



HAL
open science

Comparative single-cell transcriptomics of complete insect nervous systems

Benjamin T Cocanougher, Jason D Wittenbach, Xi Salina Long, Andrea B Kohn, Tigran P Norekian, Jinyao Yan, Jennifer Colonell, Jean-Baptiste Masson, James W Truman, Albert Cardona, et al.

► **To cite this version:**

Benjamin T Cocanougher, Jason D Wittenbach, Xi Salina Long, Andrea B Kohn, Tigran P Norekian, et al.. Comparative single-cell transcriptomics of complete insect nervous systems. 2023. pasteur-04009853

HAL Id: pasteur-04009853

<https://pasteur.hal.science/pasteur-04009853>

Preprint submitted on 1 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 Comparative single-cell transcriptomics of complete insect nervous 2 systems

3
4 Benjamin T. Cocanougher^{1,2,8}, Jason D. Wittenbach^{1,8}, Xi Salina Long^{1,8}, Andrea B.
5 Kohn³, Tigran P. Norekian^{3,4}, Jinyao Yan¹, Jennifer Colonell¹, Jean-Baptiste Masson⁶,
6 James W. Truman¹, Albert Cardona^{1,5}, Srinivas C. Turaga¹, Robert H. Singer^{1,7,9}, Leonid
7 L. Moroz^{3,5,9*}, Marta Zlatic^{1,2,9,10*}

8 Affiliations:

9 ¹Howard Hughes Medical Institute Janelia Research Campus, Ashburn, Virginia, USA.

10 ²Department of Zoology, Cambridge University, Cambridge, UK.

11 ³Department of Neuroscience and Whitney Laboratory for Marine Biosciences,
12 University of Florida, Gainesville/St. Augustine, Florida 32080, USA.

13 ⁴Institute of Higher Nervous Activity and Neurophysiology, Russian Academy of
14 Sciences, Moscow, Russia.

15 ⁵Department of Physiology, Development, and Neuroscience, Cambridge University,
16 Cambridge, UK.

17 ⁶Decision and Bayesian Computation, Institut Pasteur, CNRS USR 3756, Department of
18 Computational Biology and Neuroscience, Paris, France.

19 ⁷Department of Anatomy and Structural Biology, Albert Einstein College of Medicine,
20 Bronx, New York, USA.

21 ⁸These authors contributed equally

22 ⁹These authors contributed equally

23 ¹⁰Lead Contact

24
25 * Corresponding author. Email: moroz@whitney.ufl.edu (L.L.M.);
26 zlaticm@janelia.hhmi.org (M.Z.)

35 **Summary:**

36 Molecular profiles of neurons influence information processing, but bridging the gap
37 between genes, circuits, and behavior has been very difficult. Furthermore, the
38 behavioral state of an animal continuously changes across development and as a result
39 of sensory experience. How behavioral state influences molecular cell state is poorly
40 understood. Here we present a complete atlas of the *Drosophila* larval central nervous
41 system composed of over 200,000 single cells across four developmental stages. We
42 develop *polyseq*, a python package, to perform cell-type analyses. We use single-
43 molecule RNA-FISH to validate our scRNAseq findings. To investigate how internal
44 state affects cell state, we optogenetically altered internal state with high-throughput
45 behavior protocols designed to mimic wasp sting and over activation of the memory
46 system. We found nervous system-wide and neuron-specific gene expression changes.
47 This resource is valuable for developmental biology and neuroscience, and it advances
48 our understanding of how genes, neurons, and circuits generate behavior.

49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

66 **Introduction:**

67 Making sense of any complex system involves identifying constituent elements and
68 understanding their individual functions and interactions. Neural circuits are no
69 exception. While recent advances in connectomics (White et al., 1986; Jarrell et al.,
70 2012, Helmstaedter et al., 2013; Takemura et al., 2013; Ohyama et al., 2015; Berck et
71 al., 2016; Eichler et al., 2017; Hildebrand et al., 2017; Eschbach et al., 2019) and live
72 imaging techniques (Ahrens et al., 2013; Prevedel et al., 2014; Chhetri et al., 2015;
73 Lemon et al., 2015; Grimm et al., 2017; Vladimirov et al., 2018) offer unprecedented
74 information about neural connectivity and activity, the task of identifying cell types has
75 traditionally relied on painstaking morphological, functional, or single gene
76 histochemical taxonomy. High-throughput single-cell RNA sequencing (scRNAseq)
77 offers a new way forward by providing a molecular-level identity for each cell via its
78 transcriptomic profile. Importantly, it is also scalable to populations of millions of cells
79 without incurring exorbitant costs. These techniques have already revealed striking
80 heterogeneity in cell populations that is lost in bulk samples. In the fruit fly, efforts are
81 already well underway to produce connectomic (Takemura et al., 2013; Ohyama et al.,
82 2015; Berck et al., 2016; Eichler et al., 2017; Eschbach et al., 2019), activity (Chhetri et
83 al., 2015; Lemon et al., 2015; Grimm et al., 2017; Vladimirov et al., 2018), and behavior
84 atlases (Vogelstein et al., 2014; Robie et al., 2017) of the nervous system. Much work
85 has separately revealed the role that genes (Konopka and Benzer, 1971; Sokolowski
86 2001) and circuits (Garcia-Campmay et al., 2010, Borst 2014) play in behavior; a major
87 challenge is to combine genes, circuits, and behavior all at once. Single-cell analyses
88 have been performed in parts of adult (Croset et al., 2018; Davie et al., 2018;
89 Konstantinides et al., 2018) *Drosophila* central brain and optic lobe. One study has
90 investigated a small sample of the larval central brain (Alvalos et al., 2019). A
91 comprehensive transcriptomic atlas of the complete central nervous system is the
92 missing piece to the connectivity, activity, and behavior maps that would create the
93 required resource necessary to understand the complex interplay between genes,
94 circuits, and behavior.

95

96 To this end, we developed a protocol to capture, sequence, and transcriptionally classify
97 the molecular cell types and cell states of the entire central nervous system of the
98 *Drosophila* larva. We did this across 4 different life stages, providing a developmental
99 profile of gene expression. Given that the *Drosophila* larva has a nervous system of
100 approximately 10,000-15,000 neurons (Hartenstein and Campos-Ortega, 1984;
101 Hartenstein et al., 1987; Truman et al., 1993; Scott et al., 2001), our atlas of 202,107
102 cells has up to 20X coverage of the entire nervous system and is the largest sequencing
103 effort in *Drosophila* to date. All previously identified cell types were recognizable in our
104 atlas, including motor neurons, Kenyon cells of the mushroom body, insulin-producing
105 cells, brain dopaminergic and serotonergic cells, and all glial subtypes.

106

107 While scRNAseq provides nearly complete information about the transcriptional
108 program being used by a cell at the time of collection, a drawback to the technique is a
109 loss of spatial information. We therefore used a recently developed RNA fluorescent *in*
110 *situ* hybridization (RNA-FISH) protocol to resolve the anatomical location of molecular
111 cell types in the whole larval brain (Long et al., 2017). We combined RNA-FISH with
112 high-resolution Bessel beam structured illumination microscopy to detect and count
113 individual mRNAs within newly identified cells. This technique provides ground truth for
114 the absolute number of a particular RNA in a given molecular cell type at a particular
115 time point. It also provides an opportunity to assess the quantitative capability of our
116 scRNAseq approach.

117

118 Larval behavior after hatching is dominated by feeding; when a critical weight is
119 achieved, this behavior switches to “wandering” in preparation for pupation (Bakker et
120 al., 1953). Endocrine and neuroendocrine pathways responsible for this switch have
121 been well characterized (Truman, 2005), but the extent of molecular changes in defined
122 cell types across the nervous system that respond to this neuroendocrine signaling are
123 not known. To investigate such nervous system-wide changes during development, we
124 sequenced the nervous system at four time points in development.

125

126 Previous sensory experience alters the behavioral state of an animal. Flies that are
127 hungry form food-associated memories more easily (Krashes et al., 2009) and flies that
128 are intoxicated court more frequently (Lee et al., 2008). Male flies that lose fights are
129 more likely to subsequently exhibit submissive behaviors and to lose second contests
130 while male flies that win exhibit aggressive behavior and are more likely to win later
131 fights (Trannoy et al., 2017). Are internal states controlled transcriptionally at the level of
132 identified cell types and circuits and, if so, how? It is an open question whether memory
133 or internal state will affect gene expression globally or only in restricted cell populations.

134

135 In order to discover the nervous system-wide gene expression changes induced by
136 previous experience, we examined gene expression profiles from the nervous systems
137 of animals exposed to two experimental protocols. The first protocol involved presenting
138 repeated pain and fear, by mimicking repeated wasp sting. Fictive stings were induced
139 using optogenetic activation of a small population of well-described interneurons
140 (Ohyama et al., 2015). Of note, no mechanical damage to the animal's surface occurred
141 with this protocol. The second protocol involved repeated activation of higher-order
142 central brain neurons involved in learning. Using behavioral assays before and after the
143 stimulation, we showed that each of these protocols cause a long-lasting change in the
144 animals' behavioral state. We then analyzed the effect of fictive sting and repeated
145 activation of the learning center to search for changes in gene expression related to cell
146 state during behavioral learning. We consider these "cell state" genes and find that both
147 entire cell populations and individual neuron types can exhibit cell state changes.

148

149 Taken together, these results suggest the powerful role that transcriptomic atlases can
150 play in probing the complex interplay between cell state, circuit function, and behavior.

151

152 **Results:**

153 ***Polyseq* software performs cell type discovery**

154 A complete transcriptomic atlas of 202,107 single cells from the larval central nervous
155 system was built (Figure 1; Table S1). Nervous systems were captured at four time
156 points in development (1 hour, 24 hours, 48 hours, and 96 hours after larval hatching)

157 and for three nervous system dissections (full CNS, brain only, and ventral nerve cord
158 only). These developmental timepoints and anatomical regions were analyzed
159 separately and in combination. Four non-neuronal and non-glial tissues, including ring
160 gland cells, hemocytes, imaginal disc cells, and salivary gland cells, were also captured
161 and analyzed as outgroups.

162

163 We developed *polyseq* (github.com/jwittenbach/polyseq), an open source Python
164 package, to perform cell type analyses. *polyseq* performs functions of many popular R
165 packages such as *Seurat* or *Monocle* (Trapnell et al., 2014; Satija et al., 2015; Qui et
166 al., 2017; Cao et al., 2019) with significantly improved runtime and the extensibility and
167 modularity of Python. *polyseq* starts with a gene by cell matrix, which is then filtered for
168 high quality cells, normalized, regressed, reduced, and clustered. Visualization can then
169 be performed with tSNE or umap for the full dataset (Maaten and Hinton, 2008; McInnis
170 et al., 2018). The software also includes inbuilt functionality for violin plots and
171 heatmaps. The data remain in a form that is easy to integrate with the vast community
172 of Python packages for further visualization and analysis (full details and analysis
173 examples on github: github.com/jwittenbach/polyseq).

174

175 We first used *polyseq* to discover cell type clusters and confirmed that our findings were
176 in agreement with current state of the art analysis methods (Figure 2). In two separate
177 early third instar samples, we found the same cell types when analyzing the data in
178 *Seurat*, *Monocle*, and *polyseq* (Figure 2A). In these samples, the cells separated into
179 seven groups of developing neurons (which included subtypes of adult developing
180 neurons, neuroblasts, and ganglion mother cells), four groups of glia, immune cells, and
181 three groups of larval functional neurons (including distinct motor neuron and Kenyon
182 cell groupings).

183

184 To correct for batch effects, both the align function used in the monocle R package
185 (Haghverdi et al., 2018) and our own linear regression method in *polyseq* were tested
186 (Figure 2B,C). Both methods removed the visible batch effects in the umap plots (i.e.,
187 clusters that were made entirely of a single sample due to signal from separate batches

188 collapsed into a single co-mingled population). As additional validation of cell type
189 discovery, we used Garnett, a newly developed machine learning software package in
190 R, to build a classifier based on cell type markers (Pliner et al., 2019). We found
191 consistent results with our own annotation of known and newly discovered markers for
192 larval functional neurons, neural stem cells, motor neurons, kenyon cells, and glia
193 (Figure 2D,E). Using more specific markers for known cells which are small in overall
194 number (such as insulin-producing cells, dopaminergic cells, octopaminergic cells, etc.)
195 led to overfitting of the data. These known gene markers could be used to extract cells
196 of interest without unsupervised methods.

197

198 **Developmental profile of gene expression across four life stages**

199 Given that the analysis in *polyseq* met the standards of current state of the art methods,
200 we moved forward with an analysis of developmental timepoints. We built atlases of the
201 entire nervous system at 1 hour, 24 hours, 48 hours, and 96 hours. Within these
202 atlases, it was clear that during development, the cellular composition of the nervous
203 system changes (Figure S1). At 1 hour, the nervous system is primarily larval functional
204 neurons. As development proceeds, the absolute number of larval functional neurons
205 remains relatively constant while the proportion of developing neurons greatly expands.

206

207 Having identified the main classes of cells, we investigated developmental trajectories
208 of 12,448 neural progenitor cells (NPCs) (Figure 3). We extracted and combined NPCs
209 cells from three stages of development (1 hour, 24 hours, and 48 hours) and performed
210 an analysis in Monocle (Trapnell et al., 2014; Qiu et al., 2017; Cao et al., 2019) (Figure
211 3). Garnett was used to predict cell types (Pliner et al., 2019). Known cell ages were
212 used to anchor a pseudotime analysis, which aligned the data from early to late NPCs.
213 Gene expression in these populations revealed known markers (such as *insensible*
214 (*insb*) in Ganglion Mother Cells) and unexpected markers, including long non-coding
215 RNA (*CR31386* in early NBs). IGF-II mRNA-binding protein (*Imp*) and Syncrin (*Syp*)
216 form important gradients that mark NB age (Liu et al., 2015). *Imp* levels decrease with
217 age while *Syp* increases with age – young NBs have high levels of *Imp* and low levels of
218 *Syp*, intermediate NBs have intermediate levels of *Imp* and *Syp*, and older NBs have

219 low *Imp* and high *Syp*. These waves are evident in our data and provide an opportunity
220 to investigate further temporal gene expression gradients.

221
222 Gene modules were discovered, which characterized populations of early, intermediate,
223 and late NPCs (Figure 3D,E; Table S2). Early NPCs were characterized by the
224 expression of genes important for genome organization, chromatin remodeling, and
225 gene splicing. This allows for a future diversity of cell function and identity. Intermediate
226 NPC gene modules were characterized by genes necessary to build neurons – these
227 included expression of genes important for protein targeting and transport and
228 neurotransmitter synthesis. Late NPCs were enriched for genes which are critical for
229 newly differentiated neurons and circuit construction; significant GO terms included
230 genes required for connecting circuits, such as axon guidance molecules and synapse
231 organization genes, and genes important for circuit function, such as genes involved in
232 memory storage.

233

234 **Complete transcriptomic atlas of the larval central nervous system**

235 Next we built an atlas of all cells captured at all stages (Figure 4). Transcriptomic cell
236 types split into seventy clusters (Table S1). These seventy cell types could be grouped
237 into many recognizable groups of cells, including: (1) adult developing brain neurons,
238 (2) adult developing VNC neurons, (3) larval functional brain neurons, (4) larval
239 functional VNC neurons, (5) motor neurons, (6) kenyon cells, (7) brain neuroblasts, (8)
240 VNC neuroblasts, (9) brain ganglion mother cells, (10) VNC ganglion mother cells, (11)
241 glia, (12) hemocytes, (13) imaginal disc cells, (14) salivary gland cells, and (15) ring
242 gland cells.

243

244 Larvae spend much of their life feeding and growing. From initial hatching to pupation,
245 larvae grow significantly in length and mass (Truman et al., 2005). During this growth
246 period, the larval nervous system grows and adds developing adult neurons which
247 remain quiescent during larval life but grow and elaborate their axonal and dendritic
248 arbors during pupation into adult functional neurons (Li et al., 2014). In the atlas, we can
249 identify adult developing neurons through high expression neuronal markers (*nSyb*,

250 *elav*) and a lack of synaptic and neurotransmitter genes (e.g., *VChAT*, *VGlut*). Recent
251 work in the first instar brain showed that adult developing neurons (or undifferentiated
252 neurons) express *headcase* (*hdc*) and *unkempt* (*unk*) (Avalos et al., 2019). We see this
253 expression continues in adult developing neurons at 24 and 48 hours. Furthermore, we
254 find this group is marked by many more genes, including the actin-binding protein
255 *singed* (*sn*), the zinc finger transcription factor *jim*, and the transcriptional repressor
256 *pleiohomeotic* (*pho*).

257

258 Larval functional neurons participate in neural circuits which control sensation and
259 behavior. At larval hatching, embryonic neurons are born, and all the neurons
260 necessary for larval life are functional (Truman and Bate, 1988). These neurons will
261 continue to grow and some populations, such as the Kenyon cells of the mushroom
262 body, will add more neurons throughout development. Identifiable cells at the top level
263 include motor neurons, Kenyon cells, excitatory and inhibitory interneurons,
264 monoaminergic neurons, and neuropeptidergic neurons. A unifying feature of these
265 cells includes expression of classical *Drosophila* neuronal markers (*nSyb*, *elav*),
266 however, we also find many other genes that mark the larval functional neuron group
267 robustly. These markers include the transmembrane receptor protein tyrosine kinase
268 activator *jelly belly* (*jeb*), the protein tyrosine phosphatase *IA-2*, the ligand gated
269 chloride channel *Resistant to dieldrin* (*Rdl*), and one of the beta subunits of sodium-
270 potassium pump (*nirvana3*; *nrv3*).

271

272 Monoaminergic neurons play a key role in learning in the fly (Schwaerzel et al., 2003;
273 Selcho et al., 2009). A single top-level cluster was identified with the expression of key
274 monoaminergic synthetic enzymes (*Trh*, *ple*) and transporters (*DAT*, *SerT*).

275 Subclustering of this top-level cluster revealed three strong groups, corresponding to
276 serotonergic, dopaminergic, and octopaminergic clusters, identifying previously
277 undescribed markers of these populations of cells which separate one monoamine type
278 from another (Figure S7).

279

280 *Jhl-21*, a solute carrier 7-family amino acid transporter, is an example of a novel marker
281 found here in 5-HT neurons. This gene encodes for a protein necessary for protein
282 nutrition signaling, which was recently described (Ro et al. 2016; Ziegler et al., 2016;
283 Ziegler et al., 2018). However, these reports describe the importance of the Jhl-21
284 protein peripherally, with no mechanism for the transmitting of nutritional information to
285 the nervous system. Here we see that *Jhl-21* is expressed in the serotonin neuron itself,
286 suggesting that serotonin neurons act directly as sensors for the amino acid nutritional
287 state.

288
289 Neural progenitor cells include neuroblasts (NBs), intermediate neural progenitors, and
290 ganglion mother cells (GMCs) (Doe, 2017). NBs divide asymmetrically in three ways to
291 produce progeny: type 0 NBs divide into one self-renewing NB and one neuron; type 1
292 NBs divide into one NB and one GMC; type 2 NBs divide into one neuroblast and one
293 intermediate neural progenitor which then itself divides into a GMC (Doe, 2017). Each
294 GMC then divides terminally to form two neurons or one neuron and one glial cell.
295 Precisely timed patterns of temporal transcription factors guide this development. We
296 are able to investigate these patterns over space and time by collecting the brain and
297 VNC separately and collecting multiple stages of larval development (Figure 3). The
298 mushroom body continues to grow and develop during larval life. We were able to
299 identify mushroom body neuroblasts in our dataset, which were found in brain NB
300 clusters and characterized by high expression of the late neuroblast marker *Syp*, genes
301 for cell cycling, including *pendulin* (*Pen*) and cyclin E (*CycE*), and by the long noncoding
302 RNA *pncr002:3R* (Figure 4).

303
304 Five glial subtypes were recognizable in our atlas, including midline/cortex, astrocyte-
305 like, chiasm, peripheral/surface, and longitudinal body glia (Figure 4) (Freeman, 2015).
306 These glia were identified based on the expression of well-characterized markers, such
307 as *wrapper* and *slit* (*sli*) expression in midline/cortex glia, *alrm* expression in astrocyte-
308 like glia, *hoe1* expression in chiasm glia, *swim* in surface glia, and *vir-1* in longitudinal
309 body glia. In addition, we find *CG5955*, which codes for an L-threonine 3-

310 dehydrogenase, is highly expressed and found specifically in all glia other than
311 longitudinal body glia.

312

313 Hemocytes, imaginal disc, salivary gland, and ring gland cells were also captured and
314 sequenced. Hemocytes form the immune system in *Drosophila*. Hemocytes expressed
315 serpent (*srp*), the canonical marker of embryonic hemocytes (Fossett and Schulz,
316 2001). Hemocytes also had a very high expression of neuropeptide-like precursor 2
317 (*Nplp2*).

318

319 Imaginal discs are embryonic tissues that become adult tissues, such as wings and
320 legs, after metamorphosis. We dissected these cells and sequenced them separately.
321 We found high and specific expression of many uncharacterized genes, including
322 *CG43679*, *CG14850*, *CG44956*, and *CG31698*, among others (Table S1).

323

324 Unlike the imaginal disc, which had few genes in common with neurons, salivary gland
325 cells, surprisingly, formed a homogenous group characterized by expression of many
326 genes shared with neurons, such as the nucleo-cytoplasmic shuttling protein
327 *hephaestus* (*heph*), the RNA-binding protein *Syncrip* (*Syp*), the cadherin molecule
328 *Shotgun* (*shg*), and the cell adhesion molecule *Fasciclin 3* (*Fas3*). Given the secretory
329 nature of the salivary gland, it would be interesting to further investigate the evolutionary
330 and developmental relationship between the salivary gland and neurons, especially
331 given that in other animals, such as molluscs, salivary gland cells are secretory and
332 have action potentials (Kater et al., 1978a,b).

333

334 The ring gland is critical for transitions in development. The ring gland was
335 characterized by expression of the well-described Halloween genes, including members
336 of the cytochrome P450 family required for ecdysteroid biosynthesis, including *phantom*
337 (*phm*), *spook* (*spo*), *spookier* (*spok*), *disembodied* (*dib*), *shadow* (*sad*) and *shade* (*shd*)
338 (Gilbert, 2004). The ring gland also has a high expression of the NADP/NADPH
339 phosphatase *curled* (*cu*).

340

341 **Validating transcriptomic predictions**

342 Here, we used Insulin-producing cells (IPCs) as illustrative examples to validate scRNA-
343 seq data. IPCs consist of just fourteen neurons in the larval brain (Figure 5A) (Schlegel
344 et al., 2016). These cells participate in circuits which monitor the nutritional status of the
345 larva and function as the larval equivalent of the mammalian pancreas. If IPCs are
346 ablated, larvae and adults are smaller and have a diabetic phenotype, including
347 increased hemolymph trehalose and glucose levels (Rulifson et al., 2002). IPCs secrete
348 insulin-like peptides which regulate hemolymph sugar levels. Graph-based clustering
349 revealed a cluster defined by the strong expression of insulin-like peptide 2 (*Ilp2*) and
350 insulin-like peptide 5 (*Ilp5*) expression, which are canonical markers of IPCs (Figure
351 5B). By subsetting the data to look only at 96 cells in the putative IPCs cluster, the
352 neurotransmitters and receptors expressed by these could be analyzed (Figure 5C).

353

354 Previous reports show IPCs are regulated by canonical neurotransmitters. This includes
355 modulation by serotonin through the 5-HT1A receptor and octopamine through the
356 Octbeta1 receptor (Luo et al., 2012) and by the neuropeptide allatostatin A (Hentze et
357 al., 2015). We confirmed this known expression of 5-HT1A and Octbeta1 receptor. In
358 our atlas, we also see the strong expression of additional (previously unknown for these
359 cells) receptors for dopamine (*Dop2R*), glutamate (*GluClalpha*), and Allatostatin C
360 Receptor 2(*AstC-R2*) in IPCs (Figure 5C).

361

362 To validate the specificity of our scRNAseq approach for identifying AstC-R2 in ICP
363 cells, we probed AstC-R2 mRNA in a HaloTag reporter line for the ICPs. The overlap
364 between the neurons containing the HaloTag and FISH signals confirmed the
365 sequencing result (Figure 5D). The colocalization of AstC-R2 with 14 IPCs suggests
366 that all ICPs are regulated by AstC through AstC-R2. The discovery of regulation by
367 AstC-R2 updates our model of the regulation of IPCs by adding an additional population
368 of cells that are modulating IPC activity.

369

370 **Correlating smRNA-FISH and scRNAseq**

371 In order to determine whether scRNAseq could quantitatively capture the dynamics of
372 expression in a single cell , we compared scRNAseq expression to ground truth
373 expression levels determined using single-molecule RNA-FISH (smRNA-FISH) (Femino
374 et al., 1998). To make analyses more precise and localized, we compared the RNA
375 levels in a very small population of cells discovered in our atlas that express vesicular
376 glutamate transporter (VGlut) and the neuropeptide Allatostatin C (AstC). We quantified
377 the relative expression of these mRNAs using smFISH and compared the result with
378 scRNAseq (Figure 6).

379

380 We used smFISH to probe VGlut and AstC mRNAs simultaneously and obtain
381 quantitative expression levels. We detected 5 groups of cells that contain AstC FISH
382 signals, which was consistent with previous reports of AstC localization (Williamson et
383 al., 2001). We observed 5 pairs of cells that contained both VGlut and AstC. Among the
384 5 pairs, 1 pair belonged to previously reported SLP1 AstC cells (Figure 6A). We
385 quantified the VGlut and AstC mRNAs in these cells using a Bessel beam selective
386 plane illumination microscope (BB-SIM) (Long et al., 2017). Although quantification of
387 VGlut and AstC mRNAs within individual cell bodies could not be obtained due to the
388 difficulty of segmenting overlapping cell bodies, we were able to obtain an average
389 quantification of VGlut and AstC mRNAs within these 5 pair of cells (Figure 6B). The
390 similarity we obtained for the VGlut and AstC expression ratio between single-molecule
391 FISH and scRNAseq suggested that the relative quantification from scRNAseq was
392 compatible with single-molecule FISH (Figure 6C).

393

394 **Optogenetic sting alters expression globally**

395 To investigate if a change in internal state would alter nervous-system-wide gene
396 expression, we examined gene expression profiles from animals exposed to repeated
397 fictive sting. An optogenetic sting was induced by activation of the basin interneurons.
398 The basins are first order interneurons that receive input from nociceptive (pain) and
399 mechanosensory (vibration) sensory neurons (Ohyama et al., 2015). Such optogenetic
400 activation of the brain evokes a rolling escape response (Ohyama et al., 2015), which

401 mimics the natural response to wasp sting or nociceptor activation (Hwang et al., 2007).

402 Of note, no mechanical damage was induced in our protocol.

403

404 The basin interneurons were activated for 15 seconds, with a 45 second rest period for
405 a total of 120 activation periods (Figure 7A). Supervised machine learning was used to

406 automatically detect behavior (Jovanic et al., 2017). A rolling escape response was

407 observed at the start of the experiment; by the final activation stimulus, backing up and

408 turning were the predominate responses (Figure 7B).

409

410 Control animals of the same age were collected from the same food plate as

411 experimental animals and placed on an agar plate in the dark for two hours.

412 Immediately following the sting protocol, 2-4 animals from each group underwent the

413 scRNAseq protocol. We were primarily interested in searching for cell state genes which

414 could drive cell state clustering (Figure 7D). If cells from experimental animals and

415 controls are analyzed together, will cells cluster together based on cell type

416 (independent of treatment group) or cell state (dependent on treatment group)? If cell

417 type clustering is observed, it suggests that any changes in cell state induced by our

418 protocol are minor compared to cell type-specific features. But if cell state clustering is

419 observed, it is evidence that experience driven changes are at least of comparable

420 importance to cell type in determining genetic cell state.

421

422 The optogenetic sting protocol led to cell state clustering. Transcriptomic data from cells

423 isolated from activated and control brains were normalized and analyzed in the same

424 mathematical space, but the clustering that was observed was based on cell state (i.e.,

425 clustering was driven by whether the cells came from a “stung” animal or an “unstung”

426 control). Cell state genes that differed between the stung and unstung controls were

427 discovered (Figure 7F).

428

429 Cell state genes were most evident in larval functional and developing neurons,

430 including motor neurons, cholinergic, and neuropeptidergic cells. Genes that were

431 upregulated in motor neurons following the sting protocol included non-coding RNA

432 (*CR40469*), carbohydrate metabolic enzymes (lactate dehydrogenase, *ImpL3*), and the
433 ethanol-induced apoptosis effector, *Drat*.

434

435 In addition to cell state genes within the nervous system, a large group of immune cells
436 was sequenced in the sting state, with particularly high expression of neuropeptide-like
437 precursor 2 (*NPLP2*), a gene which has been observed in phagocytic immune cells
438 (Fontana et al., 2012). RNA-FISH in sting and control conditions revealed this transcript
439 is not higher in the neurons but suggested that the lymph gland is being activated and
440 ramping up cell numbers in the sting condition (Figure S8). Strong evolutionary
441 selection pressure exists on the larva to survive predation by parasitic wasps
442 (Kraaijeveld and Godfray, 1997). Larvae can survive by using their immune system to
443 encapsulate and prevent the hatching of internalized parasitic wasp eggs. Previous
444 work has investigated the role of signaling detected by mechanical damage of the sting
445 and the presence of the foreign body (wasp egg) inside the larva (Sorrentino et al.,
446 2002). Our transcriptomic data suggest the immune system may also respond to a
447 currently unknown signal generated directly by the nervous system.

448

449 **Learning center overactivation alters expression locally**

450 In a second behavior protocol, we activated all higher-order central brain neurons
451 involved in learning and memory, called Kenyon cells (KCs). Similar to the fictive sting,
452 we observed a change in behavioral response at the start and end of the training. At the
453 start of training, animals hunch and arrest movement at the onset of activation and
454 crawl forward at the offset of activation (Figure 7C). At the end of the training, animals
455 continue to hunch and stop at the onset, but a larger fraction (~80%) perform a small
456 motion before turning rather than crawling forward to offset. Also, this protocol not only
457 altered animals response to the optogenetic activation of KCs but also drastically
458 altered behavior after activation. Animals greatly increased the probability of stopping
459 and reduced the probability of crawling.

460

461 To discover potential molecular changes that could drive these behavioral changes, we
462 analyzed the transcriptomes of animals exposed to these optogenetic training protocols

463 and compared them to controls. Unlike the global changes in gene expression following
464 an optogenetic sting, we detected changes in the transcriptomic state of many fewer cell
465 types following repeated activation of the higher-order brain neurons involved in
466 learning and memory (Figure 7G-I). We discovered a number of interesting candidate
467 genes that were upregulated in an activity-dependent way in Kenyon cells and
468 dopaminergic neurons, which are key cell populations in the learning and memory
469 center (Figure 7; Table S3, S4).

470

471 Cell state genes with differential expression between KC activated brains and controls
472 separated local groups of cells within clusters. Changes were observed in KCs and
473 dopaminergic neurons (DANs) (Figure 7H, I). Cell state genes were not limited to
474 previously described activity-related genes. They included long non-coding RNA (*noe*,
475 *CR40469*), chromatin remodeling (*His2AV*, *mamo*), axon guidance (*trn*, *DIP1*, *fas*,
476 *dpr14*, *fz2*), and receptor genes (*Dop2R*).

477

478 **Discussion:**

479 This work makes several contributions to the field. First, we present the first full
480 transcriptomic atlas of the entire central nervous system at the single-cell resolution.
481 Second, we use super-resolution microscopy to compare single-molecule RNA-FISH
482 with scRNAseq in the *Drosophila* larva. By combining these two techniques – the first
483 providing information about the complete collection of RNA present in a cell and the
484 second providing full anatomical, subcellular, and absolute quantification of a chosen
485 RNA(s) – we provide a resource for the field of *Drosophila* neurobiology and provide an
486 example of complementary methods for building and validating single-cell molecular
487 atlases. Third, we provide an experimental paradigm for discovering a molecular
488 signature of internal state and use this paradigm to uncover drastic gene-expression
489 changes that accompany a state of stress evoked by repeated “optogenetic” predator
490 attack. Our atlas is therefore a powerful resource for developmental biology,
491 neuroscience, and evolutionary biology.

492

493 Separating cell type from cell state is a key challenge for transcriptomic cell atlases. In
494 order to understand the changes in a specific cell type between health and disease, for
495 example, it will be necessary to be able to find the same cell type among differing
496 conditions. Here we show that such a distinction can be discovered in cases where
497 molecular cell state is significantly altered. Furthermore, we show that even in a
498 nervous system with drastically different (and unnatural) activity patterns in a sizable
499 population of highly interconnected neurons (here 200 of 10,000 or 2% of the nervous
500 system), limited changes may be observed across the entire nervous system but
501 changes can be observed in specific cell types. We foresee such techniques being
502 useful to investigate a wide range of internal state and cell state changes, from sleep to
503 parasitism to circadian rhythms.

504

505 Single-cell transcriptomic atlases are the missing piece required for the combined
506 analysis of genes, circuits, and behavior. Our work here shows that transcriptomic
507 atlases can be reliably built for multiple developmental stages of the *Drosophila* larva.
508 Furthermore, we show that optogenetic manipulations of internal state can alter gene
509 expression in a context-dependent manner. By adding a transcriptomic atlas to the
510 existing atlases of neuron connectivity, neuron activity, and behavior, we have set the
511 stage for a more complete understanding of the principles that underlie the complex
512 interplay of genes, circuits, and behavior.

513

514

515 **Acknowledgments:**

516 We thank J. Grimm for sharing reagents; J. Etheredge for fly stocks; S. Harrison, M.
517 Mercer, and the Janelia Fly Core for assistance with fly husbandry; A. Lemire and K.
518 Aswath of Janelia Quantitative Genomics for assistance with sequencing. **Funding:**
519 Supported by Janelia HHMI (M.Z.), Gates Cambridge Trust (B.T.C.), HHMI Medical
520 Fellows Program (B.T.C.), NSF (1146575, 1557923, 1548121 and 1645219; L.L.M.).

521

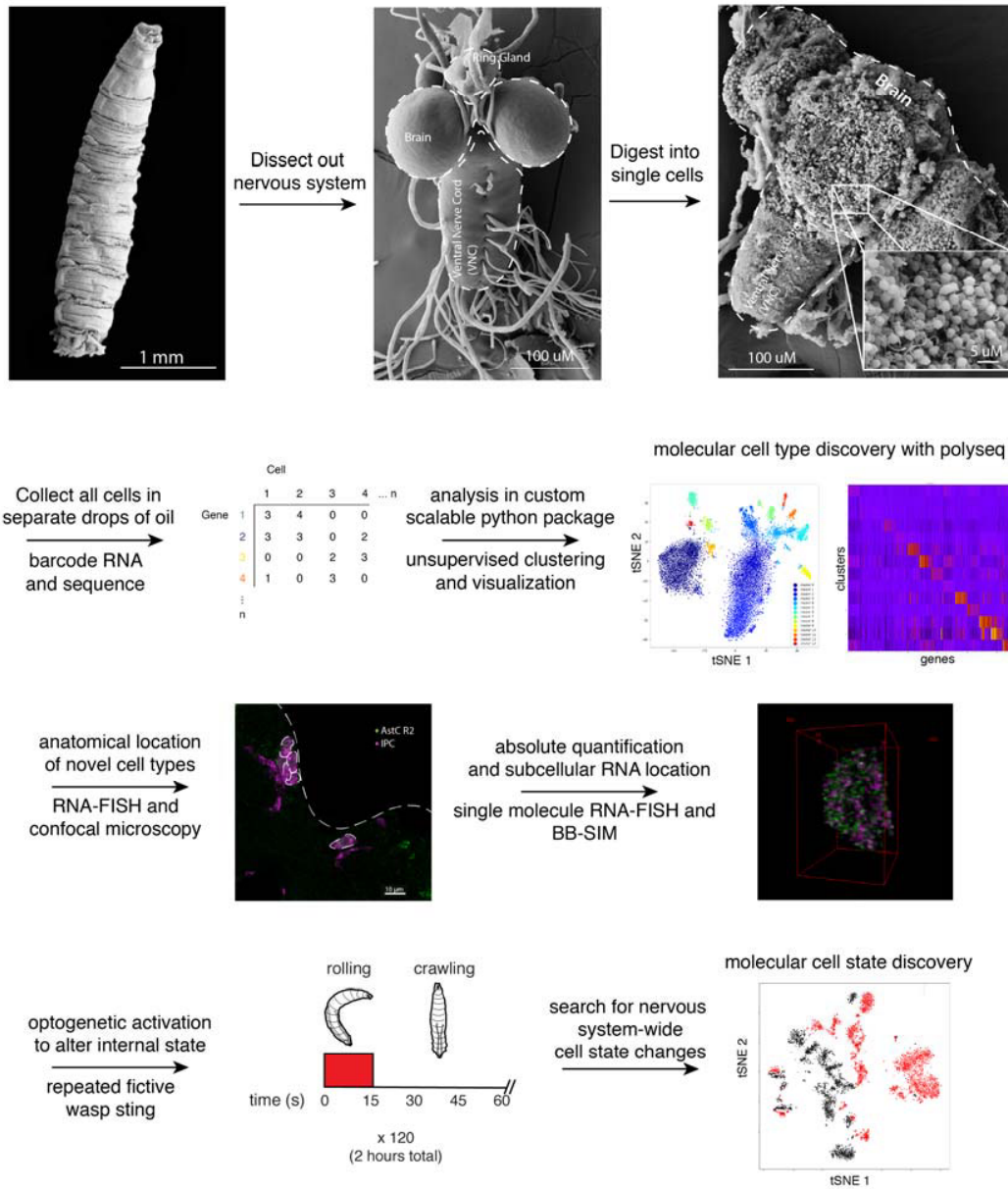
522 **Author contributions:** B.T.C., L.L.M., and M.Z. conceptualized the study; B.T.C.,
523 J.D.W., X.L., and J.Y. performed all data analysis and visualization; B.T.C., J.D.W., J.-
524 B.M., and J.Y. performed all software development, B.T.C., X.L., J.L., A.B.K., and
525 L.L.M. performed the investigation; B.T.C., J.D.W., and X.L. curated the data; B.T.C.,
526 X.L., A.B.K., and L.L.M. developed methodology and performed validation, R.H.S.,
527 L.L.M. and M.Z. acquired funding, provided resources, and performed project
528 administration; S.T., R.H.S., L.L.M. and M.Z. supervised the study; B.T.C., J.D.W. and
529 X.L. wrote the original draft; B.T.C., X.L., J.D.W., R.H.S., L.L.M. and M.Z. edited the
530 final manuscript; all authors reviewed the final manuscript.

531

532 **Declaration of Interests:** none.

533

534 **Figure titles and legends:**



535

536

537 **Figure 1. Schematic of full nervous system scRNA-seq collection and analysis.** In

538 order to develop a central nervous system-wide transcriptomic atlas of the *Drosophila*

539 nervous system with single-cell resolution, we developed a protocol to digest the entire

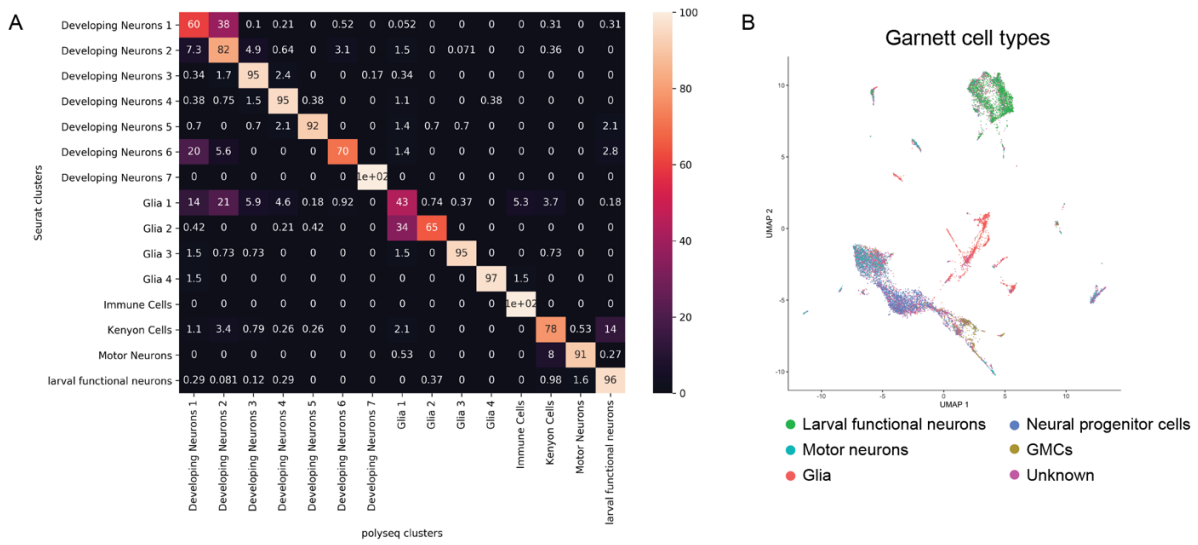
540 nervous system into single cells, collected the cells using a microfluidic device (10x

541 Chromium machine, 10x Genomics, Pleasanton, CA), and sequenced the mRNA from

542 each cell. After barcoding and sequencing, a cell by gene matrix is generated. This cell

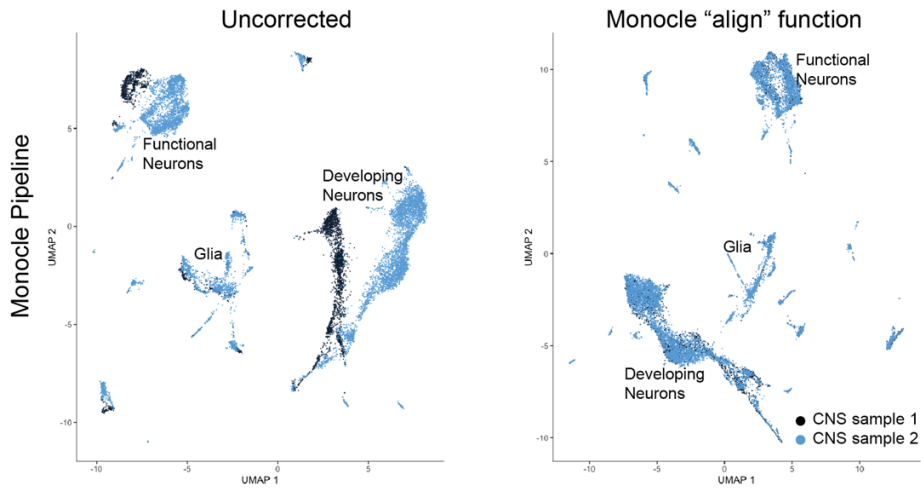
543 by gene matrix was then analyzed with *polyseq*, a custom python package. In total,
544 202,107 neurons and glia were sequenced. The anatomical location of newly defined
545 molecular cell types were validated and identified using RNA-FISH with confocal
546 imaging. To push the technique forward, RNA-FISH combined with Bessel beam
547 selective plane illumination microscope (BB-SIM) was used to obtain the absolute
548 quantification and subcellular location of transcripts in these new cell types. Optogenetic
549 manipulations were performed to alter the internal state of the animal, either with two
550 hours of fictive wasp sting or two hours of overactivation of 10% of brain neurons, and
551 scRNAseq was used once more to search for a change in molecular cell state between
552 conditions.

553

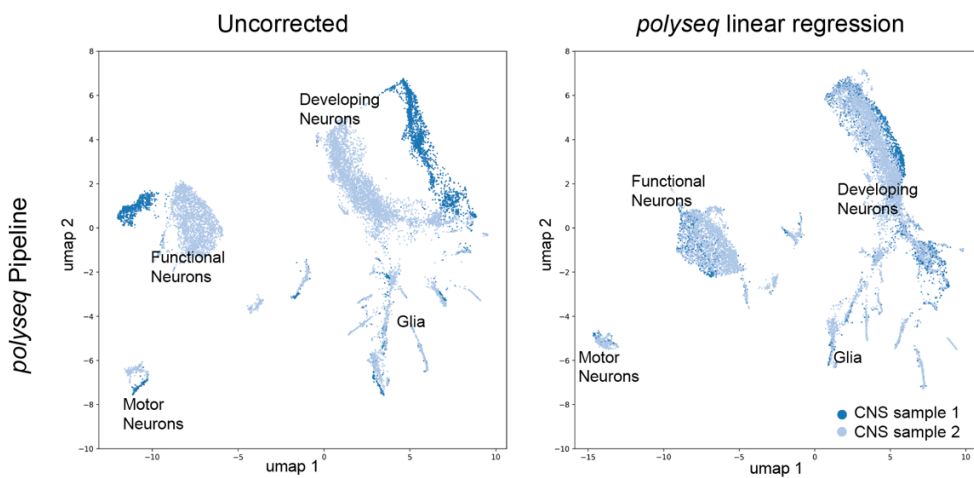


D

Batch Correction Method Comparison



E



554

555

556 **Figure 2. *Polyseq* python package performs cell type discovery and batch**
557 **correction.** A. The same dataset was analyzed in *polyseq* and *Seurat*. A confusion
558 matrix was generated to compare how cells were clustered together. Fifteen clusters
559 were found in both analyses, with 8 of 15 clusters containing more than 90% of the
560 same cells. Clusters that disagreed about cell placement were often differing between
561 two similar clusters (i.e. deciding whether a cell belonged to glia 1 vs glia 2). B. Garnett,
562 a newly developed unsupervised technique, was used to label cell groups with known
563 markers (Pliner et al., 2019). This analysis correctly annotated larval functional neurons,
564 motor neurons, glia, NPCs, and GMCs. It also provides an “unknown” label for cells with
565 low confidence. D,E. Batch correction performance was compared in *Monocle* and
566 *polyseq*. In the plots, cells are separated by a signature related to small differences in
567 sample collection rather than cell type signatures. Monocle’s align function correctly
568 collapses the separated developing neurons (blue and black in “uncorrected” plot) into a
569 comingled group. The linear regression method we implemented in *polyseq* also
570 collapses the sample separations (such as the separation of motor neurons, functional
571 neurons and developing neurons) in the “uncorrected” plot into a single, mixed group in
572 the umap plot following linear regression.

573

574

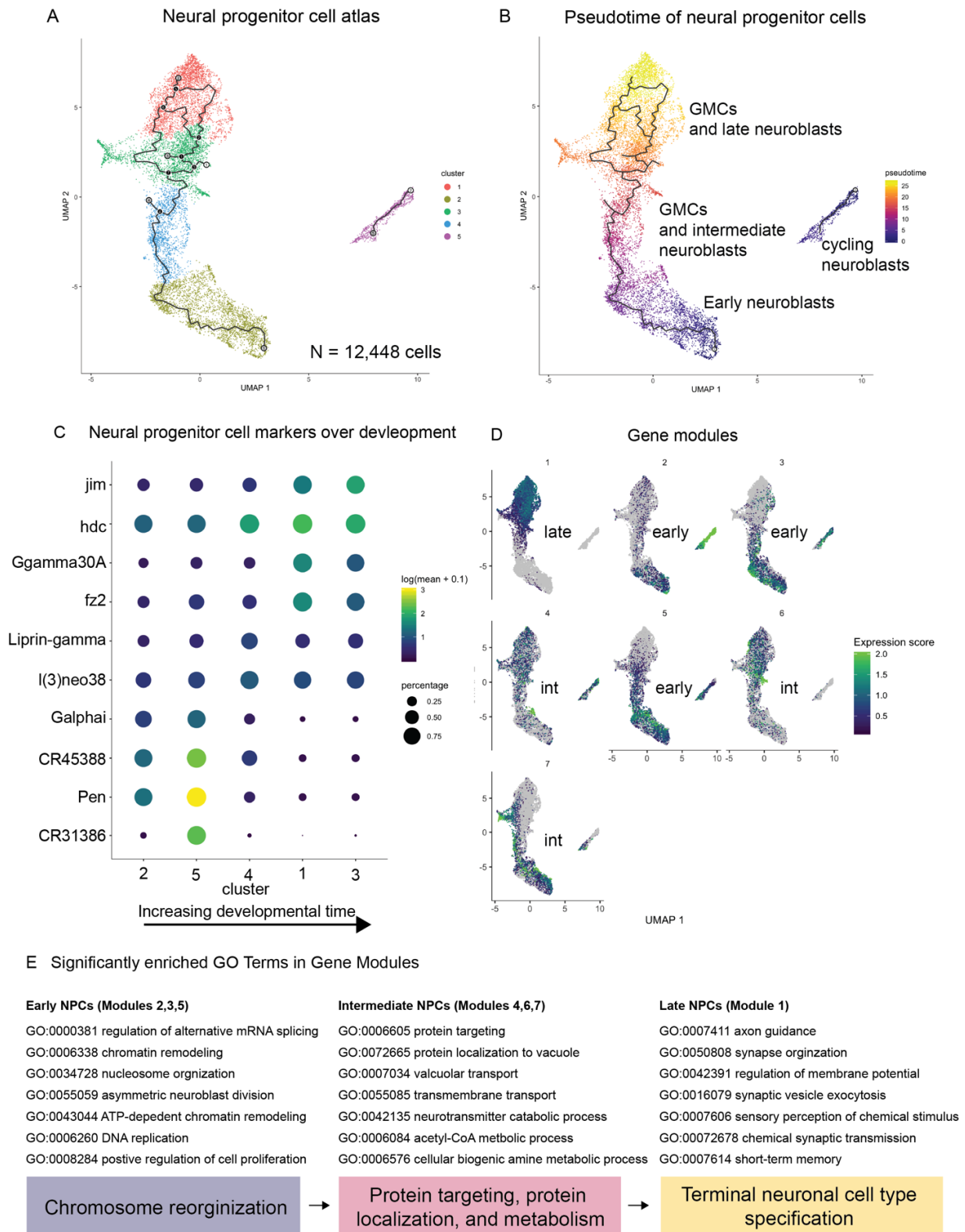
575

576

577

578

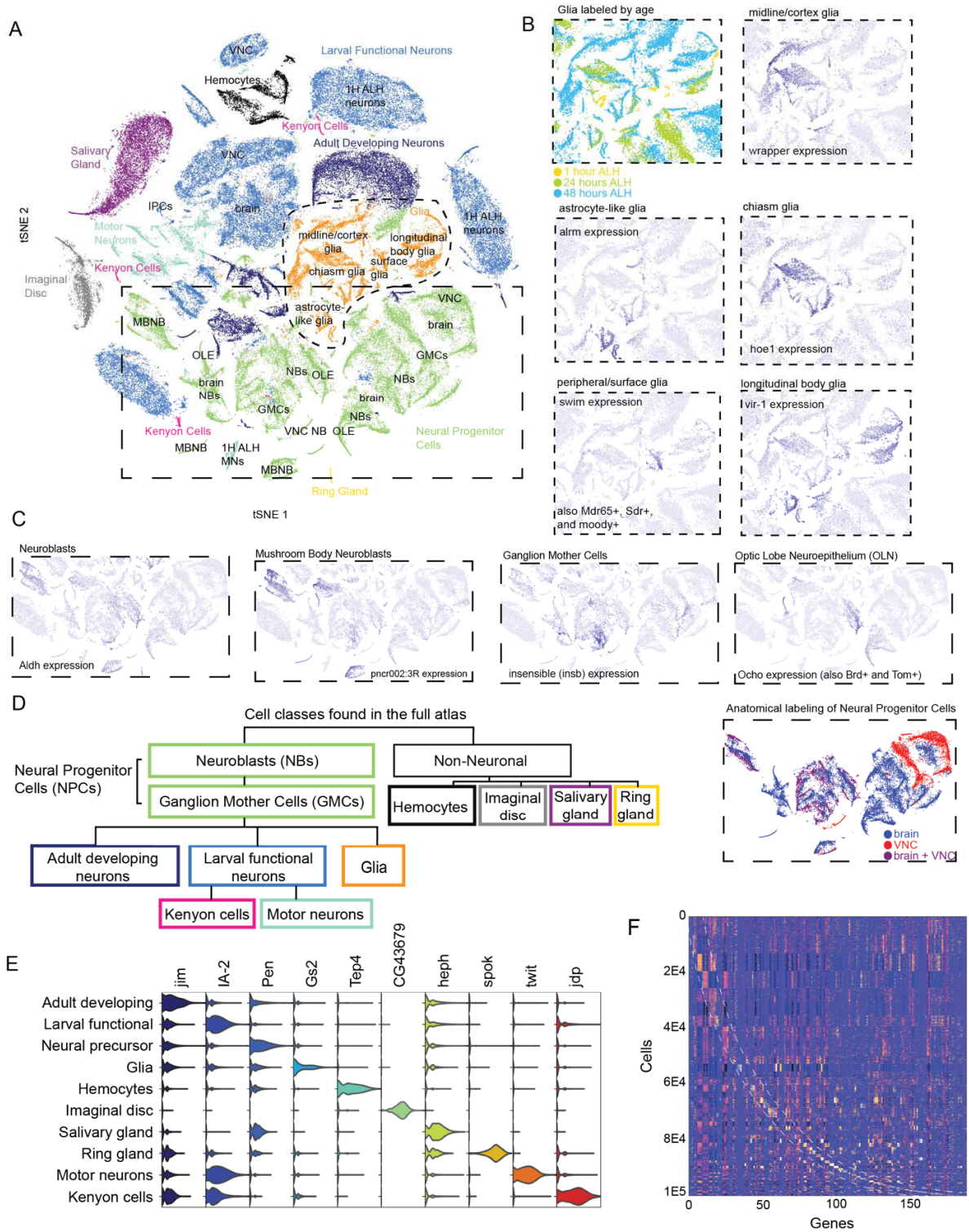
579



580

581

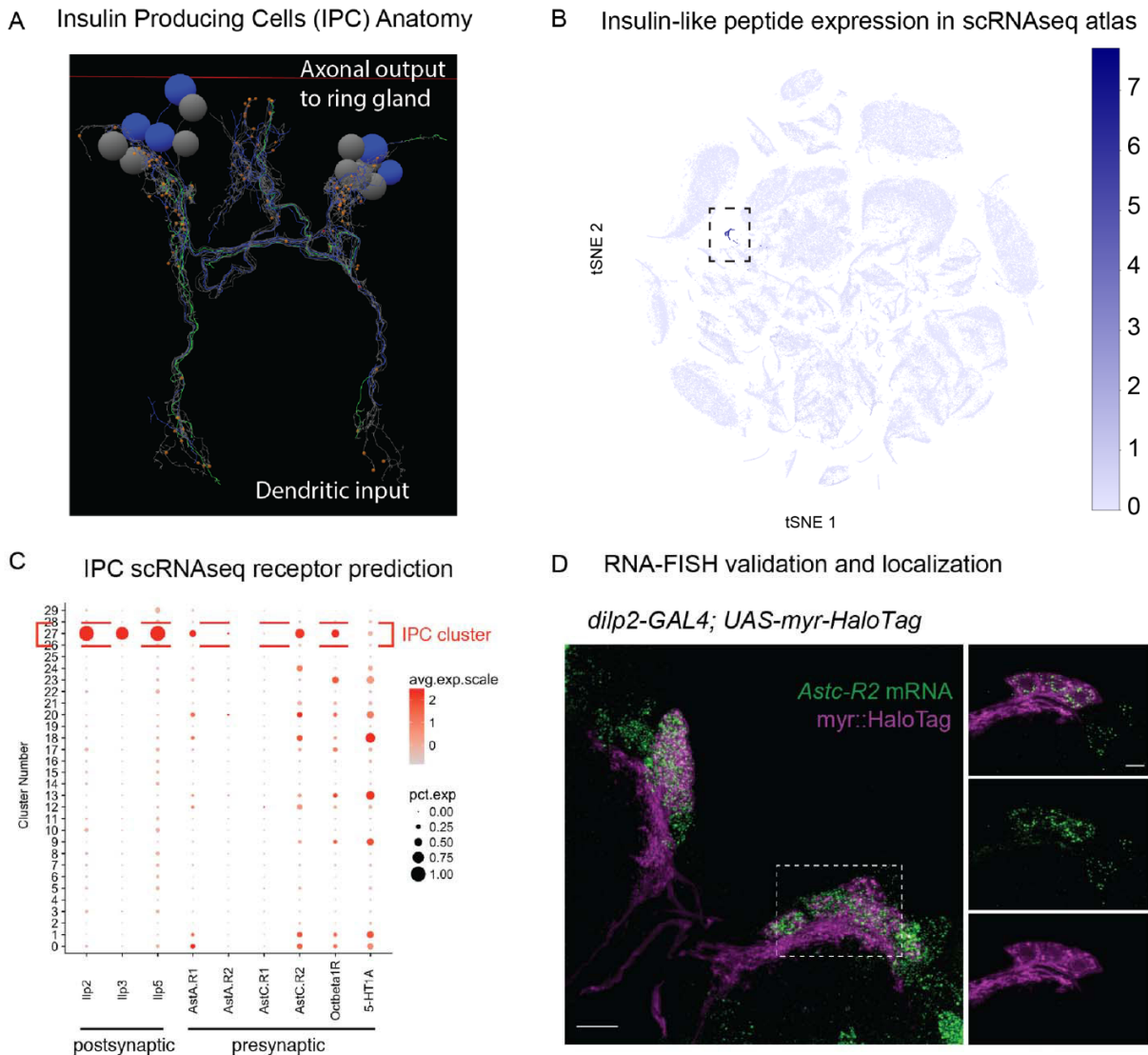
582 **Figure 3. Neural progenitor cell (NPC) atlas reveals gene modules across**
583 **developmental time.** A. A full atlas of first (1H), second (24H), and third (48H) instar
584 larval cells was built, and all NPCs were extracted. These NPCs were then analyzed
585 with Monocle and split into five clusters. B. A pseudotime analysis was performed using
586 known developmental times and separated the data into early, intermediate, and late
587 NPCs. A group of cycling neuroblasts was found to the right of the main NPC dataset in
588 UMAP space. C. Markers for each NPC cluster were extracted and revealed the change
589 in gene expression over developmental time. D. Gene modules were computed and
590 characterized early, intermediate, and late NPCs. As the gene modules represented
591 more developed cells, they were enriched for GO terms (E) which characterized more
592 developed cells (Table S2).
593
594
595



596

597

598 **Figure 4. Single-cell transcriptomic atlas of the larval central nervous system.** A. t-
599 SNE visualization of high-quality cells colored by cell class. Cells broke into 70 clusters
600 (Figure S4; Table S1) and were post-hoc identified as adult developing neurons, larval
601 functional neurons, neural progenitor cells, glia, hemocytes, imaginal disc, salivary
602 gland, ring gland cells, or subsets of these cells, such as motor neurons, Kenyon cells,
603 neuroblasts, or ganglion mother cells (Table S1). B. Five subtypes of glia, labeled in
604 orange, were found comingled in t-SNE space and could be distinguished by age and
605 function. C. Neural progenitor cells, labeled green in the t-SNE space, split into
606 recognizable classes, including neuroblasts, ganglion mother cells, and optic lobe
607 neuroepithelium. D. Diagram of cell classes contained in the atlas (see Figure S1 for
608 more information). E. Genes that define each cell cluster and cell class were
609 discoverable (Table S1). Violin plots show exemplar genes from each cell class. F.
610 Heatmap of all high-quality cells and the top 3 genes that define each cluster.
611
612
613
614

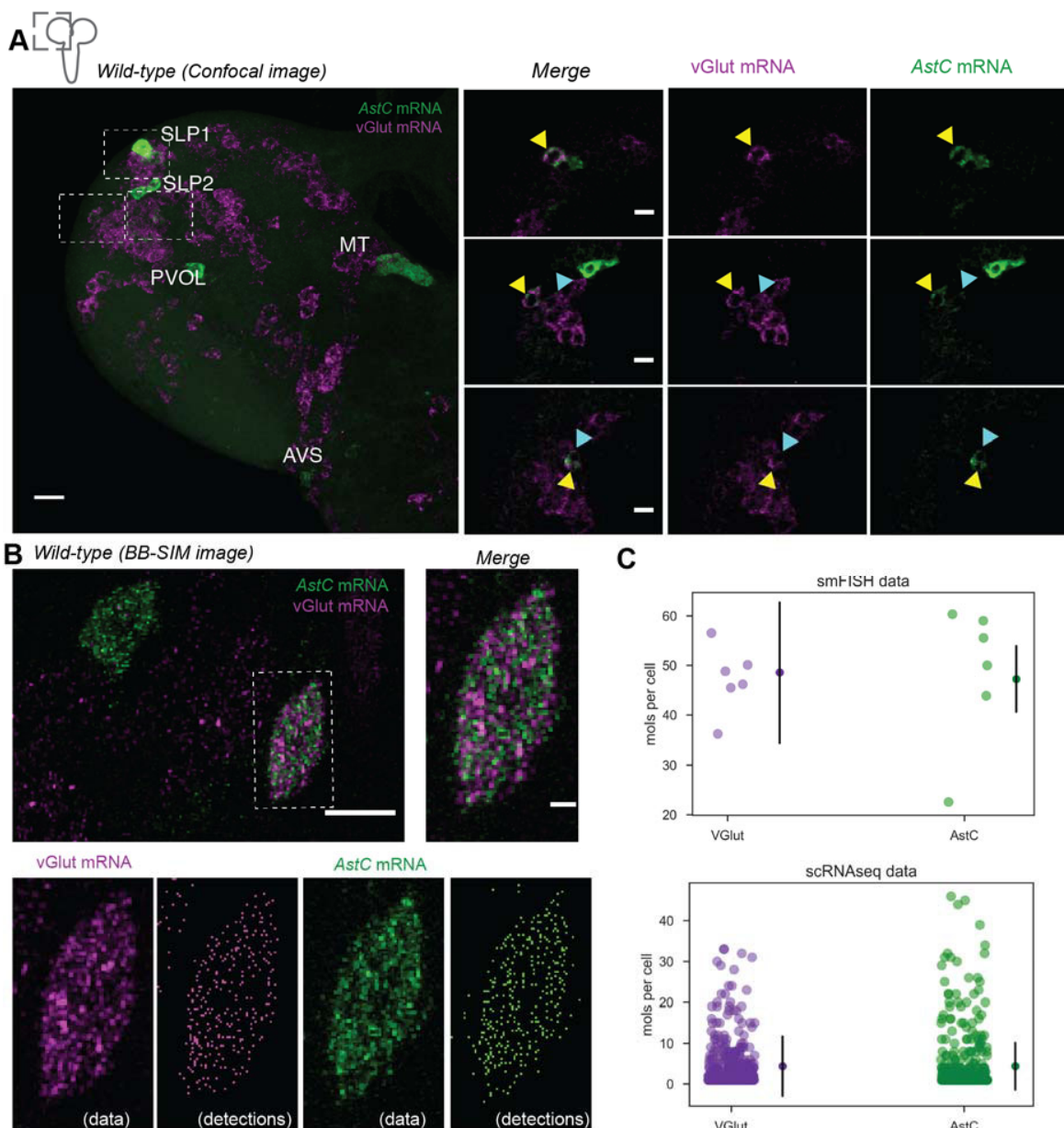


615

616

617 **Figure 5. Transcriptomic Atlas predicts previously unknown neuropeptide**
 618 **phenotype for insulin-producing cells and is verified by RNA-FISH.** A. Anatomy of
 619 insulin-producing cells (IPCs). The IPCs are a group of 7 bilaterally symmetrical
 620 neurons which receive input through their dendrites (purple) about the nutritional state
 621 of the animal and release insulin-like peptides (ILP2, ILP3, ILP5) through their axons
 622 (green), which synapse on the ring gland, to control carbohydrate balance. They are
 623 analogous to the vertebrate pancreatic beta islet cells. B,C. The RNAseq atlas built in
 624 this study discovered the IPCs as a separate cluster (cluster 27 in C) with expression in
 625 the IPCs of receptors for octopamine, serotonin, and allatostatin A, which matched

626 previous literature. Surprisingly, the atlas also suggested the presence of a previously
 627 unrecognized receptor in the IPCs for allatostatin C (AstC-R2). D. Detection of AstC-R2
 628 mRNA in IPCs. Maximum-intensity projection of the confocal stack of a brain in which
 629 the IPCs are labeled with a fluorescent HaloTag ligand (Magenta) and AstC-R2 mRNA
 630 is detected by FISH (green), bar 10 μ m. Dashed lines outline area where the single z
 631 plane is shown on the left panels, bar 5 μ m. Movie of D is in the supplement (Movie S1).
 632
 633

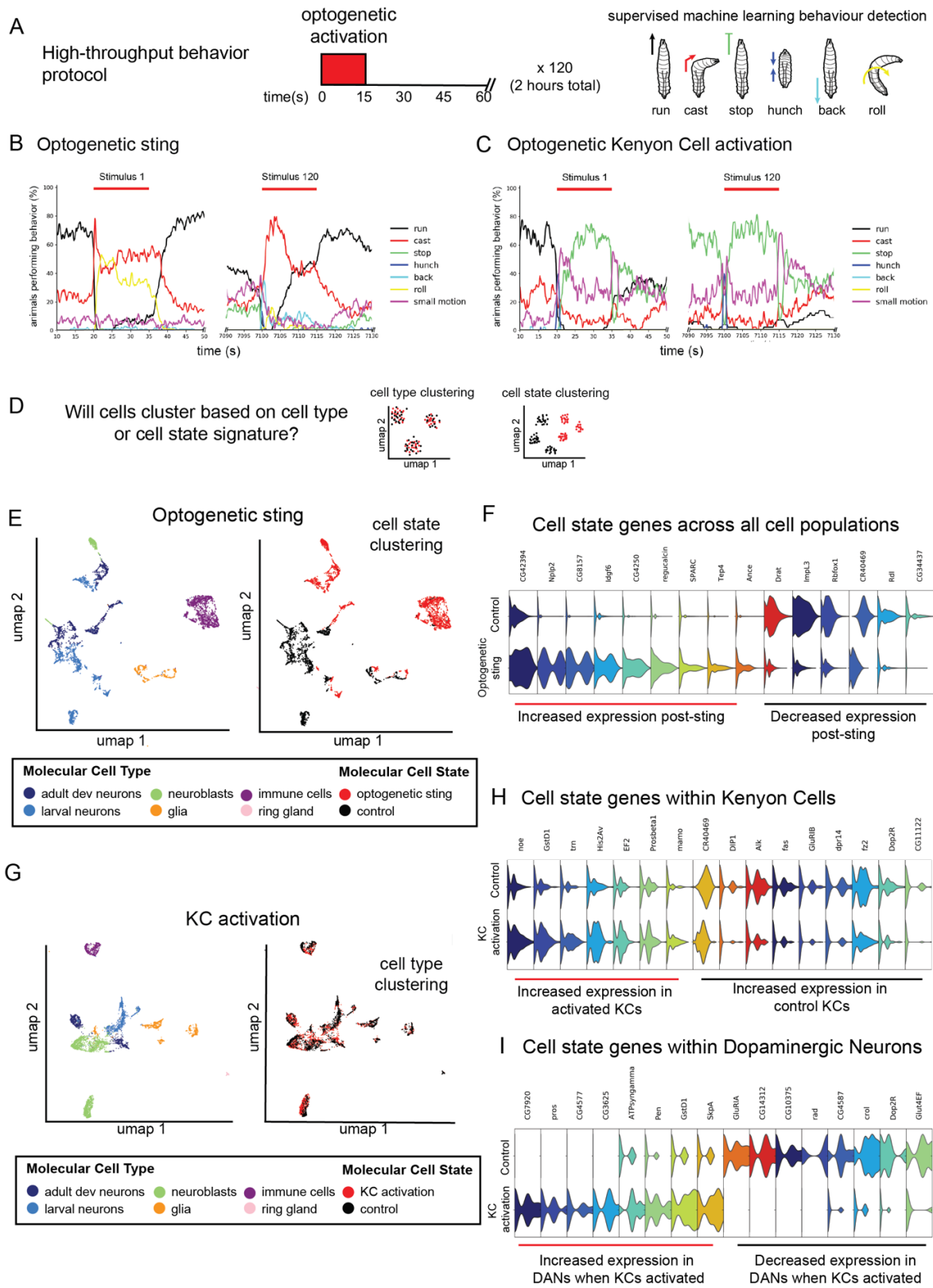


634

635

636 **Figure 6. Correlation between single-molecule FISH and scRNAseq.** A. Identifying
637 cells that have coexpression of AstC (green) and vGlut (magenta) in the whole brain B.
638 Maximum-intensity projections of BB-SIM stack of the AstC and vGlut mRNA FISH
639 channels, bar 10 μ m. Dashed lines outline 2 cells that co-express AstC and vGlut
640 mRNAs are shown on the right panels, bar 1 μ m. Lower panels show individual FISH
641 channel and the reconstructions obtained using the spot-counting algorithm. C.
642 Comparison of the quantification of AstC and vGlut mRNAs between smFISH and
643 scRNAseq.

644



646 **Figure 7. Optogenetic sting and KC activation induce a change in internal state,**
647 **which is observable as cell state changes in gene expression.** A. High-throughput
648 behavior assays were performed using previously described equipment (Ohyama et al.,
649 2013). Fifteen seconds of optogenetic activation were delivered every 60 seconds over
650 a two-hour period. Larval behavior was recorded and supervised machine learning
651 techniques were used to classify behavior as one of six behaviors. B. For the
652 optogenetic sting protocol, animals first respond to basin activation by performing an
653 escape response, including a fast turning response and rolling as previously described
654 (Hwang et al., 2007; Ohyama et al., 2013; Ohyama et al., 2015). Over the course of the
655 experiment, a switch in behavior is observed from rolling to backing up. C. Kenyon cell
656 activation led to a hunch followed by freezing. At the start of training, animals crawled
657 forward to offset. After training, the hunch and freezing response remained at the onset,
658 but the offset response switched to a turn rather than a forward crawl. D. Transcriptomic
659 atlases were produced from animals that underwent each behavior protocol and
660 matched controls. We were primarily interested in whether cell state or cell type
661 clustering would be observed after altering internal state (sting) or overactivating the
662 memory center (KC activation). Cell type clustering would result in cells from activated
663 animals and controls mixing in the same clusters while cell state clustering would lead
664 to separation of cells from activated and control animals. E,F. The optogenetic sting
665 protocol led to cell state clustering. All cell classes were identifiable for both conditions.
666 A large immune cell population was specific to the sting condition, suggesting an
667 immune response to a fictive wasp sting (Figure S8). F. Cell state genes were
668 discovered with upregulated and downregulated expression that separated all cell
669 populations. G-I. KC activation led to cell type clustering and local cell state changes.
670 Even though a larger total number of cells were activated (~200 KCs versus 64 basins),
671 a less dramatic switch was observed in behavior and in the transcriptome. There were
672 differentially expressed genes in two key populations of cells in the learning and
673 memory center: KCs and dopaminergic neurons.
674
675
676

677 **METHODS:**

678 **CONTACT FOR REAGENT AND RESOURCE SHARING**

679 Further information and requests for resources and reagents should be directed to and
680 will be fulfilled by the Lead Contact, Marta Zlatic (zlati@m@janelia.hhmi.org).

681

682 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

683 **Fly stocks**

684 *Drosophila* larvae were grown on standard fly food at 25°C and kept in 12-hour
685 day/night light and dark cycle. Vials were timed by collecting eggs on a new food plate
686 over the course of one hour.

687

688 Please see Key Resources Table for *Drosophila* lines used in this study.

689

690 **METHOD DETAILS**

691 **Single cell isolation**

692 *Drosophila* larvae were dissected at 1 hour, 24 hours, 48 hours, or 96 hours after larval
693 hatching (ALH). All dissections were performed in a cold adult hemolymph solution
694 (AHS) with no calcium or magnesium at pH 7.4. Quality of single cell isolation was
695 investigated by visual inspection with compound and confocal microscopy. Samples
696 were placed on ice during waiting periods. Samples were isolated and run on the 10x
697 Chromium Single Cell 3' immediately after cell dissociation.

698

699 First, the complete central nervous system (CNS) was dissected from every animal. The
700 dissected nervous systems were kept in cold AHS on ice. For those samples where the
701 brain and the ventral nerve cord (VNC) were sequenced separately, the separation of
702 the brain from the VNC was performed using fine-tipped forceps and MicroTools (Cat #:
703 50-905-3315, Electron Microscopy Sciences). The time from digestion (the part of the
704 protocol most likely to induce cell stress) to on the 10x Genomic instrument was never
705 longer than 30 minutes.

706

707 After separation of the brain from the VNC, the desired tissue was placed in 18 μ L of
708 AHS on ice. Once all samples were prepared, 2 μ L of 10x neutral protease (Cat #:
709 LS02100, Worthington Biochemical Corp, Lakewood, NJ, USA) was added to a final
710 volume of 20 μ L. The intact brain tissue was digested for 5 minutes. The tissue was
711 then transferred to a fresh drop of 20 μ L of AHS.

712

713 Each sample was triturated with a clean, thinly pulled glass electrode until no tissue was
714 visible under a dissection scope. All debris (pieces of nerve and undigested neuropile)
715 was removed. Samples with fluorescent markers were observed under a fluorescence
716 microscope to approximate cell density. The samples were then loaded onto the 10x
717 Chromium chip.

718

719 **10X Genomics**

720 Single cell capture and library construction was performed using the 10x Chromium
721 machine and the Chromium Single Cell 3' v2 Library and Gel Bead Kit (10x Genomics,
722 Pleasanton, CA). Manufacturer's recommendations were followed for cell collection and
723 library construction. Libraries were sequenced with an Illumina HiSeq following
724 manufacturer's instructions.

725

726 **mRNA *in situ* hybridization**

727 FISH probes were designed based on transcript sequences using the online Stellaris
728 Designer and purchased from Biosearch Technologies. Probe sequences for AstC and
729 vGlut were previously reported (Long et al., 2017; Diaz et al., 2019), and probe
730 sequences for AstC-R2, Hug, NPNL2 are in Table S5. Each probe is 18-22nt long with a
731 3' end amine-modified nucleotide that allows directly couple to an NHS-ester dye
732 according to the manufacturer's instructions (Life Technologies). Dye-labeled probes
733 were separated from the excess free dyes using the Qiagen Nucleotide Removal
734 Columns. FISH protocol was described previously (Long et al., 2017; Diaz et al., 2019).
735 The brains of 3rd instar larvae were dissected in 1xPBS and fixed in 2%
736 paraformaldehyde diluted PBS at room temperature for 55 min. Brain tissues were
737 washed in 0.5% PBT, dehydrated, and stored in 100% ethanol at 4°C. After exposure to

738 5% acetic acid at 4 °C for 5 minutes, the tissues were fixed in 2% paraformaldehyde in
739 1xPBS for 55 min at 25 °C. The tissues were then washed in 1× PBS with 1% of NaBH₄
740 at 4 °C for 30 min. Following a 2 hour incubation in prehybridization buffer (15%
741 formamide, 2× SSC, 0.1% Triton X-100) at 50 °C, the brains were introduced to
742 hybridization buffer (10% formamide, 2× SSC, 5× Denhardt's solution, 1 mg/ml yeast
743 tRNA, 100 µg/ml, salmon sperm DNA, 0.1% SDS) containing FISH probes at 50 °C for
744 10 h and then at 37 °C for an additional 10 h. After a series of wash steps, the brains
745 were dehydrated and cleared in xylenes.

746

747 **Confocal and BB-SIM Imaging**

748 For confocal imaging, the tissues were mounted in DPX. Image Z-stacks were collected
749 using an LSM880 confocal microscope fitted with an LD LCI Plan-Apochromat 25x/0.8
750 oil or Plan-Apochromat 63x/1.4 oil objective after the tissue cured for 24 hours. For
751 single-molecule imaging, we use a previous described Bessel beam selective plane
752 illumination microscope (BB-SIM). Detail construction of microscope and the imaging
753 procedure is described previously (Long et al., 2017). Briefly, this BB-SIM is engineered
754 to image in medium matched to the measured refractive index (RI) of xylene-cleared
755 *Drosophila* tissue with axial resolution of 0.3 µm and lateral resolution of 0.2 µm. For
756 BB-SIM imaging, the tissues were mounted on a 1.5x3mm poly-lysine coated coverslip
757 attached to a 30mm glass rod. The imaging process requires the objectives and tissues
758 immersed in the imaging medium consist with 90% 1,2-dichlorobenzene, 10% 1,2,4-
759 trichlorobenzene with refractive index = 1.5525. Two orthogonally mounted excitation
760 objectives are used to form Bessel beams, which are stepped to create an illumination
761 sheet periodically striped along x or y, while a third objective (optical axis along the z
762 direction) detects fluorescence. To employ structured illumination analysis, we collect
763 multiple images with the illumination stripe pattern shifted to tile the plane in x, and
764 repeat the process orthogonally to tile the plane in y. The sample is then moved in z,
765 and the imaging repeated, and so on to image the 3D volume.

766

767 **High-throughput Automated Optogenetic Behavior Experiments**

768 For the sting mimic experiments, 72F11-GAL4 males were crossed to UAS-CsChrimson
769 virgins (stock information in Key Resources Table). For Kenyon cell overactivation,
770 201Y-GAL4 males were crossed to UAS-CsChrimson virgins. Larvae were grown in the
771 dark at 25°C. They were raised on standard fly food containing trans-retinal (SIGMA
772 R2500) at a final concentration of 500 μ M. Activation was performed in a high-
773 throughput optogenetic behavior rig described previously (Ohyama et al., 2013). About
774 40 animals were placed in a 25 x 25 cm² dish covered with clear 4% agar.

775
776 Neurons were activated using a red LED at 325 μ W/cm² illuminated from below the agar
777 dish for 15 seconds with with a 45 second rest period for a total of 120 activation
778 periods (Figure 7A). Supervised machine learning was used to automatically detect
779 behavior (Jovanic et al., 2017). Control animals of the same age were collected from the
780 same food plate as experimental animals and placed on an agar plate in the dark for
781 two hours. Immediately following the sting protocol, 2-4 animals from each group were
782 dissected and cells were collected using the 10X Genomics protocol described above.

783 **QUANTIFICATION AND STATISTICAL ANALYSIS**

784 **scRNA-seq analysis**

785 Bioinformatic analysis was performed using Cell Ranger software (Version 1.3.1, 10x
786 Genomics, Pleasanton, CA, USA), the Seurat R package (Satija et al., 2015) and
787 custom software in R and Python, including the *polyseq* Python package developed
788 here. Software to train classifiers using neural networks was built with TensorFlow. The
789 *polyseq* package as well as jupyter notebooks containing code used for analysis in the
790 study are available on GitHub (<https://github.com/jwittenbach/polyseq>).

791
792 Briefly, Cell Ranger was used to perform demultiplexing, alignment, filtering, and
793 counting of barcodes and UMIs, with the output being a cell-by-genes matrix of counts.
794 To further ensure that only high-quality cells were retained, any cell that registered
795 counts in a unique number of genes below a baseline threshold was removed. To
796 reduce the dimensionality of the data for computational tractability, any gene that was
797 not expressed in a baseline number of cells was also dropped.

798

799 To account for the fact that raw counts tended to span many orders of magnitude (~
800 10^0 - 10^5), counts were transformed via $\log(\text{counts} + 1)$. To control for cell size and
801 sequencing depth, the sum of the (log-transformed) counts within each cell used as a
802 regressor for a linear regression model to predict the (log-transformed) counts for each
803 gene (one linear regression model per gene, with each cell being a sample). “Gene
804 expression levels” were then quantified as the z-scored residuals from the fitted models
805 (i.e. standard deviations above/below the predicted log-transformed counts for a
806 particular gene across all cells).

807

808 Next, to further reduce the dimensionality of the data in preparation for downstream
809 clustering and embedding operations (both of which have computational costs that
810 scale poorly with the dimensionality of the feature space), principal component analysis
811 was performed with cells as samples and gene expression levels as features. The top K
812 principal components (PCs) were retained as features for downstream analyses. For the
813 larger cell atlas dataset, K was chosen to retain a desired percentage of the total
814 variance. For the smaller cell state datasets, K was chosen automatically via a shuffle
815 test – on each shuffle, gene expression levels for each gene were randomly permuted
816 across all cells and the percent variance explained by the top PC was recorded; the 95th
817 percentile of this value across all shuffles was then used as a threshold to determine the
818 cutoff point for keeping PCs with respect to percent variance explained by a particular
819 PC.

820

821 Based on these top PCs, cells were clustered using the Louvain-Jaccard graph-based
822 clustering approach. Briefly, the k-nearest neighbor graph between cells was is
823 computed. Edge weights are then determined using the Jaccard index, which measures
824 the fraction of shared neighbors between any two nodes. Finally, the Louvain
825 community detection algorithm is applied to this graph to partition the nodes into
826 clusters; this algorithm seeks to optimize weight of connections with each cluster
827 relative to those between clusters.

828

829 In order to visualize the results of the analysis, the PC features were also used to
830 perform a nonlinear embedding into two dimensions. This was performed via either the
831 t-SNE or the UMAP algorithm.

832

833 Once cluster identities were determined, the original gene expression level data was to
834 determine important genes for defining each cluster. For each cluster, gene expression
835 levels were used as features, and a binary indicator of whether or not a cell came from
836 the cluster in question was used as a target. This data was then used to fit a linear
837 classifier (viz, a support vector classifier) to separate in-cluster cells from the rest of the
838 population. The unit normal vector from the linear classifier was then extracted and the
839 components used to rank order genes in terms of importance for defining that cluster.
840 This same technique was also used to find important genes for groups defined by
841 methods other than clustering.

842

843 **Imaging analysis**

844 To quantify the number of vGlut and AstC mRNAs in cells contain both vGlut and AstC,
845 we first manually segmented cells from BB-SIM z-stacks that have both vGlut and AstC
846 FISH signals using the Fiji plugin TrakEM2 (Schindelin et al., 2012; Meissner et al.,
847 2019). After identifying the individual fluorescent spots in segmented cells used a
848 previously described Matlab algorithm (Lionnet et al., 2011), we calculate the number of
849 mRNAs per cell. Reconstructed images were generated using Matlab code that draws
850 spots centered on each of the detected spots positions (Lionnet et al., 2011).

851

852 **DATA AND CODE AVAILABILITY**

853 All code and documentation for *polyseq* is open source and freely available on github
854 (<https://github.com/jwittenbach/polyseq>). Jupyter notebooks used for analysis are
855 available upon request. The scRNA-seq data has been deposited in GEO and is
856 accessible under the accession code GEO: GSE135810.

857

858

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
RNA-FISH probes	Biosearch Technologies	N/A
RNase-free 1x PBS	Fisher Scientific	BP2438-4
Acetic Acid, Glacial (Certified ACS), Fisher Chemical	Fisher Scientific	A38S-500
Sodium borohydride, 99%, VenPure™ SF powder	Acros Organics	AC448481000
SSC (20X)	Fisher	AM9763
Hi-Di formamide	Applied Biosystems	4311320
Denhardt's solution (50X)	Alfa Aesar	AAJ63135AD
tRNA from baker's yeast	Roche	10109495001
UltraPure™ Salmon Sperm DNA Solution	Fisher Scientific	15632011
SDS, 10%	Corning	46-040-CI
Deionizedformamide	Ambion	AM9342
Chromium Single Cell 30 Library & Gel Bead Kit v2	10x Genomics	PN-120237
Chromium Single Cell A Chip Kit	10x Genomics	PN-120236
Chromium i7 Multiplex Kit	10x Genomics	PN-120262
Deposited Data		
Raw and analyzed scRNAseq data	This paper	GEO: GSE135810
Experimental Models: Organisms/Strains		
<i>D. melanogaster</i> : w[1118]; P{y[+t7.7] w[+mC]=GMR57C10-GAL4}attP2	Bloomington <i>Drosophila</i> Stock Center	RRID:BDSC_39171
<i>D. melanogaster</i> : pJFRC29-10XUAS-IVS-myr::GFP-p10 in attP40; pJFRC105-10XUAS-IVS-nlstdTomato in VK00040	Jack Etheredge; Etheredge, 2017	N/A
<i>D. melanogaster</i> : w[1118]; P{y[+t7.7] w[+mC]=GMR72F11-GAL4}attP2	Bloomington <i>Drosophila</i> Stock Center	RRID:BDSC_39786
<i>D. melanogaster</i> : P{GawB}Tab2[201Y]	Bloomington <i>Drosophila</i> Stock Center	RRID:BDSC_4440
<i>D. melanogaster</i> : w[1118] P{y[+t7.7] w[+mC]=20XUAS-IVS-CsChrimson.mVenus}attP18	Bloomington <i>Drosophila</i> Stock Center	RRID: BDSC_55134
Software and Algorithms		
polyseq	This paper	https://github.com/jwittenbach/polyseq
bcl2fastq	Illumina	https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html ; RRID: SCR_015058

Cell Ranger	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome
Seurat	Satija et al., 2015	https://satijalab.org/seurat/ ; RRID: SCR_007322
TensorFlow	Google Brain Team	https://www.tensorflow.org/install ; RRID:SCR_016345
Monocle3	Cao et al., 2019	https://cole-trapnell-lab.github.io/monocle3/
Garnett	Pliner et al., 2019	https://cole-trapnell-lab.github.io/garnett/
topGO	Alexa and Rahnenfuhrer J, 2019	https://bioconductor.org/packages/release/bioc/html/topGO.html ; RRID:SCR_014798
Fiji	PMID:22743772	http://fiji.sc ; RRID:SCR_002285
Other		
Jupyter notebooks of data analysis used in this paper	This paper	https://github.com/jwittenbach/polyseq

859

860

861

862

863

864

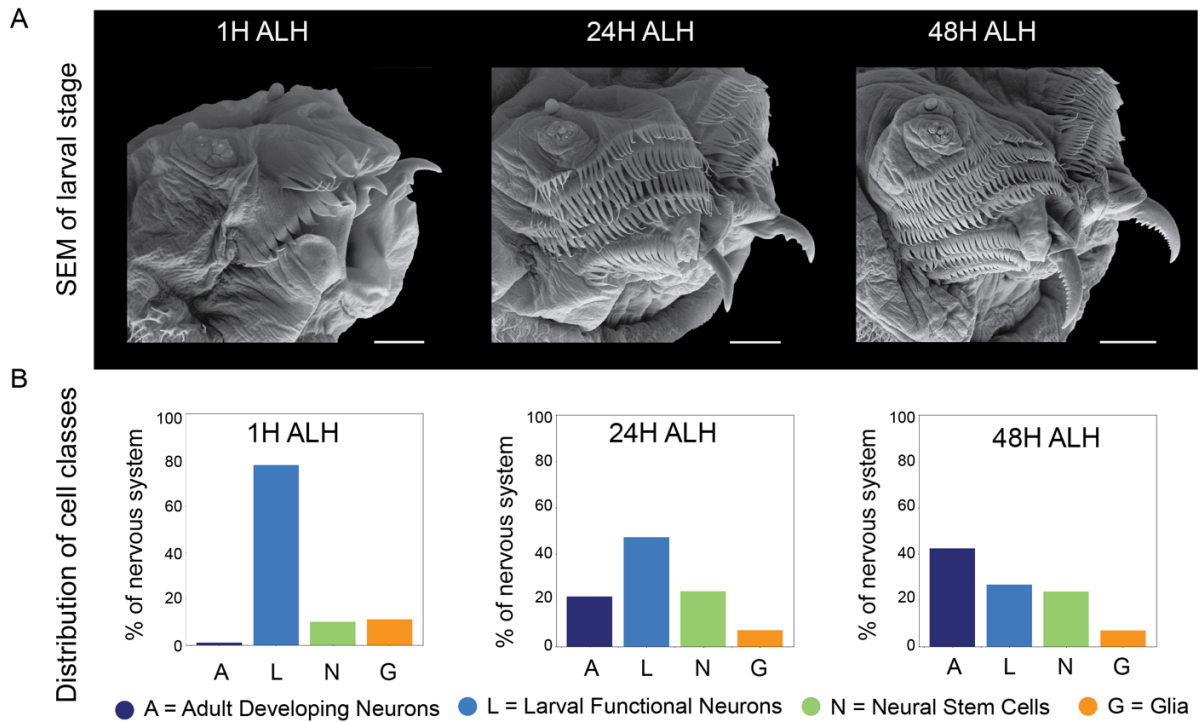
865

866

867

868 **Supplemental Information titles and legends:**

869



870

871 **Figure S1. Single cell transcriptomic atlas of the larval nervous system across**

872 **developmental stages.** A. Scanning electron microscopy (SEM) image of larval

873 anatomy during development. Three developmental stages were sequenced separately

874 to build age specific atlases (Scale bar = 10 μ m for 1H, 20 μ m for 24H and 40 μ m for 48

875 H). Larval size and elaboration of mouth hooks continues during this developmental

876 period. B. Distribution of cell classes. Just after hatching (1 hour after larval hatching

877 (ALH)), there is one primary central cluster of neurons. Neuron classes are

878 recognizable, including cholinergic neurons, motor neurons, astrocyte-like glia, and

879 neuroblasts. At 24 hours ALH multiple main groups of neurons are recognizable with

880 markers consistent with an increase in neuroblasts and ganglion mother cells (GMCs).

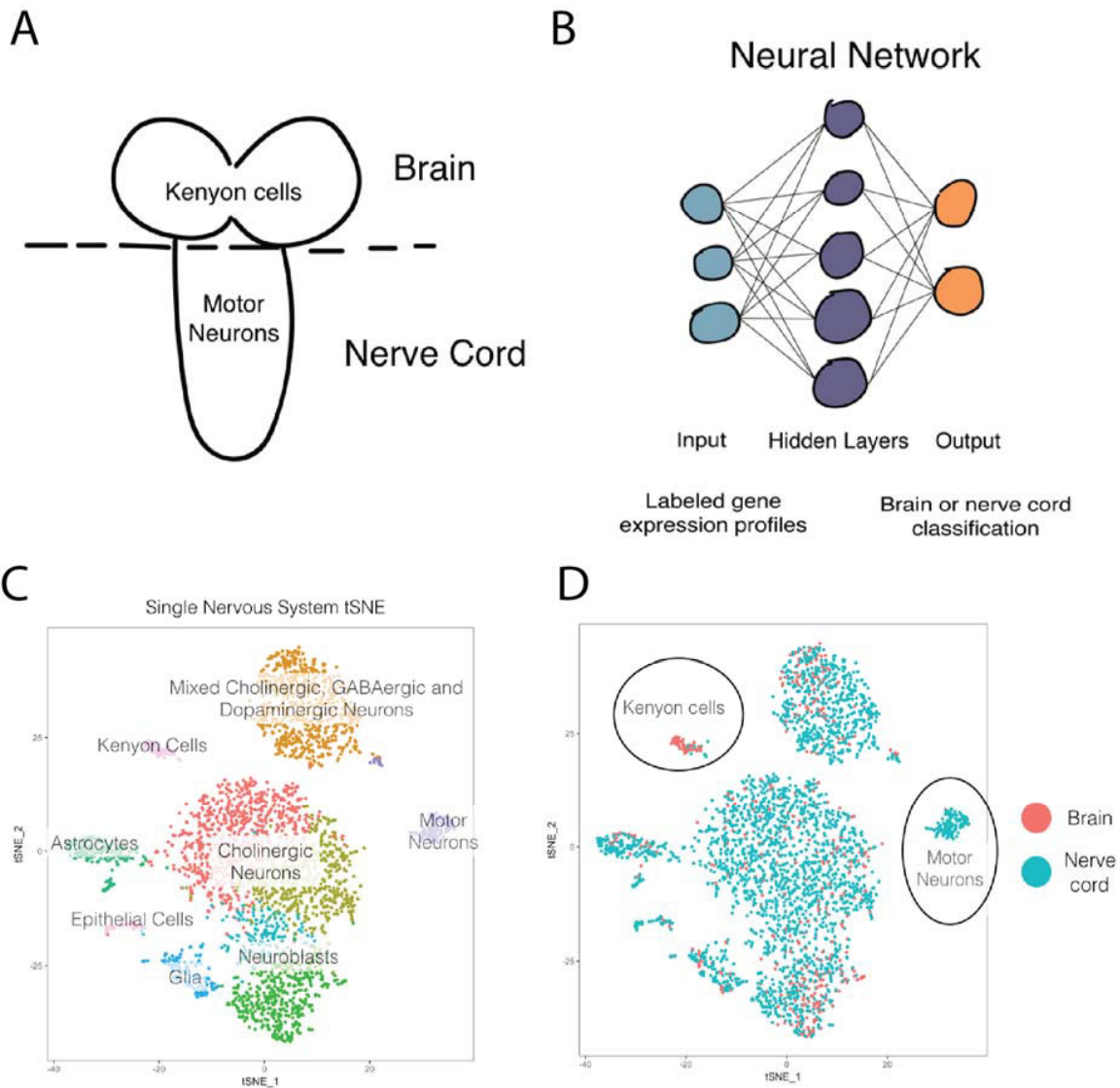
881 In addition, there are large groups of neurons, here labeled “developing adult neurons”

882 which have neuron markers but few or no genes expressed for synaptic transmitters

883 and receptors. This is consistent with the large burst of neurons born at this point in

884 development which lay dormant until adult life. At 48 hours ALH the functional larval and
885 developing adult populations are still identifiable.

886

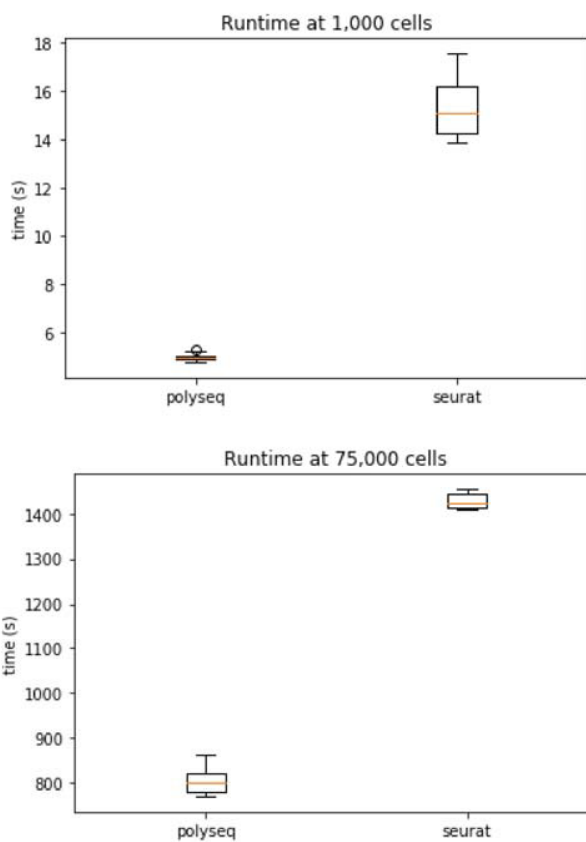


887

888 **Fig S2. Machine learning separates brain and VNC neurons.**

889 We dissected the brain from the nerve cord and sequenced the RNA from each
890 population separately. This provided ground truth labels which we could then feed into
891 (B) a neural network to train a classifier to predict spatial origin from the brain of the

892 VNC. C. We then used our classifier (built on separate data) to predict the brain or VNC
893 origin of cells from a nervous system that was sequenced in one sample (no brain or
894 VNC dissection). D. The classifier correctly separated brain from VNC cells in an intact
895 sample. Most clusters were appropriately mixed brain and VNC cells; however, one
896 pure VNC population (motor neurons) and one pure brain population (Kenyon cells)
897 were appropriately labeled by our machine learning classifier.
898
899



900

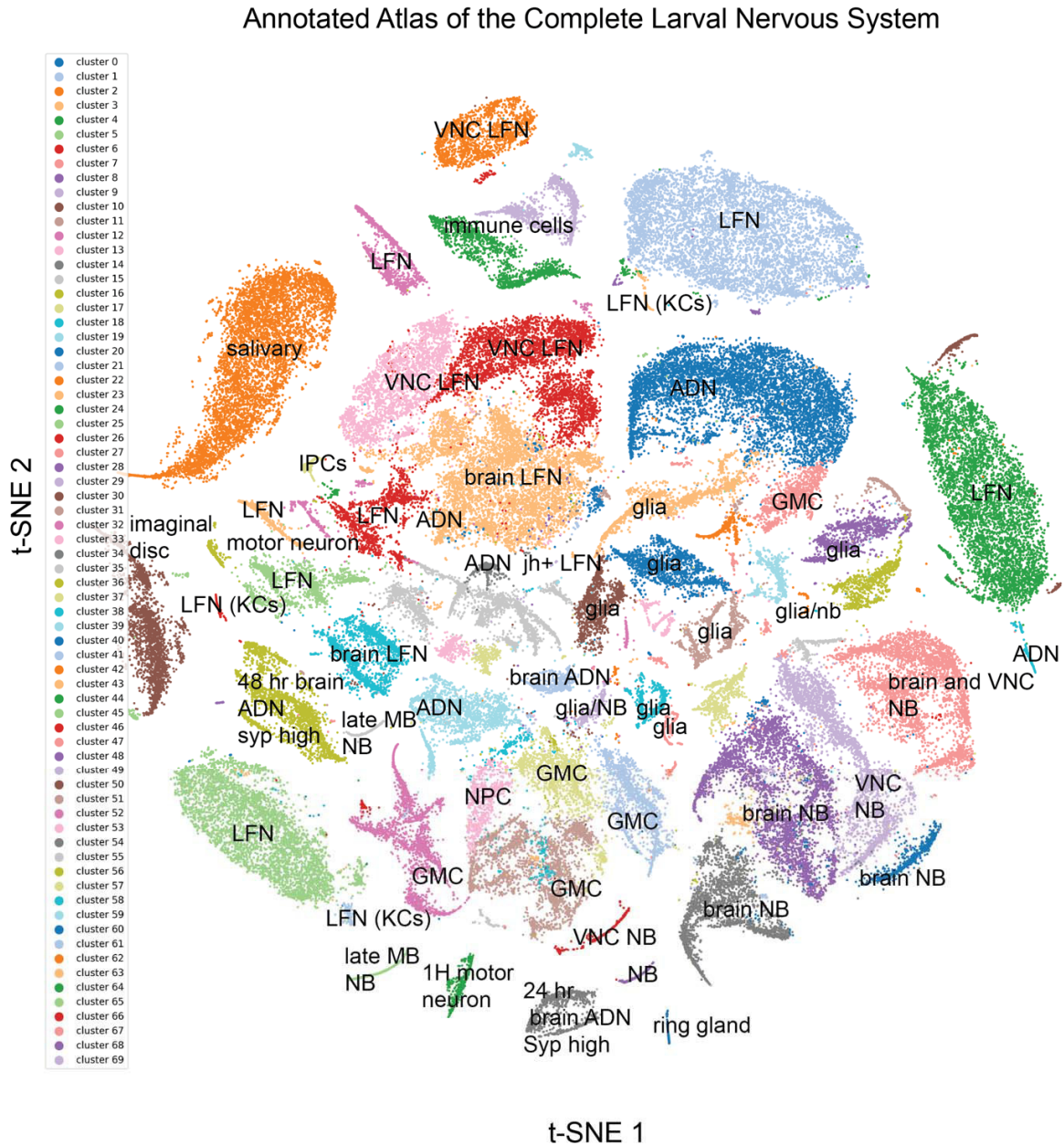
901 **Figure S3. Runtime for a typical pipeline to analyze single cell data in *Seurat* (R)**
902 **versus *polyseq* (python).**

903 Equivalent analysis pipelines were built in *Seurat* and *polyseq* to analyze single cell
904 RNAseq data for datasets of 1,000 cells and 75,000 cells on a single laptop running
905 Mac OS with 16GB RAM and a 2.6 GHz Intel Core i7. Polyseq outperformed Seurat for

906 both small and medium datasets. Jupyter notebooks used for testing that contain code
907 for Seurat and polyseq is available upon request.

908

909

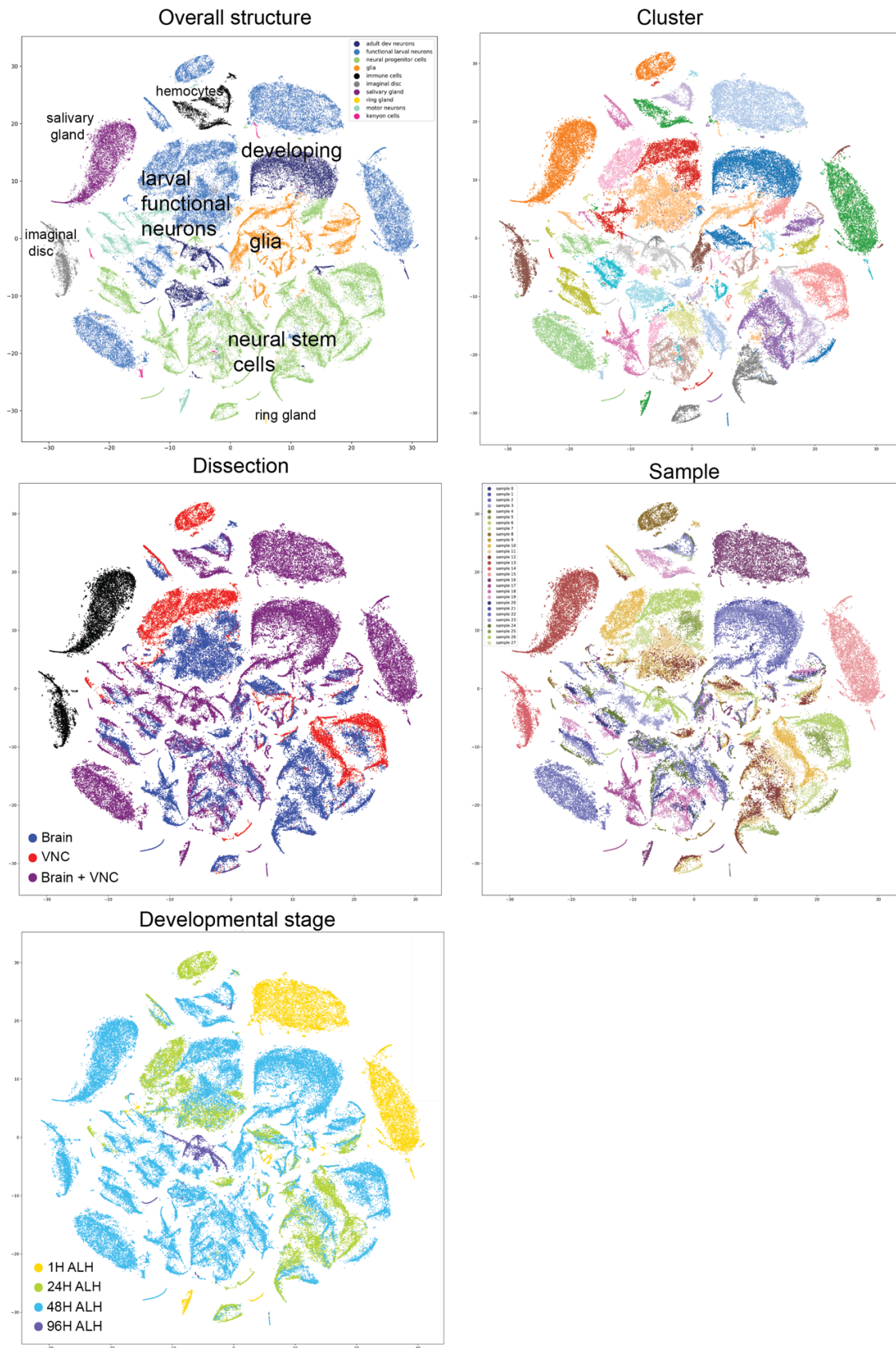


910

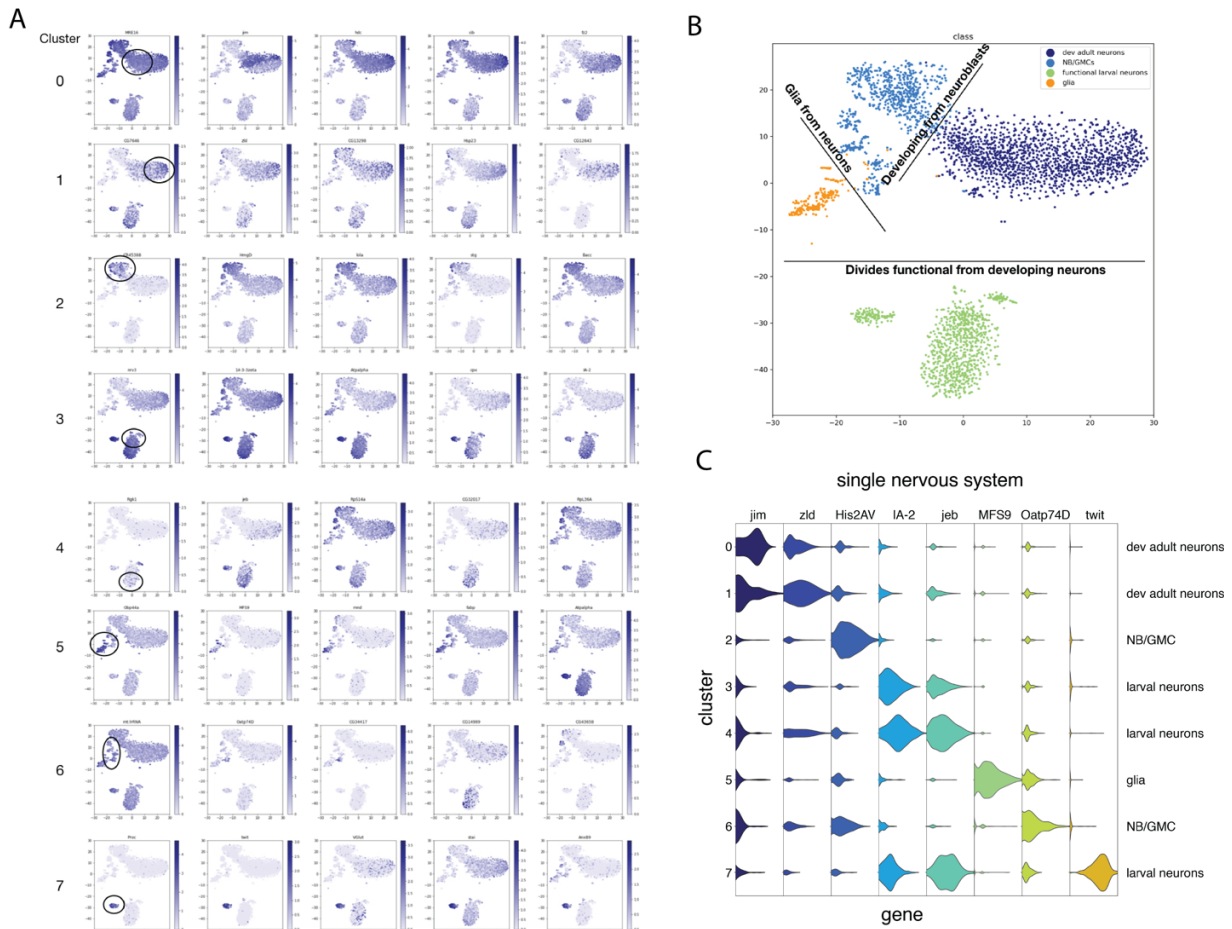
911 **Figure S4. Labeled atlas of the complete larval nervous system.** The atlas split into

912 70 clusters, which are labeled by color and class or more specific name. The top genes

913 which define each cluster, along with the mean expression within the cluster, the
914 expression outside the cluster, and the p-value can be found in Table S1.
915



917 **Figure S4. Complete single-cell transcriptomic atlas of the larval central nervous**
918 **system colored by structure, cluster, dissection, sample, and developmental**
919 **stage.** The complete atlas, described in Figures 4 and S4, colored by key
920 characteristics of the dataset. The first provides an overall structure of the cell classes
921 and where they are found in t-SNE space. The cluster coloring provides information
922 about location for each of the 70 clusters. The dissection splits the data into how the
923 data was collected – whether the sample contained only brain, only VNC, or both. The
924 sample t-SNE provides information about individual 10X genomics samples. The
925 developmental stage provides information about the age of the larvae at collection.
926
927
928
929
930



931

932

933 **Figure S6. Complete nervous system atlas from an individual animal.**

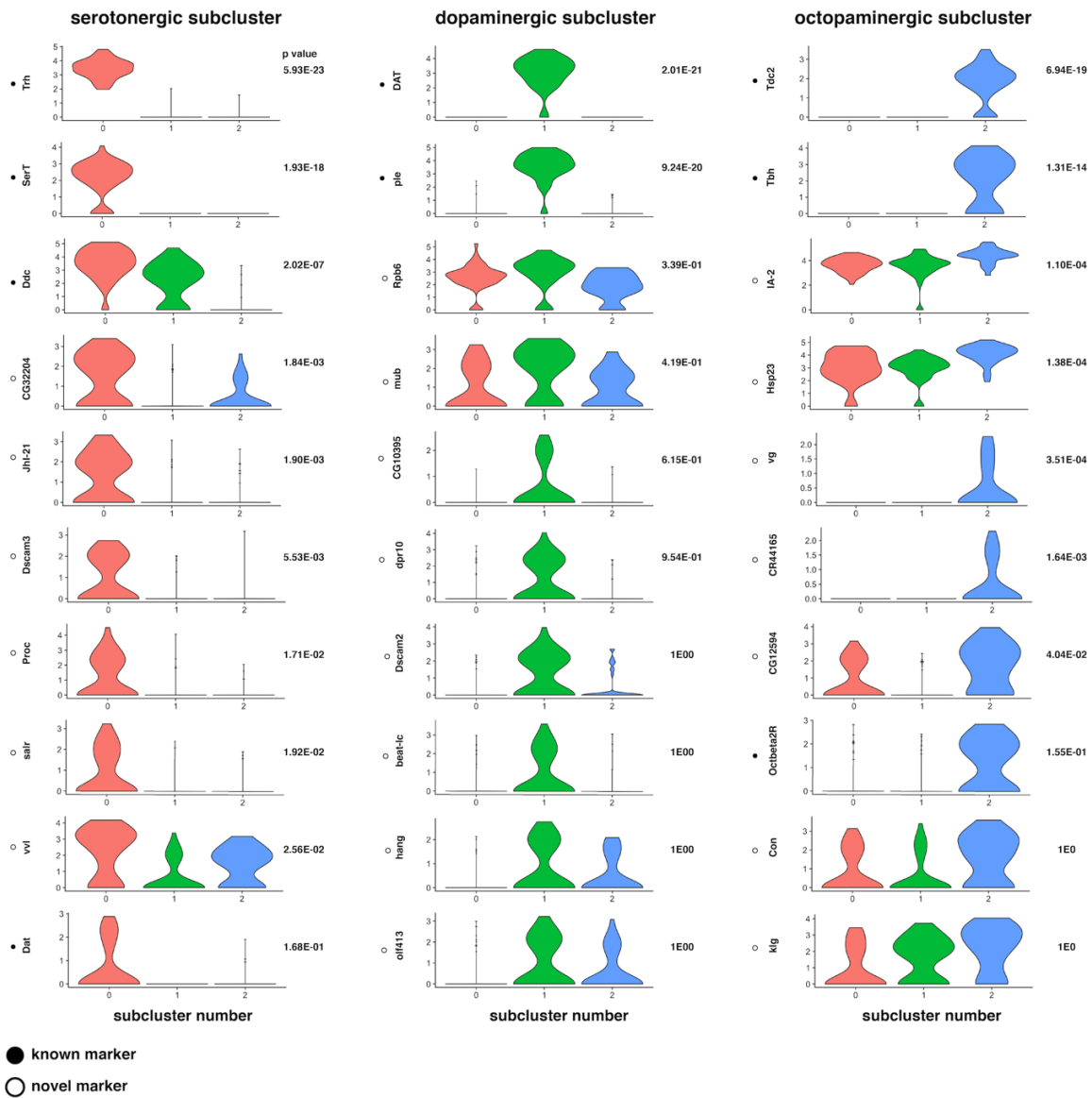
934 A. t-SNE of the complete single nervous system with cell clusters colored by gene
 935 expression of the top genes in each cluster. For each cluster in A there is a combination
 936 of genes which separate the clusters into recognizable molecular cell types and cell
 937 classes. B. Lines can be drawn in the t-SNE space that separates each of the cell
 938 classes we define here (adult developing neurons, functional larval neurons, neural
 939 stem cells, and glia). C. Violin plot of characteristic genes which separate each of the 8
 940 top level clusters.

941

942

943

944



945

946

947

948

949

950

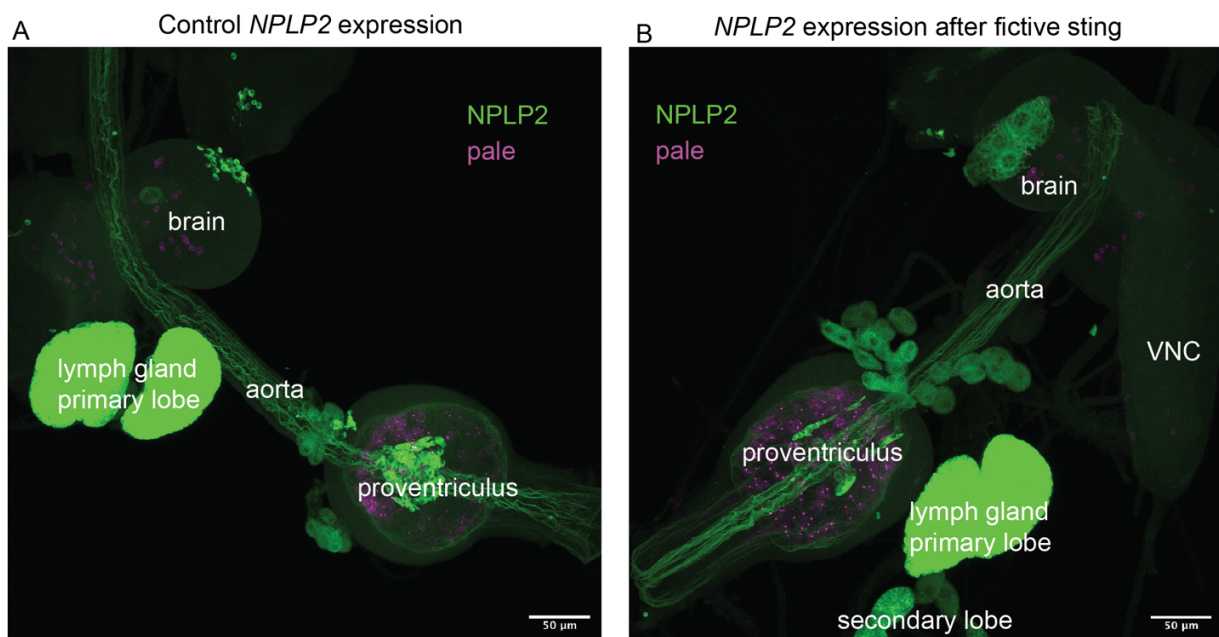
951

952

Figure S7. Unsupervised clustering separates recognizable serotonergic, dopaminergic, octopaminergic neurons into subclusters with known validated markers and novel markers. A cluster of cells was discovered using unsupervised machine learning techniques with markers indicative of dopaminergic cells. This cluster was separated and clustered once more, revealing three separate clusters with gene markers indicating a serotonergic subcluster, a dopaminergic subcluster, and an octopaminergic subcluster. The top genes that separate these clusters from one

953 another were then computed and violin plots were generated. Known markers were the
954 top genes for each subcluster and gave them recognizable identity – for example,
955 tryptophan hydroxylase (Trh), which is used to synthesize serotonin from tryptophan,
956 and serotonin transporter (SerT) were the top genes for the first subcluster, which could
957 then be appropriately labeled containing serotonergic cells. In addition to known gene
958 markers, new makers were also discovered for each group of cells.

959



960

961 **Figure S8: *NPLP2* expression is increased in the proventriculus and the brain**
962 **after fictive sting.** An RNA-FISH probe was designed for *NPLP2* and *pale* (*ple*) to
963 investigate the strong increased signal of *NPLP2* following the fictive sting. In the brain
964 lobes, the size of the *NPLP2*-positive cells is much larger. In the proventriculus, where
965 immune cells emerge, there appear to be more *NPLP2* positive cells emerging in the
966 case after fictive sting, suggesting an immune reaction related to the activation of
967 neurons alone.

968

969 **Table S1. Atlas of cell types in the complete larval nervous system.**

970 The top genes for the 70 cell type clusters are provided with the mean expression inside
971 the cluster, mean expression outside the cluster, and p-value. These clusters
972 correspond to Figure 4 and Supplementary Figure S4.

973

974 **Table S2. Gene modules that characterize neural precursor cells over**
975 **development.** Differential gene expression was used to compute gene modules in the
976 Monocle3 R package. The modules characterize early, intermediate, and late neural
977 precursor cell populations (Figure 3).

978

979 **Table S3. Fictive sting full nervous system atlas.**

980 The top genes for the 14 cell type clusters obtained from animals that were fictively
981 stung and controls. Tables include the top genes for each cluster, the mean expression
982 inside the cluster, the mean expression outside the cluster, and the p-value. These
983 clusters correspond to Figure 7E-F.

984

985 **Table S4. KC overactivation full nervous system atlas.**

986 The top genes for the 14 cell type clusters obtained from animals that had KCs
987 repeatedly activated and controls. Tables include the top genes for each cluster, the
988 mean expression inside the cluster, the mean expression outside the cluster, and the p-
989 value. These clusters correspond to Figure 7G-I.

990

991 **Table S5: RNA-FISH probe sequences.** RNA sequences of probes built to label *AstC-*
992 *R2*, *Hug*, and *NPLP2* transcripts.

993 **Movie S1. Z-stacks through RNA-FISH of AstC in insulin-producing cells.**

994 Z stacks are shown through the confocal stack of a brain in which the IPCs are labeled
995 with a fluorescent HaloTag ligand (Magenta), and *AstC-R2* mRNA is detected by FISH
996 (green), bar 10 μ m. This movie corresponds to Figure 5 in the main text.

997

998 **Movie S2. Single-molecule imaging to correlate between single-molecule FISH**
999 **and scRNAseq.**

1000 Cells were identified with coexpression of AstC (green) and vGlut (magenta) in the
1001 whole brain. This movie corresponds to Figure 6 in the main text and shows a view
1002 through z-stacks of BB-SIM images of AstC and vGlut mRNA FISH channels.

1003

1004 **References:**

1005 Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M., & Keller, P. J. (2013). Whole-brain
1006 functional imaging at cellular resolution using light-sheet microscopy. *Nature*
1007 *methods*, 10(5), 413.

1008

1009 Avalos, C. B., Maier, G. L., Bruggmann, R., & Sprecher, S. G. (2019). Single cell
1010 transcriptome atlas of the Drosophila larval brain. *eLife*, 8.

1011

1012 Bakker, K. (1959). Feeding period, growth, and pupation in larvae of *Drosophila*
1013 *melanogaster*. *Entomologia Experimentalis et Applicata*, 2(3), 171-186.

1014

1015 Berck ME, Khandelwal A, Claus L, Hernandez-Nunez L, Si G, Tabone CJ, Li F, Truman
1016 JW, Fetter RD, Louis M et al.: The wiring diagram of a glomerular olfactory system. *Elife*
1017 2016, 5.

1018

1019 Borst, A. (2014). Fly visual course control: behaviour, algorithms and circuits. *Nature*
1020 *Reviews Neuroscience*, 15(9), 590.

1021

1022 Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., ... & Trapnell, C.
1023 (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature*,
1024 566(7745), 496.

1025

1026 Chhetri, R. K., Amat, F., Wan, Y., Höckendorf, B., Lemon, W. C., & Keller, P. J. (2015).
1027 Whole-animal functional and developmental imaging with isotropic spatial
1028 resolution. *Nature methods*, 12(12), 1171.

1029

1030 Croset, V., Treiber, C. D., & Waddell, S. (2018). Cellular diversity in the *Drosophila*
1031 midbrain revealed by single-cell transcriptomics. *Elife*, 7, e34550.

1032

1033 Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., ... &
1034 Poovathingal, S. (2018). A single-cell transcriptome atlas of the aging *Drosophila*
1035 brain. *Cell*, 174(4), 982-998.

1036

- 1037 Diaz, M.M. et al. Allatostatin-C/AstC-R2 Is a Novel Pathway to Modulate the Circadian
1038 Activity Pattern in *Drosophila*. *Curr. Biol.* 29, 13-22 (2019)
1039
- 1040 Doe, C. Q. (2017). Temporal patterning in the *Drosophila* CNS. *Annual review of cell
1041 and developmental biology*, 33, 219-240.
1042
- 1043 Eichler, K, et al. "The complete connectome of a learning and memory centre in an
1044 insect brain." *Nature* 548.7666 (2017): 175.
1045
- 1046 Eschbach, C., Fushiki, A., Winding, M., Schneider-Mizell, C., Shao, M., Arruda, R., ... &
1047 Zlatic, M. (2019). Multilevel feedback architecture for adaptive regulation of learning in
1048 the insect brain. *bioRxiv*, 649731.
1049
- 1050 Etheredge, J. (2017). *Transcriptional profiling of Drosophila larval ventral nervous
1051 system hemilineages* (Doctoral dissertation, University of Cambridge).
1052
- 1053 Femino, A. M., Fay, F. S., Fogarty, K., & Singer, R. H. (1998). Visualization of single
1054 RNA transcripts in situ. *Science*, 280(5363), 585-590.
1055
- 1056 Fontana, J. R., & Crews, S. T. (2012). Transcriptome analysis of *Drosophila* CNS
1057 midline cells reveals diverse peptidergic properties and a role for castor in neuronal
1058 differentiation. *Developmental biology*, 372(1), 131-142.
1059
- 1060 Fossett, N., & Schulz, R. A. (2001). Functional conservation of hematopoietic factors in
1061 *Drosophila* and vertebrates. *Differentiation*, 69(2-3), 83-90.
1062
- 1063 Freeman, M. R. (2015). *Drosophila* central nervous system glia. *Cold Spring Harbor
1064 perspectives in biology*, 7(11), a020552.
1065
- 1066 Garcia-Campmany, L., Stam, F. J., & Goulding, M. (2010). From circuits to behaviour:
1067 motor networks in vertebrates. *Current opinion in neurobiology*, 20(1), 116-125.
1068
- 1069 Gilbert, L. I. (2004). Halloween genes encode P450 enzymes that mediate steroid
1070 hormone biosynthesis in *Drosophila melanogaster*. *Molecular and cellular
1071 endocrinology*, 215(1-2), 1-10.
1072
- 1073 Grimm, J. B., Muthusamy, A. K., Liang, Y., Brown, T. A., Lemon, W. C., Patel, R., ... &
1074 Lavis, L. D. (2017). A general method to fine-tune fluorophores for live-cell and in vivo
1075 imaging. *Nature methods*, 14(10), 987.
1076
- 1077 Haghverdi, L., Lun, A. T., Morgan, M. D., & Marioni, J. C. (2018). Batch effects in single-
1078 cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature
1079 biotechnology*, 36(5), 421.
1080
- 1081 Hartenstein, V., and Campos-Ortega, J.A. (1984). Early neurogenesis in wild-type
1082 *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* 193, 308–325.

1083
1084 Hartenstein, V., Rudloff, E., and Campos-Ortega, J.A. (1987). The pattern of
1085 proliferation of the neuroblasts in the wild-type embryo of *Drosophila melanogaster*.
1086 Wilhelm Roux's Arch Dev Bio 198, 264–274.
1087
1088 Helmstaedter M, Briggman KL, Turaga SC, Jain V, Seung HS, Denk W: Connectomic
1089 reconstruction of the inner plexiform layer in the mouse retina. Nature 2013, 500:168-
1090 174.
1091
1092 Hentze, J. L., Carlsson, M. A., Kondo, S., Nässel, D. R., & Rewitz, K. F. (2015). The
1093 neuropeptide allatostatin A regulates metabolism and feeding decisions in
1094 *Drosophila*. *Scientific reports*, 5, 11680.
1095
1096 Hildebrand DGC, Cicconet M, Torres RM, Choi W, Quan TM, Moon J, Wetzel AW, Scott
1097 Champion A, Graham BJ, Randlett O et al.: Whole-brain serial-section electron
1098 microscopy in larval zebrafish. Nature 2017, 545:345-349.
1099
1100 Hwang, R. Y., Zhong, L., Xu, Y., Johnson, T., Zhang, F., Deisseroth, K., & Tracey, W.
1101 D. (2007). Nociceptive neurons protect *Drosophila* larvae from parasitoid
1102 wasps. *Current Biology*, 17(24), 2105-2116.
1103
1104 Jarrell TA, Wang Y, Bloniarz AE, Brittin CA, Xu M, Thomson JN, Albertson DG, Hall DH,
1105 Emmons SW: The connectome of a decision-making neural network. Science 2012,
1106 337:437-444.
1107
1108 Jovanic, T., Masson, J. B., Truman, J. W., & Zlatic, M. (2017). Mapping neurons and
1109 brain regions underlying sensorimotor decisions and sequences. *bioRxiv*, 215236.
1110
1111 Kater, S. B., Murphy, A. D., & Rued, J. R. (1978). Control of the salivary glands of
1112 *Helisoma* by identified neurones. *Journal of Experimental Biology*, 72(1), 91-106.
1113
1114 Kater, S. B., Rued, J. R., & Murphy, A. D. (1978). Propagation of action potentials
1115 through electrotonic junctions in the salivary glands of the pulmonate mollusc, *Helisoma*
1116 *trivolvis*. *Journal of Experimental Biology*, 72(1), 77-90.
1117
1118 Konopka, R. J., & Benzer, S. (1971). Clock mutants of *Drosophila*
1119 *melanogaster*. *Proceedings of the National Academy of Sciences*, 68(9), 2112-2116.
1120
1121 Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R., & Desplan, C.
1122 (2018). Phenotypic convergence: distinct transcription factors regulate common terminal
1123 features. *Cell*, 174(3), 622-635.
1124
1125 Kraaijeveld, A. R., & Godfray, H. C. J. (1997). Trade-off between parasitoid resistance
1126 and larval competitive ability in *Drosophila melanogaster*. *Nature*, 389(6648), 278.
1127

- 1128 Krashes, M. J., DasGupta, S., Vreede, A., White, B., Armstrong, J. D., & Waddell, S.
1129 (2009). A neural circuit mechanism integrating motivational state with memory
1130 expression in *Drosophila*. *Cell*, 139(2), 416-427.
1131
- 1132 Lee, H. G., Kim, Y. C., Dunning, J. S., & Han, K. A. (2008). Recurring ethanol exposure
1133 induces disinhibited courtship in *Drosophila*. *PLoS One*, 3(1), e1391.
1134
- 1135 Lemon, W.C., et al. "Whole-central nervous system functional imaging in larval
1136 *Drosophila*." *Nature communications* 6 (2015).
1137
- 1138 Li, H. H., Kroll, J. R., Lennox, S. M., Ogundeyi, O., Jeter, J., Depasquale, G., & Truman,
1139 J. W. (2014). A GAL4 driver resource for developmental and behavioral studies on the
1140 larval CNS of *Drosophila*. *Cell reports*, 8(3), 897-908.
1141
- 1142 Lionnet, T. et al. A transgenic mouse for in vivo detection of endogenous labeled
1143 mRNA. *Nat Methods* 8, 165-70 (2011).
1144
- 1145 Liu, Z., Yang, C. P., Sugino, K., Fu, C. C., Liu, L. Y., Yao, X., ... & Lee, T. (2015).
1146 Opposing intrinsic temporal gradients guide neural stem cell production of varied
1147 neuronal fates. *Science*, 350(6258), 317-320.
1148
- 1149 Long, X., Colonell, J., Wong, A. M., Singer, R. H., & Lionnet, T. (2017). Quantitative
1150 mRNA imaging throughout the entire *Drosophila* brain. *nature methods*, 14(7), 703.
1151
- 1152 Luo, J., Becnel, J., Nichols, C. D., & Nässel, D. R. (2012). Insulin-producing cells in the
1153 brain of adult *Drosophila* are regulated by the serotonin 5-HT 1A receptor. *Cellular and*
1154 *Molecular Life Sciences*, 69(3), 471-484.
1155
- 1156 Maaten, L. V. D., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine*
1157 *learning research*, 9(Nov), 2579-2605.
1158
- 1159 McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and
1160 projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.
1161
- 1162 Meissner, G.W. et al. Mapping Neurotransmitter Identity in the Whole-Mount *Drosophila*
1163 Brain Using Multiplex High-Throughput Fluorescence in Situ Hybridization. *GENETICS*
1164 211:437-482
1165
- 1166 Ohyama T, Schneider-Mizell CM, Fetter RD, Aleman JV, Franconville R, Rivera-Alba M,
1167 Mensh BD, Branson KM, Simpson JH, Truman JW et al.: A multilevel multimodal circuit
1168 enhances action selection in *Drosophila*. *Nature* 2015, 520: 633-639.
1169
- 1170 Pliner, H. A., Shendure, J., & Trapnell, C. (2019). Supervised classification enables
1171 rapid annotation of cell atlases. *BioRxiv*, 538652.
1172

- 1173 Prevedel, R., Yoon, Y. G., Hoffmann, M., Pak, N., Wetzstein, G., Kato, S., ... & Vaziri, A.
1174 (2014). Simultaneous whole-animal 3D imaging of neuronal activity using light-field
1175 microscopy. *Nature methods*, 11(7), 727.
1176
- 1177 Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., & Trapnell, C. (2017).
1178 Reversed graph embedding resolves complex single-cell trajectories. *Nature*
1179 *methods*, 14(10), 979.
1180
- 1181 Ro, J., Pak, G., Malec, P. A., Lyu, Y., Allison, D. B., Kennedy, R. T., & Pletcher, S. D.
1182 (2016). Serotonin signaling mediates protein valuation and aging. *Elife*, 5, e16843.
1183
- 1184 Robie, A. A., Hirokawa, J., Edwards, A. W., Umayam, L. A., Lee, A., Phillips, M. L., ... &
1185 Reiser, M. B. (2017). Mapping the neural substrates of behavior. *Cell*, 170(2), 393-406.
1186
- 1187 Rulifson, E. J., Kim, S. K., & Nusse, R. (2002). Ablation of insulin-producing neurons in
1188 flies: growth and diabetic phenotypes. *Science*, 296(5570), 1118-1120.
1189
- 1190 Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial
1191 reconstruction of single-cell gene expression data. *Nature biotechnology*, 33(5), 495.
1192
- 1193 Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat*
1194 *Methods* 9, 676-82 (2012).
1195
- 1196 Schlegel, P., Texada, M. J., Miroschnikow, A., Schoofs, A., Hückesfeld, S., Peters, M.,
1197 ... & Truman, J. W. (2016). Synaptic transmission parallels neuromodulation in a central
1198 food-intake circuit. *Elife*, 5, e16799.
1199
- 1200 Schwaerzel, Martin, et al. "Dopamine and octopamine differentiate between aversive
1201 and appetitive olfactory memories in *Drosophila*." *Journal of Neuroscience* 23.33 (2003):
1202 10495-10502.
1203
- 1204 Scott, K., Brady Jr, R., Cravchik, A., Morozov, P., Rzhetsky, A., Zuker, C., & Axel, R.
1205 (2001). A chemosensory gene family encoding candidate gustatory and olfactory
1206 receptors in *Drosophila*. *Cell*, 104(5), 661-673.
1207
- 1208 Selcho, Mareike, et al. "The role of dopamine in *Drosophila* larval classical olfactory
1209 conditioning." *PloS one* 4.6 (2009): e5897.
1210
- 1211 Sokolowski, M. B. (2001). *Drosophila*: genetics meets behaviour. *Nature Reviews*
1212 *Genetics*, 2(11), 879.
1213
- 1214 Sorrentino, R. P., Carton, Y., & Govind, S. (2002). Cellular immune response to parasite
1215 infection in the *Drosophila* lymph gland is developmentally regulated. *Developmental*
1216 *biology*, 243(1), 65-80.
1217

- 1218 Takemura SY, Bharioke A, Lu Z, Nern A, Vitaladevuni S, Rivlin PK, Katz WT, Olbris DJ,
1219 Plaza SM, Winston P et al.: A visual motion detection circuit suggested by *Drosophila*
1220 connectomics. *Nature* 2013, 500:175-181.
1221
1222 Trannoy, S., & Kravitz, E. A. (2017). Strategy changes in subsequent fights as
1223 consequences of winning and losing in fruit fly fights. *Fly*, 11(2), 129-138.
1224
1225 Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., ... & Rinn, J.
1226 L. (2014). The dynamics and regulators of cell fate decisions are revealed by
1227 pseudotemporal ordering of single cells. *Nature biotechnology*, 32(4), 381.
1228
1229 Truman, J. W. (2005). Hormonal control of insect ecdysis: endocrine cascades for
1230 coordinating behavior with physiology. *Vitamins & Hormones*, 73, 1-30.
1231
1232 Truman, J. W., & Bate, M. (1988). Spatial and temporal patterns of neurogenesis in the
1233 central nervous system of *Drosophila melanogaster*. *Developmental biology*, 125(1),
1234 145-157.
1235
1236 Truman, J.W., Taylor, B.J., and Awad, T.A. (1993). Formation of the adult nervous
1237 system. In *The Development of Drosophila melanogaster Vol II*, M. Bate and A.M. Arias,
1238 eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 1245–1275.
1239
1240 Vladimirov, N., Wang, C., Höckendorf, B., Pujala, A., Tanimoto, M., Mu, Y., ... &
1241 Koyama, M. (2018). Brain-wide circuit interrogation at the cellular level guided by online
1242 analysis of neuronal function. *Nature methods*, 15(12), 1117.
1243
1244 Vogelstein, J. T., Park, Y., Ohshima, T., Kerr, R. A., Truman, J. W., Priebe, C. E., &
1245 Zlatic, M. (2014). Discovery of brainwide neural-behavioral maps via multiscale
1246 unsupervised structure learning. *Science*, 344(6182), 386-392.
1247
1248 White JG, Southgate E, Thomson JN, Brenner S: The structure of the nervous system
1249 of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci* 1986,
1250 314:1-340.
1251
1252 Williamson, M., Lenz, C., Winther, M. E., Nässel, D. R., & Grimmelikhuijzen, C. J.
1253 (2001). Molecular cloning, genomic organization, and expression of a B-type (cricket-
1254 type) allatostatin preprohormone from *Drosophila melanogaster*. *Biochemical and*
1255 *biophysical research communications*, 281(2), 544-550.
1256
1257 Ziegler, A. B., Augustin, H., Clark, N. L., Berthelot-Grosjean, M., Simonnet, M. M.,
1258 Steinert, J. R., ... & Grosjean, Y. (2016). The amino acid transporter Jhl-21 coevolves
1259 with glutamate receptors, impacts NMJ physiology, and influences locomotor activity in
1260 *Drosophila* larvae. *Scientific reports*, 6, 19692.
1261

1262 Ziegler, A. B., Manière, G., & Grosjean, Y. (2018). Jhl-21 plays a role in *Drosophila*
1263 insulin-like peptide release from larval IPCs via leucine transport. *Scientific reports*, 8(1),
1264 1908.