is the average
1017 expression for five cells. **(C-U)** Double FISH validates predicted transition states during ec3A
1018 differentiation. Expression states in FISH images are indicated by arrow types: closed arrows for
1019 cells expressing only the first gene, open arrows for cells expressing only the second, and double
1020 arrows for cells co-expressing both genes. **(C-G)** Validation of early ec3A transition states co-
1021 expressing *bhlha15* and *gata3*. **(C)** Visualization of *bhlha15* and *gata3* expression on the URD
1022 trajectory. Co-expressing cells (orange) are indicated with arrowheads; black represents cells
1023 expressing neither gene. **(D-G)** Confocal microscopy of double FISH for early transitions: **(D)**
1024 Body region imaged, **(E)** *bhlha15* expression (red), **(F)** *gata3* expression (yellow), and **(G)**
1025 Overlay with Hoechst-stained nuclei (gray). **(H-L)** Validation of mid ec3A transition states co-
1026 expressing *hym355* (magenta) and *gata3* (yellow). **(H)** Visualization of *hym355* and *gata3*
1027 expression on the URD trajectory. Co-expressing cells (orange) are highlighted with arrowheads.
1028 **(I-L)** Confocal microscopy of double FISH for mid transitions: **(I)** Body region imaged, **(J)**
1029 *hym355* expression (red), **(K)** *gata3* expression (yellow), and **(L)** Overlay with Hoechst-stained
1030 nuclei (gray). **(M-U)** Validation of late ec3A transition states co-expressing *hym355* (magenta)
1031 and ec3A marker *G021930* (yellow). **(M)** Visualization of *hym355* and *G021930* expression on
1032 the URD trajectory. Co-expressing cells (orange) are marked with arrowheads. **(N-U)** Confocal
1033 microscopy of double FISH for late transitions: **(N, R)** Body region imaged, **(O, S)** *hym355*
1034 expression (red), **(P, T)** *G021930* expression (yellow), and **(Q, U)** Overlay with Hoechst-stained
1035 nuclei (gray). Scale bars: 50 µm. Pink dotted lines in microscopy images (G, L, Q, U) indicate
1036 the boundary between the body column and peduncle, determined by nuclei morphology.


A
Average Accessibility Change in Gene-Proximal Peaks
***
Tg(actin1:GFP)rs3-in vs Whole Animal ATAC-seq
5
0
−5
−10
Neuronal Genes
Non-neuronal Genes
B
Average Accessibility Change in Gene-Proximal Peaks
***
Tg(tba1c:mNeonGreen)ci1-gt vs Whole Animal ATAC-seq
5
0
−5
−10
Neuronal Genes
Non-neuronal Genes
C
Normalized ATAC-seq Read Density
Whole Animal Rep 1 AEP
Whole Animal Rep 2 AEP
Whole Animal Rep 3 AEP
Neuron Enriched Rep 1 Tg(tba1c:mNeonGreen)ci1-gt
Neuron Enriched Rep 2 Tg(tba1c:mNeonGreen)ci1-gt
Neuron Enriched Rep 1 Tg(actin1:GFP)rs3-in
Neuron Enriched Rep 2 Tg(actin1:GFP)rs3-in
Neuron Enriched Rep 3 Tg(actin1:GFP)rs3-in
Neuron Enriched Rep 4 Tg(actin1:GFP)rs3-in
Gene Model
Genome Coordinates
Chr-11
tba1c G019559
5.954 mb
5.956 mb
5.958 mb
5.96 mb
5.962 mb
5.964 mb
D
tba1c G019559
ec4A
ec4B
ec3B
ec3C
ec3A
ec2
en2
en1
ec1B
en3
ec5
ec1A
ec1E
ec1C
ec1D
UMAP2
UMAP1
E
en3
en1
ec4
ec3
en2
ec1/5
ec1
ec2
F
Normalized ATAC-seq Read Density
Whole Animal Rep 1 AEP
Whole Animal Rep 2 AEP
Whole Animal Rep 3 AEP
Neuron Enriched Rep 1 Tg(tba1c:mNeonGreen)ci1-gt
Neuron Enriched Rep 2 Tg(tba1c:mNeonGreen)ci1-gt
Neuron Enriched Rep 1 Tg(actin1:GFP)rs3-in
Neuron Enriched Rep 2 Tg(actin1:GFP)rs3-in
Neuron Enriched Rep 3 Tg(actin1:GFP)rs3-in
Neuron Enriched Rep 4 Tg(actin1:GFP)rs3-in
Gene Model
Genome Coordinates
Chr-6
wnt3 G010730
9.54 mb
9.541 mb
9.542 mb
9.543 mb
9.544 mb
9.545 mb
9.546 mb
G
wnt3 G010730
ec4A
ec4B
ec3B
ec3C
ec3A
ec2
en2
en1
ec1B
en3
ec5
ec1A
ec1E
ec1C
ec1D
UMAP2
UMAP1
H
en3
en1
ec4
ec3
en2
ec1/5
ec1
ec2

1037 **Figure 6. Characterization of the neuronal chromatin landscape using ATAC-seq.**
1038 (**A-B**) Neuron-enriched ATAC-seq libraries were generated from six samples: two from the
1039 *Tg(tba1c:mNeonGreen)*$^{cj1\text{-}gt}$ transgenic line (green) and four from the *Tg(actin1:GFP)*$^{rs3\text{-}in}$
1040 transgenic line (pink) (Keramidioti et al., 2024). To assess neuronal peak enrichment, three
1041 previously published whole-animal ATAC-seq libraries (AEP1-3; Cazet et al., 2023) were used
1042 as a reference. Panels (**A**) and (**B**) show the average accessibility changes in gene-proximal
1043 peaks for *Tg(actin1:GFP)*$^{rs3\text{-}in}$ and *Tg(tba1c:mNeonGreen)*$^{cj1\text{-}gt}$ , respectively, compared to
1044 whole-animal data sets. Neuronal genes exhibit significant peak enrichment within 10,000 base
1045 pairs of the transcription start site relative to non-neuronal genes ($p < 0.001$, t-test). (**C**) ATAC-
1046 seq tracks for the *alpha tubulin* locus (*tba1c*, *G019559*) show peaks in the neuron-enriched
1047 libraries that are absent in the whole animal libraries. (**D**) UMAP visualization of the neuron-
1048 enriched scRNA-seq dataset shows that *tba1c* is ubiquitously expressed in neurons. (**E**) UMAP
1049 visualization of the whole-animal scRNA-seq dataset shows that *tba1c* expression is restricted to
1050 neurons (Cazet et al., 2023; Siebert et al., 2019). (**F**) ATAC-seq tracks for the *wnt3* locus
1051 (*G010730*) show peaks in the whole-animal libraries that are absent in the neuron-enriched
1052 libraries. (**G**) UMAP visualization of the neuron-enriched scRNA-seq dataset shows no
1053 expression of *wnt3* in neurons. (**H**) UMAP visualization of the whole animal scRNA-seq dataset
1054 shows that *wnt3* is expressed in the oral ectodermal and endodermal epithelium (Cazet et al.,
1055 2023; Siebert et al., 2019).
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067

A
ISC
ec2 Neuron
Expression
Pseudotime
Transcription Factor
Myb (G020130)
Myc3 (G003730)
Sox2 (G009896)
Noto (G003891)
Zic1/2/3/4/5 (G00445)
Arx (G021458)
Litaf (G007969)
G018876
Gsx1/2 (G023503)
ISC
ec4 Neuron
B
Sox2 (G009896)
C
Zic4 (G004456)
D
Arx (G021458)
E
Gata3 (G007969) Zic4(G023503) Gata3 + Zic4
F
Litaf (G007969) Gsx1/2 (G023503) Litaf + Gsx1/2
G
Gsx1/2 (G023503) Noto (G003891) Gsx1/2 + Noto
ec3B ec3C ec3A ec1A ec1B ec2 ec4 en1 en3 en2 ec5
H
Whole Animal Merge AEP
Neuron Enriched Merge Tg(tba1c:mNeonGreen)
Neuron Enriched Merge Tg(actin1:GFP)
scores
Predicted TFBS
Tcf7
Sox2
Myb
chr-5
Sox2 (G009896)
26342000 26343000 26344000 26345000 26346000
I
Whole Animal Merge AEP
Neuron Enriched Merge Tg(tba1c:mNeonGreen)
Neuron Enriched Merge Tg(actin1:GFP)
Predicted TFBS
Tcf7
Sox2
Myc3
Gsx2
chr-2
Noto (G003891)
38677000 38678000 38679000 38680000 38681000
J
Oral
Wnt
Bcat/Tcf
Sox2
Noto
Zic4
Arx
Gsx1/2
Myb
Myc3
Tent. Neuro Prog.
ec1B/2/4 Neuron
ec2/4 Neuron
ec4 Neuron
Neuro Prog.

1068 **Figure 7. A putative gene regulatory network (GRN) underlying ec4 neuron specification.**
1069 **(A)** Spline plots showing transcription factor expression dynamics during ec2 and ec4
1070 differentiation. **(B)** *Sox2* (*G009896*) is broadly expressed in oral neuron subtypes ec3C, ec1B,
1071 ec2, and ec4. **(C)** *Zic4* (*G004456*) is expressed in ec1B, ec2, and ec4. **(D)** *Arx* (*G021458*) is
1072 restricted to ec2 and ec4. **(E)** Co-expression trajectory tree of *Gata3* (*G022640*, magenta) and
1073 *Zic4* (*G004456*, green) show mutually exclusive expression. **(F)** Co-expression trajectory tree of
1074 *Litaf* (*G007969*, magenta) and *Gsx1/2* (*G023503*, green) reveals subtype-specific expression in
1075 ec2 and ec4 neurons. **(G)** Co-expression trajectory tree of *Gsx1/2* (magenta) and *Noto* (G003891,
1076 green) shows mutually exclusive expression in ec4. **(H)** ATAC-seq tracks at the *Sox2* locus
1077 reveal open chromatin regions containing *Tcf*, *Sox2*, and *Myb* binding motifs. (I) ATAC-seq
1078 tracks at the Noto locus highlight peaks containing *Tcf*, *Sox2*, *Myc*, and *Gsx2* binding motifs. (J)
1079 A proposed GRN model summarizing transcriptional regulators and regulatory relationships
1080 involved in ec4 neuron specification in *Hydra vulgaris*. Plots in this figure are made using the
1081 scRNA-seq portal (https://research.nhgri.nih.gov/HydraAEP/SingleCellBrowser/).
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097

1220 **References**


1221 **Badhiwala, K. N., Gonzales, D. L., Vercosa, D. G., Avants, B. W. and Robinson, J. T.**
1222 (2018). Microfluidics for electrophysiology, imaging, and behavioral analysis of Hydra.
1223 *Lab Chip* **18**, 2523–2539.

1224 **Badhiwala, K. N., Primack, A. S., Juliano, C. E. and Robinson, J. T.** (2021). Multiple
1225 neuronal networks coordinate Hydra mechanosensory behavior. *Elife* **10**, e64108.

1226 **Bode, H., Berking, S., David, C. N., Gierer, A., Schaller, H. and Trenkner, E.** (1973).
1227 Quantitative analysis of cell types during growth and morphogenesis in Hydra. *Wilhelm*
1228 *Roux Arch Entwickl Mech Org* **171**, 269–285.

1229 **Bode, H. R., Heimfeld, S., Koizumi, O., Littlefield, C. L. and Yaross, M. S.** (1988).
1230 Maintenance and Regeneration of the Nerve Net in Hydra. *American Zoologist* **28**, 1053–
1231 1063.

1232 **Bolger, A. M., Lohse, M. and Usadel, B.** (2014). Trimmomatic: a flexible trimmer for Illumina
1233 sequence data. *Bioinformatics* **30**, 2114–2120.

1234 **Buenrostro, J., Wu, B., Chang, H. and Greenleaf, W.** (2015). ATAC-seq: A Method for
1235 Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21.29.1-
1236 21.29.9.

1237 **Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R.** (2018). Integrating single-
1238 cell transcriptomic data across different conditions, technologies, and species. *Nat*
1239 *Biotechnol* **36**, 411–420.

1240 **Campbell, R. D.** (1967). Tissue dynamics of steady state growth in Hydra littoralis. II. Patterns
1241 of tissue movement. *J Morphol* **121**, 19–28.

1242 **Cazet, J. F., Siebert, S., Morris Little, H., Bertemes, P., Primack, A. S., Ladurner, P.,**
1243 **Achrainer, M., Fredriksen, M. T., Moreland, R. T., Singh, S., et al.** (2023). A
1244 chromosome-scale epigenetic map of the Hydra genome reveals conserved regulators of
1245 cell state. *Genome Res* gr.277040.122.

1246 **Chari, T., Weissbourd, B., Gehring, J., Ferraioli, A., Leclère, L., Herl, M., Gao, F.,**
1247 **Chevalier, S., Copley, R. R., Houliston, E., et al.** (2021). Whole-animal multiplexed
1248 single-cell RNA-seq reveals transcriptional shifts across Clytia medusa cell types.
1249 *Science Advances* **7**, eabh1683.

1250 **Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier,**
1251 **J., Cunha, A. and Pierce, N. A.** (2018). Third-generation in situ hybridization chain
1252 reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**,
1253 dev165753.

1254 **Cole, A. G., Steger, J., Hagauer, J., Denner, A., Ferrer Murguia, P., Knabl, P.,**
1255 **Narayanaswamy, S., Wick, B., Montenegro, J. D. and Technau, U.** (2024). Updated

single cell reference atlas for the starlet anemone Nematostella vectensis. *Frontiers in Zoology* **21**, 8.

**Corces, M. R., Trevino, A. E., Hamilton, E. G., Greenside, P. G., Sinnott-Armstrong, N. A., Vesuna, S., Satpathy, A. T., Rubin, A. J., Montine, K. S., Wu, B., et al.** (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**, 959–962.

**Das Gupta, R. R., Scheurer, L., Pelczar, P., Wildner, H. and Zeilhofer, H. U.** (2021). Neuron-specific spinal cord translatomes reveal a neuropeptide code for mouse dorsal horn excitatory neurons. *Sci Rep* **11**, 5232.

**David, C. N.** (2012). Interstitial stem cells in Hydra: multipotency and decision-making. *Int J Dev Biol* **56**, 489–497.

**David, C. N. and Gierer, A.** (1974). Cell Cycle Kinetics and Development of Hydra attenuata. III. Nerve and nematocyte differentiation.

**David, C. N. and Murphy, S.** (1977). Characterization of interstitial stem cells in hydra by cloning. *Dev Biol* **58**, 372–383.

**Dupre, C. and Yuste, R.** (2017). Non-overlapping Neural Networks in Hydra vulgaris. *Curr Biol* **27**, 1085–1097.

**Escobar, A., Kim, S., Primack, A. S., Duret, G., Juliano, C. E. and Robinson, J. T.** (2024). Terminal differentiation precedes functional circuit integration in the peduncle neurons in regenerating Hydra vulgaris. *Neural Dev* **19**, 18.

**Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A. and Schier, A. F.** (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131.

**Ferenc, J., Bonvin, M., Papasaikas, P., Ferralli, J., Nuninger, C. and Tsiairis, C. D.** (2024). A transcription factor toggle switch determines differentiated epidermal cell identities in Hydra. 2024.12.10.627691.

**Fincher, C. T., Wurtzel, O., de Hoog, T., Kravarik, K. M. and Reddien, P. W.** (2018). Cell type transcriptome atlas for the planarian Schmidtea mediterranea. *Science* **360**, eaaq1736.

**Giez, C., Pinkle, D., Giencke, Y., Wittlieb, J., Herbst, E., Spratte, T., Lachnit, T., Klimovich, A., Selhuber-Unkel, C. and Bosch, T. C. G.** (2023). Multiple neuronal populations control the eating behavior in Hydra and are responsive to microbial signals. *Curr Biol* **33**, 5288-5303.e6.

**Gonzales, D. L., Badhiwala, K. N., Vercosa, D. G., Avants, B. W., Liu, Z., Zhong, W. and Robinson, J. T.** (2017). Scalable electrophysiology in intact small animals with nanoscale suspended electrode arrays. *Nat Nanotechnol* **12**, 684–691.

1292 **Grimmelikhuijzen, C. J. P.** (1985). Antisera to the sequence Arg-Phe-amide visualize neuronal
1293 centralization in hydroid polyps. *Cell Tissue Res.* **241**, 171–182.

1294 **Hafemeister, C. and Satija, R.** (2019). Normalization and variance stabilization of single-cell
1295 RNA-seq data using regularized negative binomial regression. *Genome Biology* **20**, 296.

1296 **Hager, G. and David, C. N.** (1997). Pattern of differentiated nerve cells in hydra is determined
1297 by precursor migration. *Development* **124**, 569–576.

1298 **Hahne, F. and Ivanek, R.** (2016). Visualizing Genomic Data Using Gviz and Bioconductor.
1299 *Methods Mol Biol* **1418**, 335–351.

1300 **Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J.,**
1301 **Wilk, A. J., Darby, C., Zager, M., et al.** (2021). Integrated analysis of multimodal
1302 single-cell data. *Cell* **184**, 3573-3587.e29.

1303 **Hobmayer, E., Holstein, T. W. and David, C. N.** (1990). Tentacle morphogenesis in hydra: II.
1304 Formation of a complex between a sensory nerve cell and a battery cell. *Development*
1305 **109**, 897–904.

1306 **Hulett, R. E., Rivera-López, C., Gehrke, A. R., Gompers, A. and Srivastava, M.** (2024). A
1307 wound-induced differentiation trajectory for neurons. *Proc Natl Acad Sci U S A* **121**,
1308 e2322864121.

1309 **Juliano, C. E., Reich, A., Liu, N., Götzfried, J., Zhong, M., Uman, S., Reenan, R. A.,**
1310 **Wessel, G. M., Steele, R. E. and Lin, H.** (2014a). PIWI proteins and PIWI-interacting
1311 RNAs function in Hydra somatic stem cells. *Proc Natl Acad Sci U S A* **111**, 337–342.

1312 **Juliano, C. E., Lin, H. and Steele, R. E.** (2014b). Generation of Transgenic Hydra by Embryo
1313 Microinjection. *JoVE* 51888.

1314 **Keramidioti, A., Schneid, S., Busse, C., Cramer von Laue, C., Bertulat, B., Salvenmoser,**
1315 **W., Hess, M., Alexandrova, O., Glauber, K. M., Steele, R. E., et al.** (2024). A new
1316 look at the architecture and dynamics of the Hydra nerve net. *Elife* **12**, RP87330.

1317 **Kim, S., Badhiwala, K. N., Duret, G. and Robinson, J. T.** (2024). Phototaxis is a satiety-
1318 dependent behavioral sequence in Hydra vulgaris. *J Exp Biol* **227**, jeb247503.

1319 **Klimovich, A., Wittlieb, J. and Bosch, T. C. G.** (2019). Transgenesis in Hydra to characterize
1320 gene function and visualize cell behavior. *Nat Protoc* **14**, 2069–2090.

1321 **Klimovich, A., Giacomello, S., Björklund, Å., Faure, L., Kaucka, M., Giez, C., Murillo-**
1322 **Rincon, A. P., Matt, A.-S., Willoweit-Ohl, D., Crupi, G., et al.** (2020). Prototypical
1323 pacemaker neurons interact with the resident microbiota. *Proc. Natl. Acad. Sci. U.S.A.*
1324 **117**, 17854–17863.

1325 **Koizumi, O., Hamada, S., Minobe, S., Hamaguchi-Hamada, K., Kurumata-Shigeto, M.,**
1326 **Nakamura, M. and Namikawa, H.** (2015). The nerve ring in cnidarians: its presence
1327 and structure in hydrozoan medusae. *Zoology* **118**, 79–88.

1328 **Kondili, M., Fust, A., Preussner, J., Kuenne, C., Braun, T. and Looso, M.** (2017). UROPA: a
1329 tool for Universal RObust Peak Annotation. *Sci Rep* **7**, 2593.

1330 **Kotliar, D., Veres, A., Nagy, M. A., Tabrizi, S., Hodis, E., Melton, D. A. and Sabeti, P. C.**
1331 (2019). Identifying gene expression programs of cell-type identity and cellular activity
1332 with single-cell RNA-Seq. *eLife* **8**, e43803.

1333 **Kuehn, E., Clausen, D. S., Null, R. W., Metzger, B. M., Willis, A. D. and Özpolat, B. D.**
1334 (2022). Segment number threshold determines juvenile onset of germline cluster
1335 expansion in Platynereis dumerilii. *J Exp Zool B Mol Dev Evol* **338**, 225–240.

1336 **Landt, S. G., Marinov, G. K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S.,**
1337 **Bernstein, B. E., Bickel, P., Brown, J. B., Cayting, P., et al.** (2012). ChIP-seq
1338 guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22**,
1339 1813–1831.

1340 **Langfelder, P. and Horvath, S.** (2008). WGCNA: an R package for weighted correlation
1341 network analysis. *BMC Bioinformatics* **9**, 559.

1342 **Langmead, B. and Salzberg, S. L.** (2012). Fast gapped-read alignment with Bowtie 2. *Nat*
1343 *Methods* **9**, 357–359.

1344 **Leviev, I., Williamson, M. and Grimmelikhuijzen, C. J.** (1997). Molecular cloning of a
1345 preprohormone from Hydra magnipapillata containing multiple copies of Hydra-L
1346 Wamide (Leu-Trp-NH2) neuropeptides: evidence for processing at Ser and Asn residues.
1347 *J Neurochem* **68**, 1319–1325.

1348 **Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis,**
1349 **G. and Durbin, R.** (2009). The Sequence Alignment/Map format and SAMtools.
1350 *Bioinformatics* **25**, 2078–2079.

1351 **Li, Q., Brown, J. B., Huang, H. and Bickel, P. J.** (2011). Measuring reproducibility of high-
1352 throughput experiments. *Ann. Appl. Stat.* **5**,.

1353 **Lommel, M., Tursch, A., Rustarazo-Calvo, L., Trageser, B. and Holstein, T. W.** (2017).
1354 Genetic knockdown and knockout approaches in Hydra. 230300.

1355 **Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I.,**
1356 **Bialas, A. R., Kamitaki, N., Martersteck, E. M., et al.** (2015). Highly Parallel Genome-
1357 wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–
1358 1214.

1359 **McInnes, L., Healy, J., Saul, N. and Großberger, L.** (2018). UMAP: Uniform Manifold
1360 Approximation and Projection. *Journal of Open Source Software* **3**, 861.

1361 **Noro, Y., Yum, S., Nishimiya-Fujisawa, C., Busse, C., Shimizu, H., Mineta, K., Zhang, X.,**
1362 **Holstein, T. W., David, C. N., Gojobori, T., et al.** (2019). Regionalized nervous system
1363 in Hydra and the mechanism of its development. *Gene Expr Patterns* **31**, 42–59.

1364 **Noro, Y., Shimizu, H., Mineta, K. and Gojobori, T.** (2021). A single neuron subset governs a
1365 single coactive neuron circuit in Hydra vulgaris, representing a possible ancestral feature
1366 of neural evolution. *Sci Rep* **11**, 10828.

1367 **Rentzsch, F., Juliano, C. and Galliot, B.** (2019). Modern genomic tools reveal the structural
1368 and cellular diversity of cnidarian nervous systems. *Curr Opin Neurobiol* **56**, 87–96.

1369 **Robinson, M. D., McCarthy, D. J. and Smyth, G. K.** (2010). edgeR: a Bioconductor package
1370 for differential expression analysis of digital gene expression data. *Bioinformatics* **26**,
1371 139–140.

1372 **Salomone, J., Qin, S., Fufa, T. D., Cain, B., Farrow, E., Guan, B., Hufnagel, R. B.,**
1373 **Nakafuku, M., Lim, H.-W., Campbell, K., et al.** (2021). Conserved Gsx2/Ind
1374 homeodomain monomer versus homodimer DNA binding defines regulatory outcomes in
1375 flies and mice. *Genes Dev.* **35**, 157–174.

1376 **Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. and Regev, A.** (2015). Spatial
1377 reconstruction of single-cell gene expression data. *Nat Biotechnol* **33**, 495–502.

1378 **Sebé-Pedrós, A., Saudemont, B., Chomsky, E., Plessier, F., Mailhé, M.-P., Renno, J., Loe-**
1379 **Mie, Y., Lifshitz, A., Mukamel, Z., Schmutz, S., et al.** (2018a). Cnidarian Cell Type
1380 Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. *Cell* **173**,
1381 1520-1534.e20.

1382 **Sebé-Pedrós, A., Chomsky, E., Pang, K., Lara-Astiaso, D., Gaiti, F., Mukamel, Z., Amit, I.,**
1383 **Hejnol, A., Degnan, B. M. and Tanay, A.** (2018b). Early metazoan cell type diversity
1384 and the evolution of multicellular gene regulation. *Nat Ecol Evol* **2**, 1176–1188.

1385 **Siebert, S., Anton-Erxleben, F. and Bosch, T. C. G.** (2008). Cell type complexity in the basal
1386 metazoan *Hydra* is maintained by both stem cell based mechanisms and
1387 transdifferentiation. *Developmental Biology* **313**, 13–24.

1388 **Siebert, S., Farrell, J. A., Cazet, J. F., Abeykoon, Y., Primack, A. S., Schnitzler, C. E. and**
1389 **Juliano, C. E.** (2019). Stem cell differentiation trajectories in Hydra resolved at single-
1390 cell resolution. *Science* **365**, eaav9314.

1391 **Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y.,**
1392 **Stoeckius, M., Smibert, P. and Satija, R.** (2019). Comprehensive Integration of Single-
1393 Cell Data. *Cell* **177**, 1888-1902.e21.

1394 **Takaku, Y., Hwang, J. S., Wolf, A., Böttger, A., Shimizu, H., David, C. N. and Gojobori, T.**
1395 (2014). Innexin gap junctions in nerve cells coordinate spontaneous contractile behavior
1396 in Hydra polyps. *Sci Rep* **4**, 3573.

**Taylor, S. R., Santpere, G., Weinreb, A., Barrett, A., Reilly, M. B., Xu, C., Varol, E., Oikonomou, P., Glenwinkel, L., McWhirter, R., et al.** (2021). Molecular topography of an entire nervous system. *Cell* **184**, 4329-4347.e23.

**Trembley, A., Trembley, A., Pronk, C., Schley, J. van der and Lyonet, P.** (1744). *Mémoires pour servir à l'histoire d'un genre de polypes d'eau douce, à bras en forme de cornes*. A Leide: Chez Jean & Herman Verbeek.

**Tzouanas, C. N., Kim, S., Badhiwala, K. N., Avants, B. W. and Robinson, J. T.** (2021). Hydra vulgaris shows stable responses to thermal stimulation despite large changes in the number of neurons. *iScience* **24**, 102490.

**Vogg, M. C., Beccari, L., Iglesias Ollé, L., Rampon, C., Vriz, S., Perruchoud, C., Wenger, Y. and Galliot, B.** (2019a). An evolutionarily-conserved Wnt3/β-catenin/Sp5 feedback loop restricts head organizer activity in Hydra. *Nat Commun* **10**, 312.

**Vogg, M. C., Galliot, B. and Tsiairis, C. D.** (2019b). Model systems for regeneration: Hydra. *Development* **146**, dev177212.

**Vogg, M. C., Ferenc, J., Buzgariu, W. C., Perruchoud, C., Sanchez, P. G. L., Beccari, L., Nuninger, C., Le Cras, Y., Delucinge-Vivier, C., Papasaikas, P., et al.** (2022). The transcription factor Zic4 promotes tentacle formation and prevents epithelial transdifferentiation in Hydra. *Science Advances* **8**, eabo0694.

**Voigt, O., Erpenbeck, D. and Wörheide, G.** (2008). A fragmented metazoan organellar genome: the two mitochondrial chromosomes of Hydra magnipapillata. *BMC Genomics* **9**, 350.

**Westfall, J. A. and Epp, L. G.** (1985). Scanning electron microscopy of neurons isolated from the pedal disk and body column of *Hydra*. *Tissue and Cell* **17**, 161–170.

**Wittlieb, J., Khalturin, K., Lohmann, J. U., Anton-Erxleben, F. and Bosch, T. C. G.** (2006). Transgenic Hydra allow in vivo tracking of individual stem cells during morphogenesis. *Proc Natl Acad Sci U S A* **103**, 6208–6211.

**Yamamoto, W. and Yuste, R.** (2023). Peptide-driven control of somersaulting in Hydra vulgaris. *Curr Biol* **33**, 1893-1905.e4.

**Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., et al.** (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**, R137.

**Zhong, W., Barde, S., Mitsios, N., Adori, C., Oksvold, P., Feilitzen, K. von, O'Leary, L., Csiba, L., Hortobágyi, T., Szocsics, P., et al.** (2022). The neuropeptide landscape of human prefrontal cortex. *Proceedings of the National Academy of Sciences* **119**, e2123146119.