Single-nuclei transcriptomes in the hypothalamus and POA of a highly social cichlid

Isaac Miller-Crews[1,2] and Hans A. Hofmann[1]

1: Department of Integrative Biology, The University of Texas at Austin, Austin, TX 78712

2: Department of Biology, Indiana University, Bloomington, IN, USA

Corresponding author: imillerc@iu.edu or hans@utexas.edu

Keywords:

ABSTRACT

How does complex behavior arise from the genome across biological levels? Specific cell types have evolutionarily conserved functional roles in regulating and maintaining various aspects of social behavior. There is a long history of research demonstrating how specific cells in the hypothalamus and POA are critical in regulating social status and reproductive behavior. This cell function is dependent not just on one gene, but the entire gene expression network working within each cell. The brain is a heterogenous tissue made up of a broad diversity of cells. Targeting the cellular-level transcriptomes in socially relevant brain regions enables the identification of social status mediators. Single-nuclei RNA-sequencing (snRNA-seq) provides the resolution to identify cell-specific transcriptomes, with more practical benefits compared to single-cell approaches. Single-nucleus sequencing of neuronal tissue has remained fairly limited to model mammalian organisms. Here we performed snRNA-seq in the hypothalamus and POA of *Astatotilapia burtoni*, a fish system ideal for social neuroscience. Male *A. burtoni* are characterized by plastic phenotypes associated with social status, dominant and subordinate, with well-established differences in neural gene expression patterns and specific neuron morphology. This provides a novel opportunity to apply snRNA-seq in a complex tissue of a non-model system using hypothesis driven approaches. We show how changes in social status are linked to distinct transcriptomic profiles at the cellular level.

32 **INTRODUCTION**

33 Across vertebrates, the hypothalamus and preoptic area (POA) serve as key neural nodes
34 in modulating social behavior across vertebrates, among many other functions (Hull &
35 Dominguez, 2005). Considerable evidence suggests that the fundamental neural mechanisms that
36 regulate social behavior in the POA-hypothalamus are evolutionarily conserved (for review, see
37 (O'Connell & Hofmann, 2012a)). The hypothalamus is highly complex neural region, with
38 multiple distinguishable subdivisions assigned by neuron morphology and function. Structure-
39 function type studies are difficult to perform in the hypothalamus due to the high degree of
40 cellular heterogeneity of functionally distinct neurons compared to other brain regions (e.x.
41 hippocampus, neocortex, cerebellum, etc.) (Romanov et al., 2017). How do these systems work
42 together within and between cells to encode and regulate complex social behavior?

43 The function of a cell is dependent not just on one gene, but the entire gene expression
44 network working within each cell. This complexity is compounded when studying a tissue such
45 as the brain which is made up of a wide array of cells. Specific neural cell types have
46 evolutionarily conserved functional roles in regulating and maintaining various aspects of social
47 behavior. To resolve the complexity of neuronal tissue, we can use methods such as single-
48 nucleus RNA sequencing (snRNA-seq) to target cellular-level transcriptomes (Ding et al., 2020).
49 To date, snRNA-seq of neuronal tissue has largely been confined to standard model mammalian
50 organisms (ref).

51 Single-cell based approaches provide the resolution to identify cell-specific
52 transcriptomes and provides the mechanism to classify cell types by transcriptomic features
53 (Grün & Van Oudenaarden, 2015; Zeng et al., 2018). Other common transcriptomic techniques,
54 such as RNA-seq, give a read-out averaged across cells that will cover up any cellular

heterogeneity and will possibly mask any biological effect that occurs in a limited number of cells or from lowly expressed transcripts (Nguyen et al., 2018; Ofengeim et al., 2017). The development of single-cell RNA-sequencing aims to overcome these limitations and provide greater variability and resolution (Grindberg et al., 2013; Ofengeim et al., 2017; Tang et al., 2009).

Cells that appear to be similar morphologically can have distinct patterns of gene expression (Moffitt et al., 2018). The first example of single-cell RNA gene expression analysis started in the early 1990's (Eberwine et al., 1992), two decades before the first next-generation sequencing in mammalian cells could assess the entire transcriptome (Tang et al., 2009). Single-cell sequencing can provide insight into transcriptional response to stimuli, internal or external (Wang & Navin, 2015). In fact, single-cell sequencing has been used to study cell activation, through immediate early gene detection (Wu et al., 2017). Analysis of gene-expression at the cellular-level grants the capability to understand how specific genes change expression patterns in conjunction with alterations perceived at the phenotypic-level. These genes can then be assessed within the context of the cell with single-cell sequencing capturing the entire transcriptome, particularly when targeting genes with known functional roles, such as neuropeptides.

Neuropeptides are messenger proteins synthesized from larger precursor proteins, preprohormones, that once released from neurons may act on a variety of cells which contain the appropriate receptor. Neuropeptides, although small, are known to coordinate a plethora of behavioral and physiological processes (Smith et al., 2019). Social behavior in particular has repeatedly been shown to be regulated by evolutionarily conserved neuropeptides (e.g. AVP, OXT) (O'Connell & Hofmann, 2012a). A prior single-cell RNAseq experiment in mouse

hypothalamus demonstrated how 27 neuropeptide genes are limited to a few neuronal subtypes but co-expression of multiple neuropeptides is common in hypothalamic neurons, with some being broadly expressed across subtypes and others more specific (similar pattern found in neuropeptide receptors) (R. Chen et al., 2017). It remains unclear how consistent these co-expression networks are across individuals and how neuropeptide co-expression respond across treatment. Therefore, to understand the role that neuropeptides play in regulating social behavior the entirety of neuropeptide genes expression, the neuropeptidome, needs to be considered at the cellular level.

One of the hypothesized contributing factors to the flexibility of the hypothalamus adapting to environmental stimulus could be the common occurrence of co-expression of two or more neuropeptides within the same cells (Romanov et al., 2017). There have been dozens of cellular subtypes identified in the mouse hypothalamus, with predictions estimating that mammalian neuronal subtypes in the hypothalamus could contain as many as 100 subtypes (Romanov et al., 2017; Steuernagel et al., 2022). Studies on mouse hypothalamus found non-neuronal subtypes are similar to those in cerebral cortex, but that the neuronal cell types were highly specific to the hypothalamus, implying that differences in brain region functions are reflected with differences in neuronal composition (Zeng et al., 2018). Differentiation of these neuronal cellular subtypes typically rely on the high degree of co-expression among genes in a cellular network to assign subtypes by the expression of a singular neuropeptide (Crow & Gillis, 2018). Previous research in mouse hypothalamus goes so far as to use this reduction to a single neuropeptide genes as an annotation method on single-cell hierarchical UMAP statistical clusters by identifying specific cell-types by the expression of a single gene (ex. Pomc, Kiss1, Vip, Agrp) to identify neuronal subtypes that have been previously described by other techniques (Romanov et al., 2017). Yet, it

101 is not apparent how accurately these unsupervised statistical clusters reflect cellular subtypes or
102 the significance of neuropeptide co-expression. Due to technological constraints of techniques,
103 such as in ISH or IHC, these neuropeptides are usually studied in isolation. Therefore, we know
104 little about how neuropeptide systems interact at a cellular level.

105 Burton's Mouthbrooder cichlid, *Astatotilapia burtoni,* a highly social fish from East Africa's
106 Lake Tanganyika, has become an important model system in social neuroscience (Hofmann &
107 Fernald, 2003; Hofmann, 2006)(Maruska & Fernald, 2018). Males of this species are
108 characterized by plastic phenotypes associated with social dominance status, with well-
109 established differences in neural gene expression patterns and neuron morphology (for review,
110 see (Maruska & Fernald, 2013)). Several peptidergic systems are known to regulate social status
111 in *A. burtoni*. For example, both GnRH I and Somatostatin (SST) neurons are larger in size and
112 have increased expression in dominant over subordinate males (M. R. Davis & Fernald, 1990;
113 Hofmann & Fernald, 2000). Likewise, three morphological subpopulations of vasopressin (AVP)
114 neurons, parvo-, magno-, and giganto-cellular vary across social status (Greenwood et al., 2008).
115 Specifically, gigantocelluar AVP neurons have higher AVP expression in dominant males while
116 parvocellular AVP neurons show the opposite effect. Even though these cells are
117 morphologically distinct, and appear to have unique responses to social status, it remains
118 unknown what differences exist at other biological levels, such as the transcriptome. How do
119 these specific cell types regulate social status in the brain of *A. burtoni*?

120 In the present study, we use snRNA-seq to analyze the hypothalamus and POA of dominant
121 and subordinate male *A. burtoni*. We sought to investigate how changes in social status are
122 linked to distinct transcriptomic profiles at the cellular level. Due to the known functional role of
123 these brain regions in regulating dominance behavior, we hypothesized that neuronal gene co-

expression networks in neurons will be differentially regulated across social status and not just reflect neuronal cellular subtypes. Additionally, we predicted that morphologically-distinct, socially-responsive subpopulations of AVP neurons represented unique transcriptomic nuclei clusters that show differential expression patterns across social status at the subpopulation level rather than a ubiquitous change across all AVP expressing neurons. Given the role neuropeptides play in regulating behavioral changes, we first sought to characterize the neuropeptidome at the cellular level to understand how the variation across social states will be reflected, hypothesizing that a more complex and overexpressed neuropeptidome would occur in dominant males. Taken together, we can gain insight into the mechanisms by which these systems coordinate to build an integrated organismal phenotype.

METHODS

Animals

The animals used in this study came from a laboratory population of Burton's Mouthbrooder cichlid, *Astatotilapia burtoni*, descended for about 60 generations from a wild-caught stock of 400 individuals (Fernald & Hirata, 1977). All work was done in compliance with the Institutional Animal Care and Use Committee (IACUC) at The University of Texas at Austin.

Experimental design

We created six naturalistic communities of adult *A. burtoni* males and females within three 100-L aquaria, with each aquarium subdivided by a transparent, perforated divider to facilitate visual and chemical communication between neighboring communities. Each community consisted of one large dominant male (57-61 mm standard length) and one relatively smaller subordinate male (45-53 mm standard length), as well as three females (10-15 mm smaller than the large male). A terracotta pot served as territorial shelter and artificial aquarium plants provided a refuge for the subordinate male and females. We recorded social status and coloration of every male in each community as well as the presence of any mouthbrooding females every other day for 16 days to ensure that each male displayed a stable social phenotype and each dominant male had mated at least once. In this manner we could confirm the dominance status of every male and categorize each individual as dominant or subordinate over the entire observation period. After three weeks, we collected all males between 10:00h and 12:00h, measured their body mass and standard length, and harvested blood for subsequent hormone assays. Males were then euthanized via rapid cervical dissection and brains were

159 dissected within three minutes, embedded in OCT compound (Sakura Finetek, USA),
160 immediately frozen on dry ice, and stored at -80°C until further processing. Gonads were
161 dissected and weighed to obtain the gonadosomatic index, a proxy measure of reproductive state,
162 by dividing gonad mass by body mass. Blood samples were spun in a table centrifuge for five
163 min at 3000 rpm to separate the plasma, which was stored at -80°C until further processing.
164
165 *Hormone assay*
166 Circulating testosterone levels were assayed in duplicate from blood plasma by ELISA
167 (Cayman Chemical, USA) on a single assay plate following manufacturer's instructions (within-
168 plate coefficient of variation was 12%). Note that in contrast to many other teleost fishes,
169 testosterone is the main androgen in Haplochromine cichlids (Dijkstra et al., 2012; Kidd et al.,
170 2010).
171
172 *Tissue collection*
173 The POA and hypothalamus were micro-dissected following the protocol described in
174 (Maruska et al., 2013). Briefly, brains were sectioned into 300 µm slices on a cryostat microtome
175 onto positively-charged Superfrost Plus microscope slides (Thermo Fisher Scientific, USA),
176 which were then placed on a metal block chilled on dry ice for 10 min, or until temperature
177 reached approximately -4 to -10 °C, to allow the collection of brain tissue without fracturing
178 sections. Using a 0.75 mm tissue punch tool (Electron Microscopy Sciences, USA), the entirety
179 of the POA and hypothalamus were collected together into a 1.5 ml reaction tube (16-18 punches
180 per subordinate and 20-26 punches per dominant) and stored at -80°C until further processing.
181

182 *Nuclei isolation*

183 Neural tissue is challenging to dissociate into nuclei due to the heterogenous nature of the
184 tissue and unique interconnected structure of neurons. Fluorescence-activated cell sorting
185 (FACS) is a common method to isolate and purify single cells, for multiple parameters and cell-
186 types, but requires a starting volume larger than micro-dissected brain tissue punches can provide
187 (Grün & Van Oudenaarden, 2015). It has been shown that the transcriptomes of single cells and
188 single nuclei are highly correlated (Nguyen et al., 2018). We therefore opted to isolate single
189 nuclei from frozen tissue instead.

190 It should also be noted that statistical power of single-cell sequencing is largely
191 determined by the number of cells (nuclei) and sequencing complexity, rather than sequencing
192 depth (A. Davis et al., 2019; Grün & Van Oudenaarden, 2015; Schmid et al., 2021). Pooling
193 samples increases statistical power by increasing the number of nuclei while still capturing the
194 biological variation between samples. Pilot experiments indicated that pooling the entirety of the
195 POA and hypothalamus together for three males would be required to reach a minimum number
196 of 10,000 cells (or nuclei) per sample, as recommended for the 10X Genomics Chromium
197 platform (10x Genomics, USA). We therefore selected one dominant-subordinate pair from each
198 of our three separate aquaria and combined them into two separate pools for dominant and
199 subordinate males, respectively (1 pool per group, n=3/pool).

200 We adapted a density gradient protocol for nuclei isolation (Salem et al., 2024). Briefly,
201 all tissue samples were frozen or on ice throughout the protocol to prevent any additional
202 transcriptional activity. First, the tissue was mechanically homogenized to release the nuclei by
203 adding 1.5 ml of chilled Nuclei EZ Lysis Buffer (Sigma-Aldrich, USA) with 0.2 U/µl Protector
204 RNase Inhibitor (Roche, USA) to the tissue in a 2 mL tube and then gently homogenized 20

205 times in a chilled 7 ml Dounce homogenizer with pestle "b". Nuclei were then filtered through a
206 35 µm cell strainer cap from Falcon Round-Bottom Polystyrene Test Tubes with Cell Strainer
207 Snap Cap, 5mL (Corning, USA) into a 1.5 mL tube. Next, the nuclei were centrifuged at 900 g
208 for 5 min at 4°C before the supernatant was removed and the nuclei pellet resuspended in 750 µl
209 of 25% Optiprep (Sigma-Aldrich, USA) in nuclei resuspension buffer of 1x PBS with 2% BSA
210 and 0.2 U/µl Protector RNase Inhibitor (Roche, USA). Resuspended nuclei were then overlaid
211 on 750 µl of 29% Optiprep in nuclei resuspension buffer before centrifuging at 13,000 g for 30
212 min at 4°C. Afterwards, the supernatant was carefully removed, and the pellet quickly
213 resuspended with 20 µl nuclei resuspension buffer and transferred to a new 0.5 ml tube with 50
214 µl of nuclei resuspension buffer added. Nuclei were then filtered through a 35 µm cell strainer
215 cap (Corning, USA). Finally, the concentration of nuclei was quantified by using two 5 µl
216 aliquots with 5 µl resuspension buffer with 10 ug/mL DAPI on the Countess 3 Automated Cell
217 Counter (Invitrogen, USA) to aim for a concentration between 700-1200 nuclei/µl. The samples
218 were then submitted to UT Austin's Genomic Sequencing & Analysis Facility for quality control
219 and library construction using the 10X Genomics Chromium Next GEM Single Cell 3' reagent
220 kit v3.1 as per manufacturer's instructions. The libraries were then sequenced on the Illumina
221 NovaSeq S1platform.
222
223 *Processing of single nucleus data*
224 We obtained 926 million reads total (dominant samples pool: 460 million reads;
225 subordinate samples pool: 466 million reads). We used the MultiQC package (Ewels et al., 2016)
226 to confirm that all reads had Phred scores above 35 and passed all quality control measures. Read
227 mapping and annotation were implemented using the Cell Ranger 6.1.2 pipeline (Zheng et al.,

228 2017) and the Nile tilapia reference genome (O_niloticus_UMD_NMBU) (Conte et al., 2017).
229 Data was then processed with the Seurat integrated analysis pipeline (Satija et al., 2015) and
230 normalized with SCTransform version 2 (Choudhary & Satija, 2022). The cell by gene
231 expression matrices for each pooled sample dataset were filtered by selecting genes expressed in
232 a minimum of 3 cells and selecting nuclei with 1000-2000 detected genes (nFeature_RNA). This
233 threshold was if detected genes was determined from the distribution of nuclei to reduce
234 background and reduce doublets. We then clustered the nuclei using the standard Seurat pipeline,
235 we utilized Clustree (Zappia & Oshlack, 2018) to determine an appropriate resolution for
236 clustering by selecting a resolution that yielded the most stable number of clusters (0.8). This
237 resulted in a total of 15,392 and 12,514 single-nucleus transcriptomes for the dominant and
238 subordinate sample pool respectively.
239
240 *Genotype demultiplexing*
241 Even though our *A. burtoni* laboratory population is not fully outbred (Brawand et al.,
242 2014; Miller-Crews et al., 2021), we expected there to be sufficient genotypic variation to allow
243 demultiplexing of pooled samples by genotype, thereby recovering biological replication for
244 downstream statistical analyses. Genotypic variation in our *A. burtoni* populations allowed for
245 demultiplexing of the pooled samples using genotype of each sample. We used the souporcell
246 package (Heaton et al., 2020) for this purpose, as it does not require the genotype of each
247 individual in a pooled samples to be known. Briefly, the pipeline performs variant detection on
248 the sequence bam files to count allele support for each nucleus. Next, the nuclei are clustered
249 with a sparse mixture model to assign cells to genotypes. Together, this allows for
250 demultiplexing in the absence of prior genotype information of each sample

251

252 *Cell type annotation*

253 The lack of standardized marker genes as well as the possible presence of multiple gene
254 paralogs complicate cell type identification in non-traditional model systems. A list of known
255 cell type marker genes was obtained from the mammalian literature (Supplemental Table 1) and
256 another marker list was obtained from HypoMap (C7 markers) (Steuernagel et al., 2022), a
257 murine hypothalamus single-cell expression atlas from an 18 study meta-analysis. We
258 determined Nile Tilpaia orthologs of these genes with ensembl. We tested a few different
259 methods to identify distinct cell type classes including: Seurat's Module score (Satija et al.,
260 2015), scsorter (Guo & Li, 2021), and sctype (Ianevski et al., 2022). To best characterize nuclei
261 with unknown cell types, and to robustly identify cell types, we used the combined set of cell
262 marker gene orthologs and performed cell-type identification with sctype, an ultra-fast and
263 automated cluster-based method. Importantly, sctype default assigns a cell-type to a cluster if the
264 top positive scaled cell-type score for the entire cluster is greater than 25% the number of cells in
265 the cluster. Due to differences in the number of orthologous marker genes for each cell-type
266 marker from HypoMap, these scores can be artificially low for certain cell types and particularly
267 impacts neurons. To account for this, if a cell-type score was positive for a certain cluster and
268 above a 10% threshold the cell-type was assigned. Only three nuclei clusters (clusters 3, 7, and
269 10) were impacted by this lower threshold. Otherwise, clusters were assigned as unknown. We
270 successfully recovered all major cell types previously identified in the hypothalamus.
271 Importantly, this approach allowed us to reduce noise by characterizing nuclei with unknown cell
272 types that did not get assigned to any specific cell category.

273

274 *Differential gene expression and WGCNA*

275 We further subclustered each cell type, mediating identification of subtypes, using
276 UMAP and Clustree (Zappia & Oshlack, 2018) to determine the appropriate number of sub-
277 clusters. Next, we used limmatrend (Law et al., 2014; Soneson & Robinson, 2018), incorporating
278 the genotype data to separate out pooled samples and add statistical power, to identify
279 differentially expressed genes (DEGs) across social status for each subtype and across all nuclei
280 of each cell type.

281 Nuclei identified as GABAergic ("inhibitory") and glutamatergic ("excitatory") were
282 merged into the broader neuron category. Due to the higher degree of complexity of neurons,
283 indicated by the presence of multiple sub-clusters, these neuron nuclei were subsequently subset
284 out of the data set and re-clustered with UMAP into various sub-populations of neurons. Next,
285 we performed a Weighted Gene Co-expression Network Analysis (WGCNA) approach adapted
286 for single-nuclei data, hdWGCNA (Morabito et al., 2023), to identify gene co-expression
287 modules. Compared to traditional WGCNA designed for analysis of samples (Langfelder &
288 Horvath, 2008), hdWGCNA was developed specifically for use with single-cell and spatial
289 transcriptomics data by generating metacells from selected pools of nuclei to reduce background
290 noise of typically noisy single-nuclei transcriptomes. Metacells were pooled from individual
291 samples using the genotyping data and with a minimum cell count parameter of 50. We used
292 limmatrend (Law et al., 2014; Soneson & Robinson, 2018) to identify differentially expressed
293 genes (DEGs) and differentially expressed module eigengenes (deMEGs) for these neuron
294 populations between dominant and subordinate groups using individual genotypes. GO terms for
295 each gene was assigned through homology to mouse genes with Ensembl (Martin et al., 2023).
296 For each hdWGCNA module, gene ontology enrichment analysis was performed with GO-

297 MWU, which adaptively clusters GO categories and applies a Mann–Whitney U tests using
298 module kME value (GO-MWU, https://github.com/z0on/GO_MWU) (Wright et al., 2015).
299
300 *Spatial location prediction*
301 The hypothalamus represents a brain region comprised of multiple distinct sub-
302 populations spatial distributed with known functional differences. To assign putative spatial
303 location to neuron clusters in the *A. burtoni*, we developed a method for comparing these clusters
304 to mouse hypothalamus neuron clusters with known or imputed spatial information. Predicting
305 cluster specific spatial location relied on the HypoMap database (Steuernagel et al., 2022), which
306 uses a combination of the ISH expression values from the Allen Brain Atlas (Lein et al., 2007)
307 and described region of origin from the original dataset. Selected neuron cluster marker genes
308 compared to all other neuron clusters were assigned with Seurat. A specificity score, as described
309 in Steuernagal et al, 2022, was calculated for each marker gene. Briefly, specificity was
310 calculated as the average expression foldchange multiplied by the ratio of the percentage of
311 expressing nuclei in the cluster over the percentage of expressing nuclei in all neurons for each
312 marker gene of a given cluster. Additionally, HypoMap cluster marker genes for neurons in
313 clusters 1-50 of C66 were assigned Nile Tilapia orthologs with Ensembl. There only remained a
314 limited number of overlapping orthologous genes between the marker gene sets for neuron
315 clusters and HypoMap C66 clusters. To compare marker genes from a specific neuron cluster
316 with a HypoMap neuron cluster, we took genes that overlapped between the two lists and took
317 the sum of the specificity score for all the overlapping neuron markers and the HypoMap neuron
318 cluster markers separately. Then these two specificity sums, one from the neuron cluster and one
319 from the HypoMap neuron cluster, were multiplied together. If a gene, or couple of genes, has

320 high specificity in both the neuron cluster and the HypoMap cluster it is being compared to, the
321 score will be higher. So, each pairwise comparison of neuron cluster and HypoMap cluster has a
322 single score to indicate the strength of the specificity for the overlapping marker genes in both
323 datasets. Since only cluster level C286 is assigned brain regions, the proportion of cells assigned
324 to a specific brain region was hierarchically collapsed into cluster level C66, so that each C66
325 cluster can have a certain number of cells assigned to different brain regions. To account for this,
326 we scaled the specificity product by the proportion of cells assigned for each brain region to
327 assign a brain region value. This specificity product weighted by the proportion of assigned cells
328 was then summed for all pairwise comparisons for each neuron cluster to get the final brain
329 region values.
330
331 *Neuropeptide networks*
332 To analyze the structure of the neuropeptidome, the entirety of neuropeptide genes
333 expression within a cell, we collated known "neuropeptide genes" (encoding pre-prohormone
334 mRNAs) from several databases (Burbach, 2010; Kim et al., 2011; Seal et al., 2023) and via
335 Ensembl (Martin et al., 2023), resulting in 127 genes (Supplementary Table 2). Presence/absence
336 matrices of dominant and subordinate animals, with a threshold of two reads, were constructed.
337 Equality of the two covariance matrices across social status was tested with methods developed
338 specifically for high dimensional data using HD (Chang et al., 2017) and CLX (Cai et al., 2013)
339 statistics. The correlation between the two matrices was calculated with a mantel test. Due to the
340 higher total number of nuclei sequenced in dominant males, the number of neurons expressing
341 individual neuropeptide genes was calculated and scaled by ~60% for dominant males, the ratio
342 of dominant neuron count to subordinate neuron count. Next, we generated co-expression

343 network of just neuropeptide genes for each genotype, filtering out low expressing nuclei and
344 genes. For each edge, representing a neuropeptide gene pair, an FDR corrected t-test was used to
345 determine if the specific co-expression pair was dominant and subordinate biased. Global and
346 local clustering coefficients were calculated with transitivity in igraph (Csárdi & Nepusz, 2006).
347
348

349 **RESULTS**

350 *Behavior and hormone analyses*

351 We first examined the behavior of dominant and subordinate males. Pairs of dominant
352 and subordinate males were only selected if dominant males were observed displaying
353 dominance behavior, body coloration, and the presence of an eyebar each day of global live
354 scoring over the 16-day observation period, along with at least one occurrence of a brooding
355 female in the community. As community setup included a bias in size to ensure a dominance
356 relationship, dominant males were significantly larger in size and testes size compared to
357 subordinate males. Additionally, circulating testosterone was significantly higher in dominant
358 males compared to subordinate males (dominant males: M = 94 ng/mL, SD = 52; subordinate
359 males: M = 9.6 ng/mL, SD = 14).

360

361 *Cell type identification of sequenced hypothalamus nuclei.*

362 We identified 27,870 individual single-cell transcriptomes across all samples and
363 successfully demultiplexed three genotypes per sample pool between dominant and subordinate
364 *A. burtoni* (Fig 1a, 1b). Cell-type calling demonstrated consistent proportion of cell types across
365 social status and biological replicates (Fig 1c) for six major hypothalamic neural cell categories:
366 Excitatory neurons, Inhibitory neurons, Oligodendrocytes, Astrocytes, Vascular cells, and cells
367 of the *pars tuberalis* (Steuernagel et al., 2022), indicating that there are no cell type composition
368 differences between the two groups (Supplemental Figure 9). For each of the non-neuron cell
369 types (Oligodendrocytes, Astrocytes, Vascular cells, and cells of the *pars tuberalis*), DEG
370 analysis identified several significantly different genes across social status at the cell-type
371 specific sub-cluster level (Supplemental Figure 1-4). Yet, for the 355 DEGs between status

372 across all cell type specific sub-clusters in each of the four neuronal cell types, 29 genes are
373 present in all cell types (18 present in three and 49 present in two cell types) (Supplemental
374 Figure 8). This demonstrates how all cell types appear to be impacted to some degree by social
375 status and a core set of genes where differentially expressed across all cell types, indicating
376 common gene expression differences. Of all nuclei clusters, seven appear to have biased
377 proportion of nuclei from either dominant or subordinate males above 66% (Supplemental Figure
378 9c). Two GLU nuclei clusters, clusters 11 and 16, showed subordinate bias. Of the five clusters
379 with a dominate bias two are GLU (cluster 7 and 20), one is unknown (cluster 28), and two are
380 oligodendrocytes (cluster 17 and 27).
381
382 *Neurons*
383 Subsequent hierarchical clustering of neurons demonstrates the presence of highly structured
384 subpopulations, including excitatory and inhibitory neurons (Fig 2a). We analyzed neuron gene
385 co-expression networks using hdWGCNA by generating metacell profiles separated by social
386 status and genotype identity to reduce individual transcriptomic variability, resulting in eight
387 gene co-expression modules (Fig 2b). We found that gene modules complement UMAP clusters
388 with some module eigengene expression being highly correlated with UMAP cluster (e.g. brown)
389 while others are more broadly represented across UMAP clusters (e.g. red) (Fig 2c). Differential
390 module eigengene analysis identified certain gene modules that are significantly differentially
391 expressed across social status in both directions (Fig 2d), with the brown module increased in the
392 dominant group and the blue module increased in subordinate group. This means that certain
393 gene networks are characteristic of each of the groups. GO analysis for these differentially
394 expressed modules showed they were significantly enriched for pathways related to ion transport

395 (Supplemental Figure 5). For example, the two modules increased in dominant males, blue and
396 yellow, are significantly enriched for molecular process GO terms for voltage-gated potassium
397 channel and calcium ion transport, respectively. The turquoise module, increased in
398 subordinates, is significantly enriched for molecular process GO terms related to membrane
399 receptors and transporters.
400
401 *Neurons expressing arginine vasopressin (AVP) pre-prohormone mRNA*
402 We used hierarchical clustering of neurons that expressed *arginine vasopressin* (AVP)
403 preprohormone mRNA to determine if there are transcriptomic subpopulations of AVP neurons
404 (Fig 3a). Neuronal nuclei were filtered to identify AVP expressing neurons with an SCT
405 normalized read count above one, resulting in 254 and 236 nuclei for dominant and subordinate
406 respectively. Clustering of these AVP neurons discerned five robust clusters that mirror the
407 structure of the larger neuronal dataset and are present in all six genotypes. The average AVP
408 expression level for each cluster per genotype was calculated. We found that only AVP neuron
409 cluster 4 had a higher average AVP expression compared to all other AVP neuron clusters (Fig
410 3b). Additionally, AVP neuron cluster three is comprised of nuclei from the broader neuron
411 cluster 17, which also is the cluster with the highest relative OXT expression. In the HypoMap
412 (Steuernagel et al., 2022) dataset, there is also one cluster (C66-22: Caprin2.GLU-6) that shows
413 the highest amount of gene expression for both AVP and OXT (Supplemental Figure 6). This
414 HypoMap cluster was comprised of nuclei from primarily eight different datasets and was
415 determined to be most likely spatially restricted to the mammalian Paraventricular Hypothalamic
416 Nucleus (PVN). Since we identified a cluster expressing both high levels of AVP and OXT, and

417 a similar cluster was identified in Hypomap dataset that was suggested to be spatially restricted
418 to the PVN, we predicted that these represented the same conserved cell types across vertebrates.
419 To assign AVP neuron clusters to potential brain regions, we used a scaled specificity
420 score for the broader neuron cluster that each AVP neuron cluster is a part of to compare
421 relationship with HypoMap clusters. AVP neuron cluster zero and cluster one were excluded
422 from brain region spatial analysis as these neuron clusters were too broad to have cluster specific
423 markers. Marker genes for the remaining selected neuron clusters that represented AVP neuron
424 clusters were successfully assigned mouse orthologs with Biomart (Martin et al., 2023). AVP
425 neuron cluster two, three, and four had 8/30, 16/64, and 13/80 marker genes assigned orthologs
426 respectively (Supplemental Figure 7). Using the cross-dataset brain region specificity score
427 described above (methods), we were able to generally assign AVP neuron clusters to brain
428 regions. We found that AVP neuron cluster three likely represented the PVN, while AVP neuron
429 clusters two and four represent what HypoMap designates (Pre)Mammillary regions.
430 Differential gene expression for the top 500 variable genes across AVP neurons was
431 performed with limmatrend, across genotype, for all AVP neurons and each AVP cluster.
432 Importantly, AVP neuron subpopulations have distinct socially regulated differentially expressed
433 genes compared to other AVP neuron subpopulations and these genes are masked when
434 comparing all AVP neurons across social status (Fig 3c).
435
436 *Uncovering the Neuropeptidome*
437 Neuropeptide genes were collated from a series of mammalian databases and orthologs
438 determined. We find broad expression across neurons for the 70 annotated neuropeptide
439 preprohormones across both dominant and subordinate males with a majority of neurons

440 expressing at least one. 77% of neurons expressed at least one neuropeptide and 47% expressed
441 at least two. Notably, 95% of the neurons obtained from dominant males expressed at least one
442 neuropeptide gene compared to only 63% in subordinate males. This difference was even more
443 pronounced when we examined the co-expression of at least two or more neuropeptide genes.
444 We found neuropeptide co-expression in 70% of the dominant males neurons compared to only
445 20% of subordinate male neurons (Fig 4a). We next compared the number of neurons expressing
446 each neuropeptide across social status to find neuropeptides that are more broadly expressed in
447 one social state over the other (Fig 4b). Most neuropeptides have relatively proportional
448 expression across neurons of both social states, indicating little to no broad scale bias in which
449 neuropeptides are being expressed. However, a handful of them were significantly subordinate
450 biased (cort) or dominant biased (pomc, GnRH1, prl). We then analyzed the neuropeptide co-
451 expression network to quantify this difference across the neuropeptidome and found that
452 dominant males have a significantly more integrated co-expression network of neuropeptides,
453 indicating broad-scale increased co-expression of neuropeptide genes as males attain social
454 dominance (Fig 4c). Almost all neuropeptides display a significant increase (FDR adjusted p-
455 value < 0.05) in relative neuropeptide co-expression in dominant males, except for the
456 significantly higher co-expression of cort and cbln1 in subordinate males. Comparing the
457 average dominant and subordinate neuropeptide co-expression networks found that these
458 networks are significantly positively correlated (Mantel test: $r = 0.18$, 95% CI for $r = 0.07$-0.25;
459 $p = 0.003$). This indicates that neuropeptide gene co-expression networks are similarly structured
460 across social status. Additionally, equality of the two high-dimensional covariance matrices
461 found that there is a significant similarity, p-value = 0, with two methods: HD statistic of ~55
462 (Chang et al., 2017) and CLX statistic of ~2998 (Cai et al., 2013). Therefore, both methods

463 indicate that the overall gene co-expression matrices are significantly different across social
464 status when testing large covariance and sparse but large magnitude differences, respectively.
465 Global transitivity of these average networks indicates a dominant male bias in correlation
466 coefficient (0.853 in dominants compared to 0.748 in subordinates), with strongest gene local
467 transitivity in dominant biased genes (gnrh1, prl, and pomc) (Fig 4d). This indicates that genes in
468 dominant male neuropeptidome tend to cluster together more relative to subordinate male, an
469 indicator of increased connectivity. Likewise, the presence of dominant biased neuropeptide
470 genes with high degree of local transitivity, and effectively no local transitivity in subordinate,
471 highlights how the neurons expressing these genes have status-specific gene co-expression
472 networks.
473
474

475 **DISCUSSION**

476 In this work, we demonstrate how differences in social status are regulated through
477 transcriptomic profiles at the cellular level. We successfully implemented single-nuclei RNA-seq
478 in a non-model organism to identify the transcriptional changes associated with social dominance
479 in the hypothalamus and POA of males in *A. burtoni*. We observed notable changes across social
480 status in the expression of gene co-expression networks within neurons. Our approach was able
481 to target specific known functional neural populations of AVP neurons to assess the impact of
482 the social environment on gene expression. We also applied a novel framework for
483 understanding the role the entire neuropeptidome plays in regulating social behavior.

484 We demonstrate how all cell types appear to be impacted to some degree by social status,
485 as a common core set of genes appeared significantly differentiated across all cell types. This
486 aligns with previous work on the role of non-neuronal cell types in regulating social behavior.
487 Importantly, of the seven clusters that had a bias in the proportion of nuclei from a certain social
488 status two are oligodendrocytes and both are biased towards dominate males. In this study we
489 primarily focused on the gene networks occurring with neurons due to their direct functional role
490 in regulating brain activity and social behavior. Particularly since neurons exhibit a relatively
491 high degree of transcriptomic variability and subpopulation structure (Romanov et al., 2017).
492 Using a specialized gene co-expression network approach, hdWGCNA (Morabito et al., 2023),
493 we were able to characterize gene networks that go beyond cell-type or statistical cluster. We
494 predicted that spatially restricted modules would represent neuron sub-type specific gene
495 networks while the broader modules represent gene pathways involved in ubiquitous cell
496 function. Therefore, we were surprised to find that only the brown module appears to be mainly
497 restricted to one of the broad neuronal cell types, that being inhibitory neurons. Additionally, we

498 found that certain module eigengenes were significantly correlated within cell clusters indicating
499 co-expression of some gene networks are within cells. This implies that most gene co-expression
500 networks do not necessarily represent major cell-type stratification, but are rather tied into cell
501 state, function, and likely spatial distribution. This result is supported by the differential
502 expression of module eigengenes across social status. Further work will need to be done to
503 understand what role these gene networks play in facilitating cell function. Based on the GO
504 analysis, most of the significantly differential expressed gene modules were significantly
505 enriched for receptors and transporters across the cellular membrane. Taken together, these
506 results highlight the role of these gene co-expression modules in cell signaling. This could
507 indicate that the underlying differences in neurons driving regulation of social status are integral
508 to the connectivity and interaction between neurons.

509 The vasopressin system has long been studied for its role in regulating social behavior.
510 Of course, this neuropeptide does not occur in isolation but within the transcriptome that a cell
511 expresses. We find evidence that there are multiple subpopulations of AVP expressing neurons
512 based on the cellular-level transcriptome. This is consistent with morphological and spatial data
513 identifying multiple subpopulations of AVP expressing neurons (Braford & Northcutt, 1983;
514 Moore & Lowry, 1998; O'Connell & Hofmann, 2011).

515 Previous research in *A. burtoni* demonstrated that the three main morphological
516 subpopulations of vasopressin neurons, parvo-, magno-, and giganto-cellular have distinct AVP
517 expression patterns across social status (Greenwood et al., 2008). Similarly, AVP, and not OXT,
518 has been shown to have a status dependent role in male social dominance in specific regions of
519 the POA in other species such as the weakly electric fish (Pouso et al., 2024).While we did not
520 identify any statistical clusters that matched these patterns in terms of gene expression, we did

521 identify that AVP neuron cluster 3 had the highest AVP expression. Importantly, parvocellular
522 neurons in the homolog of the mammalian paraventricular nucleus (PVN) do exhibit the highest
523 overall AVP expression (Greenwood et al., 2008; O'Connell & Hofmann, 2011)(O'Connell &
524 Hofmann, 2011; Greenwood, 2008). Since AVP neuron cluster 3 has highest the AVP
525 expression, is likely spatially assigned to the PVN, we could therefore predict that it likely
526 represents a homolog cell population to the parvocellular AVP neurons. Spatial assignment
527 applied in this study are predicted through integrating gene expression from this study and data
528 from the Allen brain atlas. Therefore, it is important to note that the spatial assignment applied in
529 this paper are by no means complete and require interpretation through multiple rounds of
530 abstraction (Allen brain atlas ISH expression to hypomap, hypomap markers to ensemble
531 orthologs, orthologs to neuron clusters markers). Ultimately, as spatial transcriptomics are
532 becoming more broadly accessible to researchers and non-model organisms, public databases
533 will incorporate more spatial data and this type of analysis will become more refined.
534 Inherent to all RNA studies is the understanding that transcriptome does not equate
535 proteome. Importantly, gene expression does not necessarily lead to mature peptides or guarantee
536 secretion (Meissner et al., 2013). This is particularly notable for peptidergic genes which
537 individually code for multiple possible peptides and are commonly co-expressed (Smith et al.,
538 2019). With single-nucleus RNA sequencing, we are limited to preprocessed nuclear mRNA of
539 preprohormones and therefore cannot determine which specific, or combination of, peptide
540 hormones would have ultimately been sequestered and secreted. Yet, activation patterns of
541 transcriptomic pathways and networks grant insight into how cells are responding and
542 maintaining status at that moment in time.

543 While trying to understand the impact of the neuropeptidome, it is important to
544 understand that these gene expression levels represent preprohoromone mRNA, not mature
545 peptide levels. While the OXT gene is broadly expressed, peptide studies indicate that only a
546 subset of cells in the POA contain the peptide in fish (Batten et al., 1990; Holmqvist & Ekström,
547 1991; van den Dungen et al., 1982). In general, the OXT system has not been studied to a great
548 extent in *A. burtoni* (Huffman et al., 2012). There is little evidence of AVP and OXT being co-
549 expressed at peptide level, even though they appear to be highly co-expressed in this single-
550 nuclei data. Although this is not entirely unprecedented as small populations of neurons co-
551 expressing AVP and OXT has previously been shown in rats (Kiyama & Emson, 1990; Mezey &
552 Kiss, 1991). This indicates that the expression of these genes are highly responsive to social
553 stimuli, the mechanisms that maintain and regulate peptide translation add an additional layer of
554 complexity. Similarly, while previous work has demonstrated in *A. burtoni* a dominant-bias in
555 peptide levels of GnRH1 (M. R. Davis & Fernald, 1990) and SST (Hofmann & Fernald, 2000),
556 only GnRH1 preprohormone mRNA levels reflect this difference at the cellular level.
557 Prior to single-cell transcriptomic approaches, cellular-level resolution of neuropeptide
558 expression has been limited to a single gene, or at most a couple of genes, making wide-scale co-
559 expression analysis difficult (K. H. Chen et al., 2015). Here, we take a novel approach targeting
560 the neuropeptidome in its entirety within individual cells. While only a few neuropeptide genes
561 show strong expression patterns based on social status, gene network approaches demonstrate
562 that dominant males maintain a more highly integrated network. In fact, this elevated gene
563 network integration pattern in dominant male *A. burtoni* has previously been demonstrated with
564 select neuropeptides (O'Connell & Hofmann, 2012b; Solomon-Lane & Hofmann, 2019). The
565 combinatorial nature of these cellular-level neuropeptide gene networks, taken together with this

566 more integrated network, implies that social status reflects more than simply the activation of a

567 few well-studied gene pathways but a complex interplay across the genome.

568

569 **CONCLUSIONS**

570 Here, we used single-nuclei RNA-seq in the hypothalamus and POA of *Astatotilapia*
571 *burtoni* to understand the effect that social status has on distinct cellular-level transcriptomic
572 profiles. We demonstrate how the impact of social dominance manifests from broader neuronal
573 gene networks to targeted changes in gene expression between known socially-relevant
574 populations of neurons. For the first time, we assess the entirety of the neuropeptidome of the
575 hypothalamus to understand the interaction of neuropeptide genes in two distinct social-status
576 populations.

577

578 **DATA AVAILABILITY STATEMENT**

579 The filtered deduplicated reads have been deposited to the NCBI Short Read Archive
580 (SRA), bioproject XXXXXX. The code used in this paper together with documentation can be
581 accessed as the GitHub repository, https://github.com/imillercrews/

582

583 **ACKNOWLEDGEMENTS**

584 We thank members of the Hofmann Lab for guidance and discussion along with Holly
585 Stevenson and the GSAF for assistance with sequencing. This work was supported by a US
586 department of Justice graduate fellowship and an EEB grant to IMC and NSF-IOS grant 1326187
587 to HAH.

588 **REFERENCES**

589 Batten, T. F. C., Cambre, M. L., Moons, L., & Vandesande, F. (1990). Comparative distribution of
590 neuropeptide-immunoreactive systems in the brain of the green molly, Poecilia latipinna.
591 *Journal of Comparative Neurology*, *302*(4), 893–919. https://doi.org/10.1002/cne.903020416
592 Braford, M. J., & Northcutt, R. G. (1983). Organization of the diencephalon and pretectum of the ray-
593 finned fishes. *Fish Neurobiology*, *2*, 117–163.
594 Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A. Y., Lim, Z. W.,
595 Bezault, E., Turner-Maier, J., Johnson, J., Alcazar, R., Noh, H. J., Russell, P., Aken, B., Alföldi, J.,
596 Amemiya, C., Azzouzi, N., … Di Palma, F. (2014). The genomic substrate for adaptive radiation in
597 African cichlid fish. *Nature*, *513*(7518), 375–381. https://doi.org/10.1038/nature13726
598 Burbach, J. P. H. (2010). Neuropeptides from concept to online database www.neuropeptides.nl.
599 *European Journal of Pharmacology*, *626*(1), 27–48. https://doi.org/10.1016/j.ejphar.2009.10.015
600 Cai, T., Liu, W., & Xia, Y. (2013). Two-Sample Covariance Matrix Testing and Support Recovery in High-
601 Dimensional and Sparse Settings. *Journal of the American Statistical Association*, *108*(501), 265–
602 277.
603 Chang, J., Zhou, W., Zhou, W.-X., & Wang, L. (2017). Comparing large covariance matrices under weak
604 conditions on the dependence structure and its application to gene clustering. *Biometrics*, *73*(1),
605 31–41. https://doi.org/10.1111/biom.12552
606 Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., & Zhuang, X. (2015). Spatially resolved, highly
607 multiplexed RNA profiling in single cells. *Science*, *348*(6233), aaa6090.
608 https://doi.org/10.1126/science.aaa6090
609 Chen, R., Wu, X., Jiang, L., & Zhang, Y. (2017). Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity.
610 *Cell Reports*, *18*(13), 3227–3241. https://doi.org/10.1016/j.celrep.2017.03.004

611 Choudhary, S., & Satija, R. (2022). Comparison and evaluation of statistical error models for scRNA-seq.
612 *Genome Biology*, *23*(1), 27. https://doi.org/10.1186/s13059-021-02584-9
613 Conte, M. A., Gammerdinger, W. J., Bartie, K. L., Penman, D. J., & Kocher, T. D. (2017). A high quality
614 assembly of the Nile Tilapia (Oreochromis niloticus) genome reveals the structure of two sex
615 determination regions. *BMC Genomics*, *18*(1), 341. https://doi.org/10.1186/s12864-017-3723-5
616 Crow, M., & Gillis, J. (2018, November 1). Co-expression in Single-Cell Analysis: Saving Grace or Original
617 Sin? *Trends in Genetics*, *34*(11), 823–831. https://doi.org/10.1016/j.tig.2018.07.007
618 Csárdi, G., & Nepusz, T. (2006). *The igraph software package for complex network research*.
619 https://www.semanticscholar.org/paper/The-igraph-software-package-for-complex-network-
620 Cs%C3%A1rdi-Nepusz/1d2744b83519657f5f2610698a8ddd177ced4f5c
621 Davis, A., Gao, R., & Navin, N. E. (2019). SCOPIT: Sample size calculations for single-cell sequencing
622 experiments. *BMC Bioinformatics*, *20*(1), 566. https://doi.org/10.1186/s12859-019-3167-9
623 Davis, M. R., & Fernald, R. D. (1990). Social control of neuronal soma size. *Journal of Neurobiology*, *21*(8),
624 1180–1188. https://doi.org/10.1002/neu.480210804
625 Dijkstra, P. D., Verzijden, M. N., Groothuis, T. G. G., & Hofmann, H. A. (2012). Divergent hormonal
626 responses to social competition in closely related species of haplochromine cichlid fish.
627 *Hormones and Behavior*, *61*(4), 518–526. https://doi.org/10.1016/j.yhbeh.2012.01.011
628 Ding, J., Adiconis, X., Simmons, S. K., Kowalczyk, M. S., Hession, C. C., Marjanovic, N. D., Hughes, T. K.,
629 Wadsworth, M. H., Burks, T., Nguyen, L. T., Kwon, J. Y. H., Barak, B., Ge, W., Kedaigle, A. J.,
630 Carroll, S., Li, S., Hacohen, N., Rozenblatt-Rosen, O., Shalek, A. K., … Levin, J. Z. (2020).
631 Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nature*
632 *Biotechnology*, *38*(6), 737–746. https://doi.org/10.1038/s41587-020-0465-8

633 Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., & Coleman, P. (1992). Analysis
634 of gene expression in single live neurons. *Proceedings of the National Academy of Sciences of*
635 *the United States of America*, *89*(7), 3010–3014.
636 Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for
637 multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048.
638 https://doi.org/10.1093/bioinformatics/btw354
639 Fernald, R. D., & Hirata, N. R. (1977). Field study of Haplochromis burtoni: Quantitative behavioural
640 observations. *Animal Behaviour*, *25*, 964–975. https://doi.org/10.1016/0003-3472(77)90048-3
641 Greenwood, A. K., Wark, A. R., Fernald, R. D., & Hofmann, H. A. (2008). Expression of arginine vasotocin
642 in distinct preoptic regions is associated with dominant and subordinate behaviour in an African
643 cichlid fish. *Proceedings of the Royal Society of London B: Biological Sciences*, *275*(1649), 2393–
644 2402. https://doi.org/10.1098/rspb.2008.0622
645 Grindberg, R. V., Yee-Greenbaum, J. L., McConnell, M. J., Novotny, M., O'Shaughnessy, A. L., Lambert, G.
646 M., Arauzo-Bravo, M. J., Lee, J., Fishman, M., Robbins, G. E., Lin, X., Venepally, P., Badger, J. H.,
647 Galbraith, D. W., Gage, F. H., & Lasken, R. S. (2013). RNA-sequencing from single nuclei.
648 *Proceedings of the National Academy of Sciences*, *110*(49), 19802–19807.
649 https://doi.org/10.1073/pnas.1319700110
650 Grün, D., & Van Oudenaarden, A. (2015). Design and Analysis of Single-Cell Sequencing Experiments.
651 *Cell*, *163*(4), 799–810. https://doi.org/10.1016/j.cell.2015.10.039
652 Guo, H., & Li, J. (2021). scSorter: Assigning cells to known cell types according to marker genes. *Genome*
653 *Biology*, *22*(1), 69. https://doi.org/10.1186/s13059-021-02281-7
654 Heaton, H., Talman, A. M., Knights, A., Imaz, M., Gaffney, D. J., Durbin, R., Hemberg, M., & Lawniczak, M.
655 K. N. (2020). Souporcell: Robust clustering of single-cell RNA-seq data by genotype without

656 reference genotypes. *Nature Methods 2020 17:6*, *17*(6), 615–620.
657 https://doi.org/10.1038/s41592-020-0820-1
658 Hofmann, H. a, & Fernald, R. D. (2000). Social status controls somatostatin neuron size and growth. *The*
659 *Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *20*(12), 4740–4744.
660 Holmqvist, B. I., & Ekström, P. (1991). Galanin-like immunoreactivity in the brain of teleosts: Distribution
661 and relation to substance P, vasotocin, and isotocin in the atlantic salmon (Salmo salar). *Journal*
662 *of Comparative Neurology*, *306*(3), 361–381. https://doi.org/10.1002/cne.903060302
663 Huffman, L. S., O'Connell, L. A., Kenkel, C. D., Kline, R. J., Khan, I. A., & Hofmann, H. A. (2012).
664 Distribution of nonapeptide systems in the forebrain of an African cichlid fish, Astatotilapia
665 burtoni. *Journal of Chemical Neuroanatomy*, *44*(2), 86–97.
666 https://doi.org/10.1016/j.jchemneu.2012.05.002
667 Ianevski, A., Giri, A. K., & Aittokallio, T. (2022). Fully-automated and ultra-fast cell-type identification
668 using specific marker combinations from single-cell transcriptomic data. *Nature*
669 *Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-28803-w
670 Kidd, C. E., Kidd, M. R., & Hofmann, H. A. (2010). Measuring multiple hormones from a single water
671 sample using enzyme immunoassays. *General and Comparative Endocrinology*, *165*(2), 277–285.
672 https://doi.org/10.1016/j.ygcen.2009.07.008
673 Kim, Y., Bark, S., Hook, V., & Bandeira, N. (2011). NeuroPedia: Neuropeptide database and spectral
674 library. *Bioinformatics*, *27*(19), 2772–2773. https://doi.org/10.1093/bioinformatics/btr445
675 Kiyama, H., & Emson, P. C. (1990). Evidence for the co-expression of oxytocin and vasopressin
676 messenger ribonucleic acids in magnocellular neurosecretory cells: Simultaneous demonstration
677 of two neurohypophysin messenger ribonucleic acids by hybridization histochemistry. *Journal of*
678 *Neuroendocrinology*, *2*(3), 257–259. https://doi.org/10.1111/j.1365-2826.1990.tb00401.x

679 Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis.
680 *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559
681 Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). voom: Precision weights unlock linear model analysis
682 tools for RNA-seq read counts. *Genome Biology*, *15*(2), R29. https://doi.org/10.1186/gb-2014-
683 15-2-r29
684 Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S.,
685 Brockway, K. S., Byrnes, E. J., Chen, L., Chen, L., Chen, T.-M., Chi Chin, M., Chong, J., Crook, B. E.,
686 Czaplinska, A., Dang, C. N., Datta, S., … Jones, A. R. (2007). Genome-wide atlas of gene
687 expression in the adult mouse brain. *Nature*, *445*(7124), Article 7124.
688 https://doi.org/10.1038/nature05453
689 Martin, F. J., Amode, M. R., Aneja, A., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Becker, A., Bennett,
690 R., Berry, A., Bhai, J., Bhurji, S. K., Bignell, A., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju,
691 S. B., Charkhchi, M., Cockburn, A., Da Rin Fiorretto, L., … Flicek, P. (2023). Ensembl 2023. *Nucleic*
692 *Acids Research*, *51*(D1), D933–D941. https://doi.org/10.1093/nar/gkac958
693 Maruska, K. P., & Fernald, R. D. (2013). Social Regulation of Male Reproductive Plasticity in an African
694 Cichlid Fish. *Integrative and Comparative Biology*, *53*(6), 938–950.
695 https://doi.org/10.1093/icb/ict017
696 Maruska, K. P., & Fernald, R. D. (2018). Astatotilapia burtoni: A Model System for Analyzing the
697 Neurobiology of Behavior. *ACS Chemical Neuroscience*, *9*(8), 1951–1962.
698 https://doi.org/10.1021/acschemneuro.7b00496
699 Maruska, K. P., Zhang, A., Neboori, A., & Fernald, R. D. (2013). Social Opportunity Causes Rapid
700 Transcriptional Changes in the Social Behaviour Network of the Brain in an African Cichlid Fish.
701 *Journal of Neuroendocrinology*, *25*(2), 145–157. https://doi.org/10.1111/j.1365-
702 2826.2012.02382.x

703 Meissner, F., Scheltema, R. A., Mollenkopf, H. J., & Mann, M. (2013). Direct proteomic quantification of
704 the secretome of activated immune cells. *Science*, *340*(6131), 475–478.
705 https://doi.org/10.1126/SCIENCE.1232578/SUPPL_FILE/MEISSNER_SM.PDF
706 Mezey, E., & Kiss, J. Z. (1991). Coexpression of vasopressin and oxytocin in hypothalamic supraoptic
707 neurons of lactating rats. *Endocrinology*, *129*(4), 1814–1820. https://doi.org/10.1210/endo-129-
708 4-1814
709 Miller-Crews, I., Matz, M. V., & Hofmann, H. A. (2021). A 2b-RAD parentage analysis pipeline for complex
710 and mixed DNA samples. *Forensic Science International: Genetics*, *55*.
711 https://doi.org/10.1016/j.fsigen.2021.102590
712 Moffitt, J. R., Bambah-Mukku, D., Eichhorn, S. W., Vaughn, E., Shekhar, K., Perez, J. D., Rubinstein, N. D.,
713 Hao, J., Regev, A., Dulac, C., & Zhuang, X. (2018). Molecular, spatial and functional single-cell
714 profiling of the hypothalamic preoptic region. *Science (New York, N.Y.)*, eaau5324.
715 https://doi.org/10.1126/science.aau5324
716 Moore, F. L., & Lowry, C. A. (1998). Comparative neuroanatomy of vasotocin and vasopressin in
717 amphibians and other vertebrates. *Comparative Biochemistry and Physiology Part C:*
718 *Pharmacology, Toxicology and Endocrinology*, *119*(3), 251–260. https://doi.org/10.1016/s0742-
719 8413(98)00014-0
720 Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E., & Swarup, V. (2023). hdWGCNA identifies co-
721 expression networks in high-dimensional transcriptomics data. *Cell Reports Methods*, *3*(6),
722 100498. https://doi.org/10.1016/j.crmeth.2023.100498
723 Nguyen, Q. H., Pervolarakis, N., Nee, K., & Kessenbrock, K. (2018). Experimental Considerations for
724 Single-Cell RNA Sequencing Approaches. *Frontiers in Cell and Developmental Biology*, *6*, 108.
725 https://doi.org/10.3389/fcell.2018.00108

O'Connell, L. A., & Hofmann, H. A. (2011). The Vertebrate mesolimbic reward system and social behavior network: A comparative synthesis. *The Journal of Comparative Neurology*, *519*(18), 3599–3639. https://doi.org/10.1002/cne.22735

O'Connell, L. A., & Hofmann, H. A. (2012a). Evolution of a Vertebrate Social Decision-Making Network. *Science*, *336*(6085), 1154–1157. https://doi.org/10.1126/science.1218889

O'Connell, L. A., & Hofmann, H. A. (2012b). Social status predicts how sex steroid receptors regulate complex behavior across levels of biological organization. *Endocrinology*, *153*(3), 1341–1351. https://doi.org/10.1210/en.2011-1663

Ofengeim, D., Giagtzoglou, N., Huh, D., Zou, C., & Yuan, J. (2017). Single-Cell RNA Sequencing: Unraveling the Brain One Cell at a Time. *Trends in Molecular Medicine*, *23*(6), 563–576. https://doi.org/10.1016/j.molmed.2017.04.006

Pouso, P., Cabana, Á., Francia, V., & Silva, A. (2024). Vasotocin but not isotocin is involved in the emergence of the dominant-subordinate status in males of the weakly electric fish, Gymnotus omarorum. *Hormones and Behavior*, *158*, 105446. https://doi.org/10.1016/j.yhbeh.2023.105446

Romanov, R. A., Zeisel, A., Bakker, J., Girach, F., Hellysaz, A., Tomer, R., Alpár, A., Mulder, J., Clotman, F., Keimpema, E., Hsueh, B., Crow, A. K., Martens, H., Schwindling, C., Calvigioni, D., Bains, J. S., Máté, Z., Szabó, G., Yanagawa, Y., … Harkany, T. (2017). Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nature Neuroscience*, *20*(2), 176–188. https://doi.org/10.1038/nn.4462

Salem, N. A., Manzano, L., Keist, M. W., Ponomareva, O., Roberts, A. J., Roberto, M., & Mayfield, R. D. (2024). Cell-type brain-region specific changes in prefrontal cortex of a mouse model of alcohol dependence. *Neurobiology of Disease*, *190*, 106361. https://doi.org/10.1016/j.nbd.2023.106361

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, *33*(5), Article 5. https://doi.org/10.1038/nbt.3192

750 Schmid, K. T., Höllbacher, B., Cruceanu, C., Böttcher, A., Lickert, H., Binder, E. B., Theis, F. J., & Heinig, M.
751 (2021). scPower accelerates and optimizes the design of multi-sample single cell transcriptomic
752 studies. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-26779-7
753 Seal, R. L., Braschi, B., Gray, K., Jones, T. E. M., Tweedie, S., Haim-Vilmovsky, L., & Bruford, E. A. (2023).
754 Genenames.org: The HGNC resources in 2023. *Nucleic Acids Research*, *51*(D1), D1003–D1009.
755 https://doi.org/10.1093/nar/gkac888
756 Smith, S. J., Smbül, U., Graybuck, L. T., Collman, F., Seshamani, S., Gala, R., Gliko, O., Elabbady, L., Miller,
757 J. A., Bakken, T. E., Rossier, J., Yao, Z., Lein, E., Zeng, H., Tasic, B., & Hawrylycz, M. (2019). Single-
758 cell transcriptomic evidence for dense intracortical neuropeptide networks. *eLife*, *8*.
759 https://doi.org/10.7554/ELIFE.47889
760 Solomon-Lane, T. K., & Hofmann, H. A. (2019). Early-life social environment alters juvenile behavior and
761 neuroendocrine function in a highly social cichlid fish. *Hormones and Behavior*, *115*, 104552.
762 https://doi.org/10.1016/j.yhbeh.2019.06.016
763 Soneson, C., & Robinson, M. D. (2018). Bias, robustness and scalability in single-cell differential
764 expression analysis. *Nature Methods 2018 15:4*, *15*(4), 255–261.
765 https://doi.org/10.1038/nmeth.4612
766 Steuernagel, L., Lam, B. Y. H., Klemm, P., Dowsett, G. K. C., Bauder, C. A., Tadross, J. A., Hitschfeld, T. S.,
767 del Rio Martin, A., Chen, W., de Solis, A. J., Fenselau, H., Davidsen, P., Cimino, I., Kohnke, S. N.,
768 Rimmington, D., Coll, A. P., Beyer, A., Yeo, G. S. H., & Brüning, J. C. (2022). HypoMap—A unified
769 single-cell gene expression atlas of the murine hypothalamus. *Nature Metabolism 2022 4:10*,
770 *4*(10), 1402–1419. https://doi.org/10.1038/s42255-022-00657-y
771 Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui,
772 A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell.
773 *Nature Methods*, *6*(5), 377–382. https://doi.org/10.1038/nmeth.1315

774 van den Dungen, H. M., Buijs, R. M., Pool, C. W., & Terlou, M. (1982). The distribution of vasotocin and
775 isotocin in the brain of the rainbow trout. *Journal of Comparative Neurology*, *212*(2), 146–157.
776 https://doi.org/10.1002/cne.902120205
777 Wang, Y., & Navin, N. E. (2015). Advances and applications of single-cell sequencing technologies.
778 *Molecular Cell*, *58*(4), 598–609. https://doi.org/10.1016/j.molcel.2015.05.005
779 Wright, R. M., Aglyamova, G. V., Meyer, E., & Matz, M. V. (2015). Gene expression associated with white
780 syndromes in a reef building coral, Acropora hyacinthus. *BMC Genomics*, *16*(1), 371.
781 https://doi.org/10.1186/s12864-015-1540-2
782 Wu, Y. E., Pan, L., Zuo, Y., Li, X., & Hong, W. (2017). Detecting Activated Cell Populations Using Single-Cell
783 RNA-Seq. *Neuron*, *96*(2), 313-329.e6. https://doi.org/10.1016/j.neuron.2017.09.026
784 Zappia, L., & Oshlack, A. (2018). Clustering trees: A visualization for evaluating clusterings at multiple
785 resolutions. *GigaScience*, *7*(7), giy083. https://doi.org/10.1093/gigascience/giy083
786 Zeng, Z., Miao, N., & Sun, T. (2018). Revealing cellular and molecular complexity of the central nervous
787 system using single cell sequencing. *Stem Cell Research & Therapy*, *9*(1), 234.
788 https://doi.org/10.1186/s13287-018-0985-z
789 Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D.,
790 McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G.,
791 Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H.
792 (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*,
793 *8*(1), Article 1. https://doi.org/10.1038/ncomms14049
794
795

FIGURES

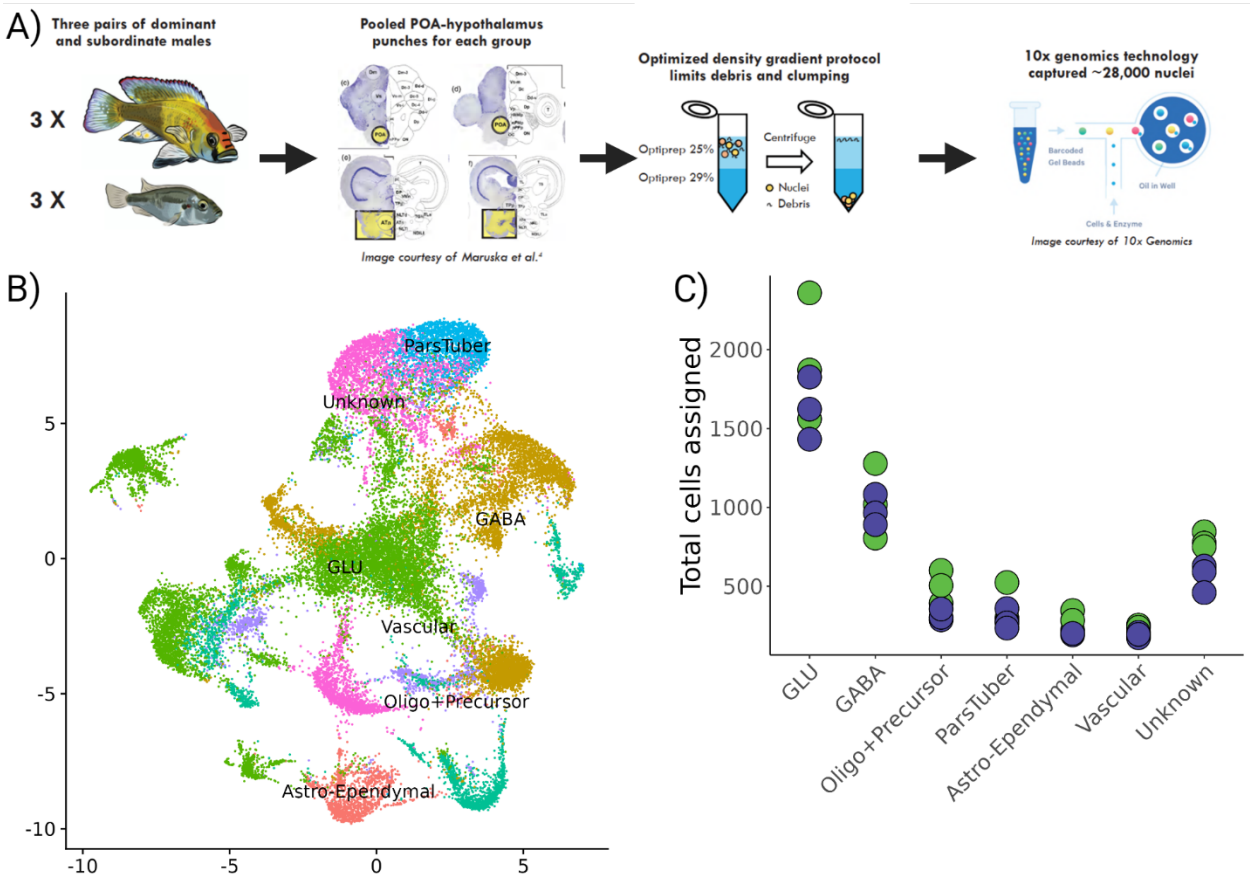

Figure 1: Experimental overview for snRNA-seq pipeline and cell type identification. A)Three pairs of dominant and subordinate male *A. burtoni* were collected and the POA-hypothalamus were pooled for each group prior to single nuclei isolation and snRNA-seq with 10x genomics. Fish illustrations curtesy of Dr. Caitlin Friesen. B) UMAP projection of single nuclei data with color-coded and labelled cell type assignments from scsorter using marker genes from HypoMap. C) Cell type identification resulted in roughly equal total cells assigned across all genotypes for dominant (purple) and subordinate (green) males.

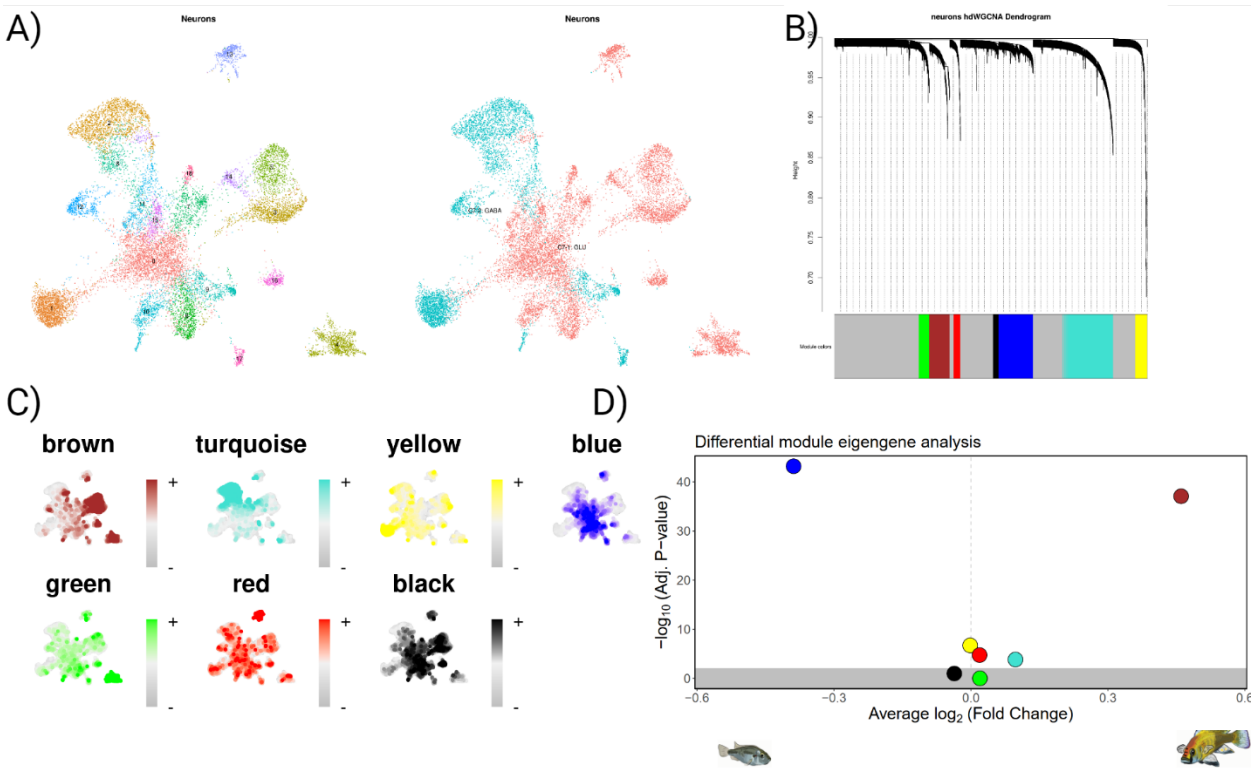

Figure 2: Neuron gene network analysis. A) UMAP projections for all neuron nuclei demonstrate the degree of statistical clustering (left) and separation into excitatory (pink) and inhibitory (blue) neurons (right). B) Dendrogram representing implementation of single-cell optimized weighted gene co-expression network analysis, hdWGCNA, identified six gene expression modules. C) hdWGCNA module eigengene values across neuronal UMAP space illustrate how most gene co-expression networks are expressed across neurons. D) Gene modules show significant differential expression across social status.

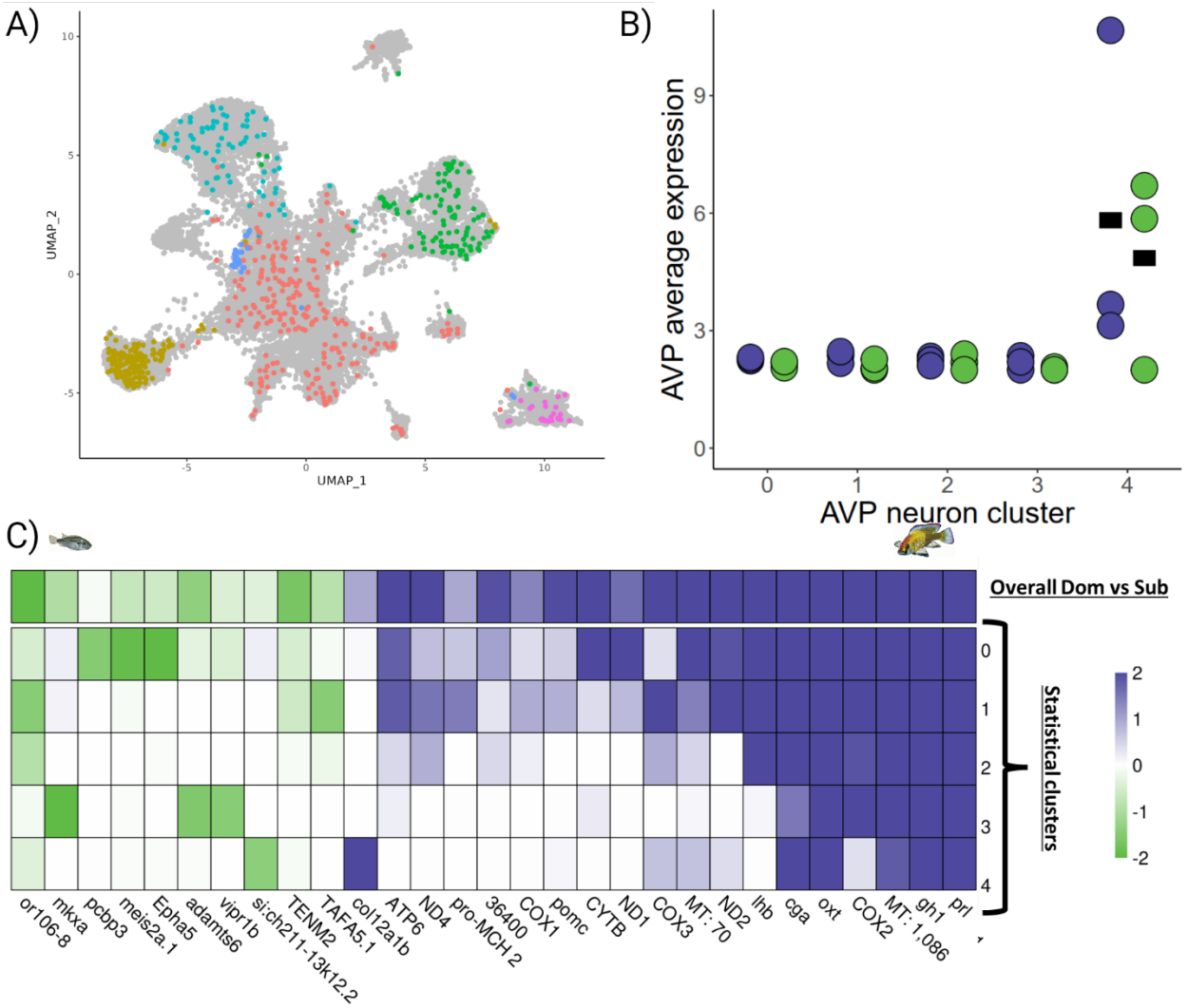

Figure 3: AVP expressing neuron subpopulations. A) Projection on neuron UMAP of all AVP expressing neurons colored by the five statistical subpopulation clusters. B) Average AVP expression for each genotype across social status and AVP neuron statistical clusters. C) Heatmap of statistically significant (adjusted p-value < 0.05) DEGs determined by limmatrend analysis across genotypes for all AVP expressing and for each AVP neuron statistical cluster.

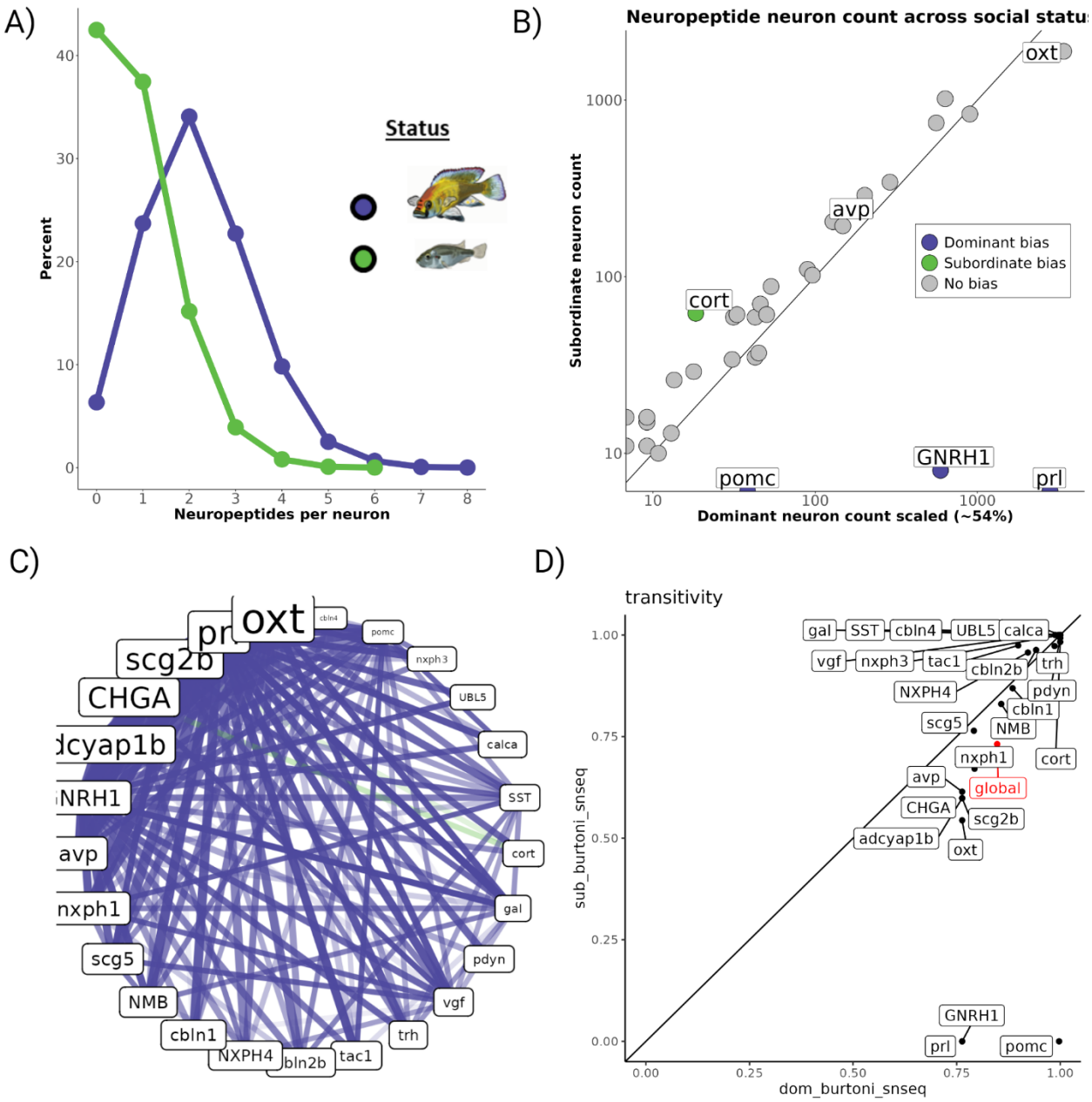

Figure 4: Uncovering the Neuropeptidome. A) Neurons broadly express neuropeptide genes, dominant male neurons have a higher percentage of neuropeptide co-expression. B) The number of neurons expressing each neuropeptide across social status. C) Significantly (FDR p-value < 0.05) different edges between the dominant (purple) and subordinate (green) neuropeptide gene co-expression network. D) Comparison of both average global and average local transitivity across social status of neuropeptide co-expression networks from each genotype.