

The enteric nervous system of the human and mouse colon at a single-cell resolution

Eugene Drokhlyansky^{1*}, Christopher S. Smillie^{1*}, Nicholas Van Wittenberghe^{1*}, Maria Ericsson², Gabriel K. Griffin^{3,4}, Danielle Dionne¹, Michael S. Cuoco¹, Max N. Goder-Reiser⁵, Tatyana Sharova⁵, Andrew J. Aguirre^{3,6}, Genevieve M. Boland⁸, Daniel Graham^{3,8,9}, Orit Rozenblatt-Rosen^{1,13}, Ramnik J. Xavier^{3,8,10,11,13}, Aviv Regev^{1,12,13}

¹Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA

²Department of Cell Biology, Harvard Medical School, Boston, MA, USA

³Broad Institute of MIT and Harvard, Cambridge, MA, USA

⁴Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA

⁵Massachusetts General Hospital, Cambridge, MA, USA

⁶Department of Medical Oncology, Dana Farber Cancer Institute, Boston, MA, USA

⁷Department of Surgery, Massachusetts General Hospital, Cambridge, MA, USA

⁸Department of Molecular Biology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

⁹Gastrointestinal Unit and Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

¹⁰Harvard Medical School, Boston, MA, USA.

¹¹Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, MA, USA.

¹²Howard Hughes Medical Institute and Koch Institute for Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA.

¹³Corresponding authors: orit@broadinstitute.org (O.R.R.), xavier@molbio.mgh.harvard.edu (R.J.X.), aregev@broadinstitute.org (A.R.).

* These authors contributed equally to this work.

Abstract

As the largest branch of the autonomic nervous system, the enteric nervous system (ENS) controls the entire gastrointestinal tract, but remains incompletely characterized. Here, we develop RAISIN RNA-seq, which enables the capture of intact single nuclei along with ribosome-bound mRNA, and use it to profile the adult mouse and human colon to generate a reference map of the ENS at a single-cell resolution. This map reveals an extraordinary diversity of neuron subsets across intestinal locations, ages, and circadian phases, with conserved transcriptional programs that are shared between human and mouse. These data suggest possible revisions to the current model of peristalsis and molecular mechanisms that may allow enteric neurons to orchestrate tissue homeostasis, including immune regulation and stem cell maintenance. Human enteric neurons specifically express risk genes for neuropathic, inflammatory, and extra-intestinal diseases with concomitant gut dysmotility. Our study therefore provides a roadmap to understanding the ENS in health and disease.

Introduction

The enteric nervous system (ENS) is an extensive network of neurons and glia along the gastrointestinal (GI) tract, which coordinates motility, digestion, nutrient absorption, and barrier defense (1). Robert Remak was likely the first to recognize the presence of enteric neurons in the intestine; his students, Leopold Auerbach and Georg Meissner, subsequently embarked on the rich effort to map the ENS using neuroanatomy and neurophysiology (reviewed in (2)). These and subsequent studies showed that the human ENS rivals the spinal cord in complexity (3). In humans, the ENS is broadly partitioned into the myenteric (Auerbach's) plexus and submucosal (Meissner's) plexus (4), with anecdotally reported differences in anatomy and composition within ganglia, across intestinal regions, and among species (5). In addition, other factors are thought to contribute to ENS heterogeneity, including age (6), sex (7), circadian phase (8), and functional motility disorders (9).

The ENS is implicated in a broad range of intestinal and extraintestinal disorders. Primary enteric neuropathies, including Hirschsprung's disease, chronic intestinal pseudo-obstruction, Waardenburg syndrome type IV, and MASH1 deficiency, directly affect enteric neurons, resulting in agangliosis and impaired GI transit (10), but are poorly understood (11). Moreover, ENS dysfunction may impact local inflammation, as suggested by studies of neuro-epithelial and neuro-immune interactions (12), such as neuronal activation of group 2 innate lymphoid cells (ILC2s) (13), motivating ENS characterization in other diseases that affect the gut (14). Intriguingly, several extraintestinal disorders, including those affecting the central nervous system (CNS), such as autism spectrum disorders (15) and Parkinson's disease (16), are associated with early GI

motility dysfunction. However, the pathophysiology of the ENS across these disorders, including affected cell types, is poorly understood.

Despite its central role in intestinal function and health, our understanding of the ENS is incomplete due to longstanding technical challenges. First, enteric neurons are rare in any sample because the ENS is dispersed among other cell types within the intestine (*e.g.*, myocytes and fibroblasts). Moreover, they are exceptionally challenging to isolate, limiting our ability to enrich populations for genomic profiling in either bulk or even single cells, and restricting most work on the ENS to rodent models with relatively few human studies (17). Even recent efforts to study the ENS using single cell genomics have been restricted to a limited number of neurons from embryonic animals or the myenteric plexus of early post-natal mice (18, 19).

Here, we generated a reference map of the ENS at single cell resolution across age, gender, location, circadian phases, and species (**Fig. 1**). We first developed a new method, Ribosomes And Intact Single Nucleus (“RAISIN”) RNA-seq, and applied it to generate a high quality single-cell census of the ENS in adult humans and mice, overcoming challenges in single-cell and single-nucleus RNA-seq (scRNA-Seq, snRNA-seq; (20–24) of the ENS. In the mouse, we used genetic tools to directly enrich for and profile 2,447 enteric neurons and 2,710 glia using deep, full-length snRNA-seq spanning four colon segments (proximal to distal) of three transgenic models (both sexes, multiple ages, two phases of the circadian rhythm). In humans, where transgene-based enrichment was not possible, we sequenced 163,741 single RAISINs (*i.e.* nuclei and attached ribosomes) from the muscularis propria of 10 individuals (men and women; 35-90 years old) and identified diverse cell types, among them 831 enteric neurons and 431 rare Interstitial Cells of

Cajal (ICCs). Enteric neurons partitioned into 24 murine and 11 human subsets, which we annotated with putative functions (*e.g.*, motor, sensory, secretomotor) using literature-derived marker genes, and matched between the two species based on conserved transcriptional programs. We mapped signaling interactions between human enteric neurons and other cell types in the colon, identifying possible neuro-immune, neuro-adipose, neuro-epithelial, neuro-muscular, and neuro-ICC pathways. Finally, we show that enteric neurons express genes specifically associated with primary enteroneuropathies, inflammatory disorders of the gut, and CNS disorders with early gut motility dysfunction, highlighting their potential roles in these diseases.

RESULTS

Validation of murine models to label the ENS

Because neurons comprise less than 1% of all colon cells, we first devised a strategy to enrich for the mouse ENS using three transgenic models, which label both the developing and mature ENS. Because the ENS is derived from the neural crest, we first targeted the pre-migratory and migratory neural crest using *Wnt1-Cre2* or *Sox10-Cre* drivers, respectively (25). As the *Wnt1-Cre* driver leads to neurodevelopmental defects, we used the corrected *Wnt1-Cre2* driver (26). For the *Sox10* driver, we used a *Sox10-Cre* line (27) that robustly targets ~80% of migratory neural crest cells (28) and has been successfully used to study the ENS (29). For both Cre-driver lines, we tagged nuclei in cells with a history of Cre-expression using the conditional *INTACT* (Isolation of Nuclei Tagged in specific Cell Types) allele (30). Finally, to enrich for mature neurons, we used *Uchl1-Histone2BmCherry:GFP-gpi* (subsequently referred to as *Uchl1-H2B mCherry*) mice, which were reported to label neurons and possibly neuroendocrine cells (31), although we found no evidence of the latter (see below).

For all three transgenic lines, we validated nuclei labeling within TUBB3⁺ neurons and confirmed our ability to enrich for extracted labeled nuclei using FACS (**Fig. 2a**). For the *Sox10-Cre* driver, we confirmed extensive neuron labeling by generating a triple transgenic animal harboring *Sox10-Cre*, *INTACT*, and conditional *tdTomato* (32) alleles, to label both the nuclei (*i.e.* *INTACT*) and cell bodies and their projections (*i.e.* *tdTomato*) of the ENS. There was excellent concordance between TUBB3 (neuron) immunostaining and reporter expression within the mouse colon (**Fig. 2b**; *tdTomato*⁺/*TUBB3*⁻ cells represent glia). For the *Wnt1-Cre2* driver, we observed labeled neuron nuclei, and also extensive signal in the colon mucosa (**Fig. 2a**); we validated that the *Wnt1-Cre2* driver also labeled colon epithelial cells by snRNA-seq (below). This off-target labeling may explain why a previous study using the *Wnt1-Cre* driver to target the ENS removed the mucosa when profiling enteric neurons of early post-natal mice with scRNA-seq (33). Lastly, for the *Uchl1-H22B mCherry* mice, we observed labeling of enteric neurons but not of enterendocrine cells (the main neuroendocrine type in the intestine; (34), by histology (**Fig. 2a**) and snRNA-seq (below).

Systematic optimization of nuclei extraction enables profiling of single ENS colonic nuclei

We next applied two previously published snRNA-seq protocols (35, 36) to FACS-enriched labeled-nuclei from the colon of *Sox10-Cre;INTACT* mice. Unfortunately, neither protocol performed well on ENS nuclei from the colon, in contrast to their excellent performance on labeled nuclei from the brain (**Supp. Fig. 1a**). In addition, the *Sox10-Cre* driver labeled both neurons and oligodendrocytes in the brain (**Supp. Fig. 1a**), whereas we anticipated recovering only brain oligodendrocytes (37). These limitations raised the need to develop new snRNA-seq approaches.

To develop snRNA-seq methods that are compatible with a broader range of tissues, including colon, we performed an optimization with nuclei from adult *Sox10-Cre;INTACT* mice, systematically varying the detergent (NP40, CHAPS, Tween, or Digitonin), detergent concentrations, buffer (HEPES, Tris, Tricine), mechanical extraction conditions (dounced, chopped, or ground tissue), and added modifiers (*e.g.* salts, polyamines) used in nuclei isolation (**Materials and Methods**), and compared to published protocols (38, 39). We profiled 5,236 nuclei isolated across 104 preparations spanning 36 extraction conditions (mean = 145 nuclei per condition) using SMART-Seq2 (**Supp. Table 1, Materials and Methods**). We scored conditions by (1) the recovery rate of neurons and glia relative to other cells (*i.e.* damaged or contaminating cells), (2) the number of genes detected per cell; and (3) an ENS signature score of known markers of enteric neurons and glia (**Fig. 3a-c; Supp. Fig. 1b-d; Supp. Table 1; Materials and Methods**).

Detergent type, detergent concentration, buffer, and mechanical force each impacted quality metrics (**Supp. Fig. 1b-d and 2**) and we identified two conditions with high ENS recovery and low contamination rates (~20% neurons, 55% glia, 25% contamination across both conditions, **Fig. 3b**), which also yielded high-quality profiles enriched in the ENS signature score (**Fig. 3c**). We termed these preparations “CST” (0.49% CHAPS detergent, Salts, Tris buffer, and “chopped” tissue) and “TST” (0.03% Tween-20 detergent, Salts, Tris buffer, and “chopped” tissue). Both preparations yielded higher numbers of detected genes than published methods (mean = 2,486 for CST and 2,542 for TST *vs.* 1,502 for published protocols on average across all nuclei; $p < 10^{-10}$ for both comparisons; Wilcoxon test).

Preservation of ribosomes or rough endoplasmic reticulum on the nuclear envelope allows for mature mRNA capture

To understand the basis for these performance differences among nuclei preparations, we compared nuclei structure between CST, TST, and published preparations for snRNA-seq (40, 41), using ultrathin-section transmission electron microscopy (TEM) (**Materials and Methods; Fig. 3d; Supp. Fig. 2**). As expected, the two published methods yielded isolated intact nuclei (**Fig. 3d**). In contrast, CST preserved not only the nuclear envelope, but also the ribosomes (42) on the outer nuclear membrane (**Fig. 3d**); we thus termed this method RAISIN (Ribosomes And Intact Single Nucleus) RNA-seq. TST maintained both the rough ER and its attached ribosomes (43) on the outer nuclear membrane (**Fig. 3d**); we thus termed this method, INNER Cell (Intact Nucleus and Endoplasmic Reticulum from a single Cell) RNA-seq. Consistent with the TEM results, both RAISIN RNA-seq and INNER Cell RNA-seq yielded higher exon:intron ratios than the published methods (**Fig. 3e**; 41% and 64% increases, respectively), suggesting greater recovery of mRNA relative to pre-mRNA.

Of the two methods, we opted to use RAISIN RNA-seq to profile the mouse and human ENS, because it captures more neurons and has fewer contaminants than INNER Cell RNA-seq (**Fig. 3b; Supp. Fig. 1b-d**). To test whether RAISIN RNA-seq is compatible with massively parallel droplet-based scRNA-seq, we also sequenced 10,889 unsorted RAISINs from the mouse colon (**Materials and Methods**). We recovered most major cell types in the colon, including epithelial cells, myocytes, fibroblasts, endothelial cells, immune cells, mesothelial cells, neurons, and glia (**Fig. 3f**), without any apparent “doublet” clusters (**Materials and Methods**), indicating that RAISINs correspond to single nuclei rather than to cellular aggregates. Therefore, even though

RAISIN RNA-seq captures RNA both inside and outside the nuclear envelope, it is compatible with droplet-based scRNA-seq and yields little observed contamination.

RAISIN RNA-seq survey of the ENS from adult mice identifies 24 neuron and 3 glia subsets

We used RAISIN RNA-seq with SMART-Seq2 to profile 5,181 high-quality transcriptomes from the ENS of 24 adult mice, spanning a range of ages (11-52 weeks), both males and females, and two phases of the circadian rhythm (morning or evening), and dividing each colon specimen into four equally sized segments along the proximal-distal axis to capture differences in anatomical location (**Fig. 1; Supp. Table 2; Materials and Methods**). We initially used *Wnt1-Cre2;INTACT* and *Sox10-Cre;INTACT* mice to label both neurons and glia, and *Uchl1-H2B mCherry* mice to enrich for enteric neurons (**Fig. 2a,b**); however, because the *Wnt1-Cre2* driver also targeted epithelial cells (**Fig. 2a**), we subsequently focused on the other two transgenic mouse models.

Among the 5,181 transcriptomes, we identified 2,447 neurons and 2,710 glia, with an average of 7,491 and 4,732 genes detected per RAISIN, respectively, which partitioned into 24 and 3 subsets, respectively (**Fig. 4a-e**). The clusters were enriched for markers of neuron and glia transcriptomes from scRNA-seq studies (**Fig. 4a,b**) (44, 45), with no detectable epithelial or enteroendocrine contamination, except for 8 contaminating cells in the “Other 2” cluster (**Fig. 4a-c**). Neurons and glia clustered primarily by cell subsets, rather than by mouse, intestinal region, or other known technical covariates (**Supp. Fig. 3a,b**). We estimate that enteric neurons comprise less than 1% of all nuclei in the murine colon after adjusting the numbers of FACS-sorted nuclei by the proportions of neurons identified in each mouse model (**Supp. Fig. 3c**).

Major neuron subsets are distinguished by expression of canonical neurotransmitters

We next arranged neurons into a hierarchy (**Fig. 4f**), and annotated it *post-hoc* using canonical ENS marker genes from the rich area of enteric research (*e.g.*, Chat, Cck, Elavl3, Elavl4, Gal, Grp, Nmu, Nos1, Npy, Pdyn, Penk, Phox2b, Ret, Snap25, Sst, Tac1, Uchl1, and Vip (46–50); discussed below), as well as data-driven genes that were differentially-expressed between clusters (**Fig. 4b,f; Supp. Table 2,3; Materials and Methods**). We validated many of the data-driven markers *in situ* (discussed below).

Broadly, neurons partitioned into either cholinergic (Chat⁺) or nitrergic (Nos1⁺) subsets (**Fig. 4f**, Ach and NO producing, respectively), except for four subsets expressing both Chat and Nos1 (defined as $\log_2(\text{TP10K}+1) > 0.5$), which we validated *in situ* (**Supp. Fig. 5a**), and one subset that expressed neither marker. Based on expression of known marker genes, we defined putative neurons subsets (**Fig. 4d,e; Supp. Table 2,3; Materials and Methods**), including: (1) Chat⁺Tac1⁺ excitatory motor neurons (PEMNs; 6 subsets), and (2) Nos1⁺Vip⁺ inhibitory motor neurons (PIMNs; 7 subsets), which together coordinate muscle contraction and relaxation; (3) CGRP⁺ sensory neurons (PSNs; 4 subsets), which sense and respond to chemical and mechanical signals in the intestine; (4) interneurons (PINs; 3 subsets), which relay signals between neurons; and (5) Glp2r⁺ secretomotor and vasodilator neurons (PSVNs; 2 subsets), which trigger secretions and fluid movement in other cell types.

The only major marker that we could not detect was the neuronal enzyme for serotonin synthesis, Tph2 (51, 52). We probed for Tph2 *in situ* in the colon as well as targeted brain regions, which served as positive (raphe nuclei) and negative (pontine reticular nucleus) controls (**Supp. Fig. 4a-**

e), but only observed Tph2 signal in the brain. We considered the possibility that Tph2-expressing enteric neurons are rare (53, 54), and examined published bulk RNA-seq data (55), finding Tph2 expression in the brain, but not the colon (**Supp. Fig. 4f**). Lastly, an independent scRNA-seq study of the small intestine myenteric plexus did not yield serotonergic neurons (56). However, we cannot exclude the possibility that Tph2 is expressed only under different physiological conditions, in other locations, or cannot be captured using current genomic and RNA-FISH tools. One possibility is that serotonergic neurons only populate the small intestines, as conditional *Tph1* knock-out mice crossed with a *Villin-Cre* driver, which lack serotonin production by the mucosa, have detectable serotonin in the duodenum and jejunum; although these regions still had detectable *Tph1* mRNA in the conditional knock-out (57).

ENS composition and expression programs vary by region and with circadian oscillations

To systematically assess sources of variation in the ENS, we leveraged the fact that our atlas comprises samples that vary by genetic background, age, sex, circadian phase, and intestinal location, to test how each factor impacts ENS composition (*i.e.* the relative proportions of neuron subsets) or gene expression within each neuron subset.

The transgenic background had profound effects on neuron composition (**Fig. 4f**), suggesting distinct developmental origins for some neuron subsets. In particular, two subsets of putative sensory neurons (PSN1 and PSN2) were nearly absent from *Sox10-Cre* mice (**Fig. 4f**), suggesting they may arise from distinct lineages (58, 59). ENS composition also varied significantly along the length of the colon within each of the *Sox10-Cre; INTACT* and *Uchl1-H2B mCherry* lines, with distinct neuron subsets enriched in different regions (**Fig. 4f**). For example, PSN1 and PSN2

were enriched in the proximal colon ($P < 10^{-22}$ and 10^{-6} , respectively; Fisher's exact test), whereas distinct subsets of putative motor neurons (PMNs) were enriched in either the proximal or distal colon (**Fig. 4f**).

We next used a regression framework to identify genes that were differentially expressed (DE) with respect to age, sex, circadian phase, and colon location, in a manner shared across neuron subsets (**Materials and Methods**). Overall, few DE genes were associated with age or sex beyond genes on the sex chromosomes (**Supp. Table 3**); however, the circadian clock and colon location had substantial impacts on gene expression of many neuron genes (**Supp. Table 3**). For example, core clock regulators were among the most DE genes during morning (*Arntl*) and evening (*Per1*, *Per2*, *Per3*) (**Fig. 5a**). In the morning, there was also increased expression of cytoskeleton-associated genes (*e.g.*, *Tubb3*, *Prph*, *Tubb2a*, *Cfl1*), suggesting circadian regulation of structural remodeling (60), and genes involved in neuronal signaling (*e.g.*, *Scg2*, *Pcsk1n*, and *Slc7a11*). In PSN1 and PSN2, we also observed morning upregulation of genes involved in neuro-immune signaling (*e.g.*, *Calcb*, *Il13ra1*) (**Fig. 5b; Supp. Table 2**) (61, 62). In the evening, several TFs were upregulated relative to morning, including *Nr1d2*, *Tef*, *Rfx2*, and *Dbp* (**Fig. 5a; Supp. Table 2**), many of which are known circadian regulators (63).

In addition, there were significant changes in gene expression across colon regions (**Fig. 5c; Supp. Table 2**), after controlling for differences in ENS composition (which itself varies by location; **Fig. 4f**). Most notably, neurons in the distal mouse colon had higher expression of several neurotransmitter receptors, including serotonin receptors (*Htr3a*, *Htr3b*), glutamate receptors

(Gria3, Grid1), acetylcholine receptors (Chrna7, Chrm1), and potassium and sodium channels (Kcnq5, Scn5a), suggesting electrophysiological differences along the ENS.

Motor neuron expression profiles suggest that mechanosensation drives the peristaltic reflex

The myenteric plexus is a major functional unit of the ENS, moving luminal contents along the intestine through coordinated muscle contraction and relaxation (64). The canonical model of the peristaltic reflex (**Fig. 5d**) (65–67) begins with the release of serotonin (5HT) by enterochromaffin cells in the mucosa, which acts on sensory neurons. Interneurons then relay this signal to ascending and descending motor neurons, which elicit muscle contraction and relaxation, respectively. This model is based on associations between 5HT stimulation and muscle contraction (initially described (68, 69), but was subsequently questioned because (1) removal of the mucosa does not ablate contraction (70–74), (2) the timing of serotonin release does not precede contraction (75), and (3) deletion of *Tph1* (mucosal serotonin synthetic enzyme) has no effect on GI transit *in vivo* (76); although the knock-out may have been incomplete (77). We therefore hypothesized that the molecular signatures of neuron subsets could help build and test models of peristalsis.

The transcriptional profiles of putative motor neurons suggest possible revisions to the peristaltic model, with a potential role for direct mechanosensation of gut distention by motor neurons in driving peristaltic reflexes. First, nearly all putative motor neurons express the mechanosensitive ion channel, Piezo1 (**Fig. 5e**, PEMNs and PIMNs; confirmed *in situ*, **Supp. Fig. 5b**), suggesting they have the capacity to directly sense distention. Previous studies reported Piezo1 expression in the ENS (78, 79), which our atlas allows us to map to specific subsets. This refinement raises the

hypothesis that peristalsis is at least partially driven by distention, specifically via motor neuron depolarization through Piezo1.

Moreover, 5HT receptor 4 (Htr4) expression on enteric neurons is well-documented and its expression has been attributed to sensory or secretomotor neurons (80, 81); also possibly expressed more broadly (82). We find Htr4 expression in two putative sensory neuron (PSN) subsets and all putative excitatory motor neuron (PEMN) subsets (**Fig. 5e**); we confirmed this *in situ* in Chat⁺ neurons of the myenteric plexus (**Supp. Fig. 5c**). Because Htr4 activation facilitates acetylcholine release in enteric neurons (83) we hypothesize a model where distention evokes motor neuron activity (through Piezo1) and mucosal serotonin enhances this distention-evoked stimulation of muscle contraction. This model was previously proposed for sensory neurons (83) and our atlas has allowed its refinement to excitatory motor neurons.

Sensory neurons express key regulators of ILC responses and tissue homeostasis

We identified four subsets of putative sensory neurons (PSNs) by expression of calcitonin gene-related peptide (CGRP), a marker of sensory neurons expressed in two forms (Calca, Calcb), which is involved in feeding, pain sensation, hormone secretion, and inflammation (84). While all four subsets express Calcb, only PSN3 expresses Calca at significant (but low) levels (**Fig. 4f; Supp. Fig. 6a**), which we confirmed *in situ* (**Fig. 5f; Supp. Fig. 5d**). The CGRP receptor (Calcrl) and one of its three co-receptors (Ramp1) are expressed in all neurons, except putative secretomotor neurons (**Supp. Fig. 6a**).

We inferred the likely target cells for each PSN subset based on the signaling molecules and receptors that they express (**Fig. 4f; Supp. Fig. 6a,b; Supp. Table 3**). For example, most sensory neuron subsets express receptors for glucagon (Gcgr), glucagon-like peptide 1 (Glp1r), and galanin (Galr) (**Fig. 4f; Supp. Fig. 6a**), peptides that are produced by enteroendocrine cells with roles in hunger and satiety (85). One subset, PSN3, co-expresses Cck and Vip (**Fig. 5e**), markers of intestinofugal neurons that innervate the prevertebral ganglia (86), thus supporting connections to the sympathetic nervous system. This subset also uniquely expresses brain-derived neurotrophic factor (Bdnf, **Supp. Fig. 6b**), which is elevated in patients with irritable bowel syndrome (IBS), where it is correlated to abdominal pain (87), and Piezo2 (**Fig. 5e**), a mechanosensitive ion channel, which may help detect and regulate smooth muscle tone (88); confirmed *in situ*; **Supp. Fig. 5e**). Another Calcb⁺ subset, PSN4, uniquely expresses somatostatin (Sst; **Fig. 4f; Supp. Fig. 6b**; validated *in situ*, **Supp. Fig. 5f**), previously attributed to interneurons (89); the role of SST in the GI tract is poorly understood, but has been broadly linked to regulating most GI functions, including motility, secretion, absorption and the sensation of visceral pain (90). Localization of Sst expression to a single neuron subset will help dissect its function in the ENS.

One sensory neuron subset, PSN1, uniquely expresses Noggin (Nog) and Neuromedin U (Nmu) (**Fig. 4f; Supp. Fig. 6b**; validated *in situ* **Fig. 5g,h**), and is rarely marked by the *Sox10-Cre* driver (**Fig. 4f**; validated *in situ*; **Fig. 5i**): these genes are known regulators of epithelial stem cells (91) and immune cells (92), respectively. In particular, Noggin is a BMP antagonist that is necessary for maintaining the intestinal stem cell niche, but whose cellular source is unknown. Noggin expression by sensory neurons raises the hypothesis that these neurons could help regulate the positioning or differentiation of intestinal stem cells. Furthermore, the neuropeptide NMU

regulates type 2 cytokine responses via activation of innate lymphoid cells (ILCs) (93). Expression of its receptors, Nmur1 and Nmur2, on excitatory motor (PEMN1, PEMN2; **Supp. Fig. 6a**) and sensory (PSN1, PSN2, PSN3; **Supp. Fig. 6b**) neurons, respectively, suggests diverse neuronal targets of NMU, that may help orchestrate inflammation. PSN1 cells also express additional genes that may interact with ILCs, including Calcb, both subunits of the Il-13 receptor (Il4ra and Il13ra1, **Supp. Fig. 6a**), and Il-7 (**Supp. Fig. 6b**), a major regulator of ILC differentiation and survival (94). Lastly, both PSN1 and PSN2 cells express gastrin-releasing peptide (Grp, **Fig. 4f**), which in the lung is produced by neuroendocrine cells and contributes to the response to tissue injury (95).

Secretomotor neurons may integrate epithelial and immune signals

Secretomotor/vasodilator neurons (SVNs) integrate signals from the mucosa and sympathetic ganglia to regulate fluid movement between the body and the lumen. We identified two subsets of putative secretomotor/vasodilator neurons (PSVNs) corresponding to non-cholinergic (PSVN1) and cholinergic (PSVN2) subtypes (96) (**Fig. 4d,f**). Both subsets uniquely express receptors for GLP-2 (Glp2r) and secretin (Sctr), hormones released by enteroendocrine cells that stimulate blood flow (97) and epithelial secretions (98), respectively (**Fig. 4f; Supp. Fig. 6a**). Most local reflexes regulating water and electrolyte balance likely act through non-cholinergic SVNs (99), and our data suggest that cholinergic SVNs may support tissue homeostasis. Specifically, the GM-CSF receptor (Csf2rb, Csf2rb2, **Supp. Fig. 6b**) and Thymic Stromal Lymphopoietin (Tslp, **Supp. Fig. 6a**) are expressed by PSVN2s, suggesting these neurons participate in GI immune responses (100, 101).

Profiling the human muscularis propria using RAISIN RNA-seq

Next, we profiled human colon enteric neurons. Unlike genetic mouse models, we could not enrich for nuclei from human enteric neurons, and thus opted to profile the muscularis propria (MP), which has a higher proportion of neurons than the submucosa or mucosa. We isolated and profiled nuclei from cancer-adjacent normal colon segments from colorectal cancer resections from both genders (5 male, 5 female) and a range of ages (35 - 90) (**Supp. Table 4**). Based on our mouse data (**Supp. Fig. 3c**), we conservatively estimated a 0.5% capture rate for neuron nuclei, such that in order to capture 500 human neurons, we would need to profile at least 100,000 unsorted nuclei.

Profiling 134,835 human RAISINs from the muscularis propria recovered transcriptomes from neurons, resident adipocytes, endothelial cells (lymphatic, vascular), fibroblasts, glia, immune cells (macrophages, mast cells, lymphoid cells), interstitial cells of Cajal (ICCs), myocytes, and pericytes (**Fig. 6a**), each annotated by expression of marker genes (**Fig. 6a; Supp. Table 4**). Some subsets were enriched in specific patients (**Supp. Fig. 7a-f**), which may be due to differences in sampled locations, variable cellular states or variation in the sampling of rare cells. Additionally, human RAISIN RNA-seq data contained more background contamination than either mouse RAISIN SMART-Seq2 or droplet data, possibly due to delayed tissue freezing time following resection.

Human enteric neurons cluster into 11 subsets with distinct transcriptional programs

The 134,835 RAISINs include 831 human enteric neurons (0.6%), which clustered into 11 subsets (**Fig. 6b**) after correcting for putative differences in cell quality (**Supp. Fig. 7g-j; Materials and**

Methods). The neuron recovery rate in humans slightly exceeded our original estimate, likely because the muscularis propria is somewhat enriched for neurons relative to the rest of the colon.

Although we detect many hallmark neurotransmitters, CHAT was lowly expressed (**Supp. Fig. 8a**), either due to actual low expression in human cells, reduced levels in the nucleus, an alternative 3' end compared to the current annotation (102), or cancer-adjacent effects. We do detect SLC5A7 (**Supp. Fig. 8a**), a transporter that mediates choline uptake into cells, and is known to be expressed with CHAT in cholinergic neurons (103), and is co-expressed with CHAT in human neurons in our data (Pearson's $r=0.60$); likewise, Slc5a7 is co-expressed with Chat in mouse neurons. We therefore used SLC5A7 as a surrogate marker for CHAT in human neurons. Interestingly, we observed broad, albeit low, levels of expression of tryptophan hydroxylase 2 (TPH2; required for serotonin biosynthesis) across almost all human neuron subsets (**Supp. Fig. 8b**), but not in mouse neurons (above, **Supp. Fig. 4**), suggesting differences in serotonergic signaling between the two species.

Human ENS contains sensory, motor, interneuron, and secretomotor/vasodilator subsets that share core transcriptional programs with mouse

We used a classification-based approach (**Materials and Methods**) to map the 11 subsets of human neurons onto the 24 mouse subsets (**Fig. 6c**), leveraging the larger number of cells and deeper sequencing data in mouse to annotate the human cells. The neurons partitioned into 2 PEMN subsets, 5 PIMN subsets, 1 PSN subset, 2 PIN subsets, and 2 PSVN subsets (**Fig. 6b**), annotated with known markers (**Supp. Fig. 8a,b**). Despite representing distinct regions of the colon (*i.e.* full colon vs. muscularis propria), both species contained similar neuron compositions,

with excitatory and inhibitory motor neurons being the most abundant classes (**Fig. 6c**). However, sensory neurons were more abundant and more diverse in mouse: humans contained only one sensory subset, whereas mice contained four. This may be due to removal of the human submucosa, although we cannot entirely rule out the possibility that the different number of profiled neurons may contribute to this difference as well. Furthermore, while the fraction of secretomotor/vasodilator neurons was similar across both species, the human muscularis propria lacked the cholinergic subtype, as assessed by CHAT and SLC5A7 expression, whereas mice contained both cholinergic and non-cholinergic subsets.

We leveraged the human-mouse mapping to identify conserved (core) programs for each of five major neuron types (**Fig. 6d; Supp. Table 5; Materials and Methods**). For example, the core transcriptional program for excitatory motor neurons ($n = 75$ genes) includes acetylcholine, various receptors (*e.g.*, GFRA2, OPRK1, HTR4), solute transporters (*e.g.*, SLC5A7), transcription factors (*e.g.*, CASZ1), and COLQ, which tethers acetylcholinesterase within the neuromuscular junction (104) (**Fig. 6d; Supp. Fig. 8a; Supp. Table 5**). In addition, human PEMNs uniquely express the mechanosensitive ion channel, PIEZO2 (**Supp. Fig. 8a**), whereas mice express Piezo1 (**Fig. 5e**). Similarly, we defined core transcriptional programs for inhibitory motor neurons ($n = 89$ genes; *e.g.*, VIP, NOS1, CARTPT, GFRA1, OPRD1, ETV1), sensory neurons ($n = 76$ genes; *e.g.*, CALCB, NMU, NOG, SST, VIPR2), interneurons ($n = 57$ genes; *e.g.*, PENK, TAC1, ADRA2A), and secretomotor/vasodilator neurons ($n = 46$ genes; *e.g.*, VIP, GAL, SCGN, CALB2) (**Fig. 6d; Supp. Table 5; Materials and Methods**).

Human Interstitial Cells of Cajal (ICCs) may underlie smooth muscle relaxation

Our reference map of the human muscularis propria includes 431 KIT⁺ANO1⁺ ICCs (**Fig. 6a**), which are regarded as pacemaker cells that rhythmically alter the excitability of smooth muscle tissue (105, 106). Two major models have been proposed for ICC function (107): either (1) neurons signal directly to smooth muscle, with an indirect role for ICCs (*e.g.*, to generate motor patterns), or (2) neurons signal to ICCs, which then relay signals to smooth muscle to coordinate peristalsis.

To help distinguish between these possibilities, we defined a gene signature for ICCs (**Fig. 6e**) and mapped known ligand-receptor pairs onto neurons, ICCs, and smooth muscle cells (**Materials and Methods**). Although motor activity requires both excitatory (*i.e.* cholinergic) and inhibitory (*i.e.* nitrergic) signals to elicit contraction and relaxation, respectively, smooth muscle cells only expressed the receptors for acetylcholine (**Fig. 7a**). In contrast, the receptor for nitric oxide was expressed by ICCs (**Fig. 7a**), which we validated *in situ* (**Fig. 7b**). As a positive control, we note that nitric oxide receptors are detected in pericytes and neurons (**Fig. 7a**) (108). These results suggest a revised model of smooth muscle function, where enteric neurons directly activate smooth muscle contraction, but elicit smooth muscle relaxation indirectly via ICCs (**Fig. 7c**). Consistent with this hypothesis, smooth muscle-specific knockout of the $\beta 1$ subunit of the nitric oxide receptor only partially reduces relaxation, whereas its global knockout nearly abolishes relaxation (109).

Enteric neurons can interact with diverse stromal and immune cells in the colon

To systematically examine interactions between the enteric nervous system and other cell types in the human colon, we analyzed profiles from the 134,835 RAISINs from the muscularis propria (above) together with scRNA-Seq profiles from 115,517 cells from the colon mucosa (*i.e.*

epithelium and lamina propria) (110). In total, these data span a wide range of cell types in the human colon, including 16 epithelial subsets, 26 immune subsets (myeloid and lymphoid), 7 endothelial subsets, 9 fibroblast subsets, myocytes, ICCs, resident adipocytes, 2 glia subsets (muscularis propria and lamina propria), and 11 neuron subsets. We mapped thousands of receptor-ligand pairs onto this dataset and identified pairs of cell subsets expressing a significantly greater number of cognate receptor-ligand pairs than is expected under a null model (**Fig. 7d; Materials and Methods**).

Broadly, neurons were enriched for putative interactions with other cells from the muscularis propria rather than from the mucosa, suggesting the recovery of local interactions. This approach highlighted known interactions between excitatory motor neurons and smooth muscle (111), secretomotor/vasodilator neurons and both epithelial cells (*i.e.* tuft and enteroendocrine) and lymphatics (112), and glia and multiple subsets of neurons (**Fig. 7d**).

More unexpectedly, we found statistically enriched interactions between neurons and diverse stromal cells, most notably resident adipocytes and fibroblasts (**Fig. 7d,e**), the two largest producers of neurotrophic growth factor (NGF) outside of the ENS in our data (**Supp. Table 4**). Potential enteric neuron signaling to resident adipocytes spanned neuropeptides that regulate appetite and energy metabolism (CGRP/CALCRL, NPY/NPYR1) (113, 114), and two neurotransmitters (glutamate/GRM8, GABA/GABRE) (**Fig. 7e**). Resident adipocytes reciprocally are inferred to signal to neurons via the leptin pathway, with all neuron subsets expressing the leptin receptor (LEPR) (**Fig. 7e**). In addition, inferred neuron signaling to fibroblasts included neuropeptides (PACAP/VIP/VIPR2) (**Fig. 7e**), neurotransmitters (glutamate/GRIA4, nitric

oxide/GUCY1A3), growth factors (FGF1/FGFR1, PDGF/PDGFRB), guidance cues (SLIT2/ROBO1, SLIT3/ROBO2), and IL15/IL15R (**Fig. 7e**).

Even if cell subsets are not enriched for interactions, they may still interact through a more limited, but functionally important, receptor-ligand repertoire. Given recent reports describing neuro-immune crosstalk (115), we searched for specific examples of interactions between neurons and immune cells (**Fig. 7e**). We identified potential neuron signaling to (1) T cells via IL7/IL7R, IL12A/IL12RB1 (neuronal expression validated *in situ*, **Fig. 7f,g**), and PENK/OPRM1, (2) dendritic cells via CHAT/CHRNA, and (3) B cells via TPH2/HTR3A (**Fig. 7e**). Both IL-7 and IL-12 have key roles in lymphocyte and ILC survival and Th1 polarization (116), suggesting key pathways by which enteric neurons may regulate adaptive immunity. Finally, human PSN1s express NMU, which in mouse activates ILC2s (117).

Human enteric neurons express risk genes for enteric neuropathies, intestinal inflammatory disorders, and extra-intestinal disorders with GI dysmotility

To interrogate potential contributions of the ENS to human diseases, we examined whether enteric neurons expressed any genes associated with diseases with varying degrees of known ENS involvement. These ranged from Hirschsprung's disease (HSCR), a primary enteroneuropathy that directly affects the ENS, to autism spectrum disorder (ASD) and Parkinson's disease (PD), which are extra-intestinal CNS disorders that are associated with dysfunctions in gut motility that occur early in disease progression (118–120). In addition, because the ENS is thought to play a pivotal role in inflammation – for example, through the activation of ILCs (121) – we also examined whether IBD-associated genes are expressed by enteric neurons.

Mapping a curated list of 185 disease-associated genes (**Materials and Methods**) onto cell subsets from the muscularis propria, lamina propria, and epithelium (as above), we identified many genes that were specifically enriched in enteric neurons (**Fig. 8a**). For example, even though it is a neurodevelopmental disorder, most HSCR-associated genes were expressed in adult enteric neurons, including RET, PHOX2B, GFRA1, ZEB2, and ECE1 (**Fig. 8a**). The two exceptions, EDN3 and EDNRB, mediate endothelin signaling in the embryonic neural crest (122). Although most IBD risk genes are expressed in epithelial and immune cells, a subset of genes were most highly expressed in neurons, including GRP, BTBD8, KSR1, NDFIP1, and REV3L (**Fig. 8b**). In particular, GRP products stimulate GI hormone release, muscle contraction, and epithelial cell proliferation (123). Another such gene, REV3L, is also perturbed in the craniofacial neurologic disorder Möbius syndrome (124). Indeed, increased expression of many neuropeptides (*e.g.*, tachykinin and galanin) has been reported in IBD patients (125).

The risk genes for CNS diseases with concomitant GI dysfunction were predominantly expressed in enteric neurons, with some notable exceptions in ASD and PD (*e.g.*, P2RX5 and IL1R2 in B cells and epithelial cells, respectively) (**Fig. 8c**). CNS disease risk genes that mapped specifically to enteric neurons include ANK2, DSCAM, and NRXN1 for ASD, and DLG2, SCNA and SCN3A for PD (**Fig. 8c**). Expression of these risk genes specifically by enteric neurons, compared with a colon reference map, motivate further investigation of the role that enteric neurons play in the development and progression of dysmotility in intra- and extra- intestinal disorders; it also suggests that studying the ENS, in the far more accessible GI, can provide a window for investigation of GWAS genes of some CNS disorders.

Discussion

Here, we constructed reference maps of the colon ENS of adult mice and humans at single cell resolution, revealing the broad potential of neurons to orchestrate tissue homeostasis. Isolating individual enteric neurons from adult animals for transcriptional profiling has not been possible due to technical limitations, and recent efforts using whole-cell dissociations have been limited to embryonic or post-natal animals (126, 127). Our development of RAISIN and INNER Cell RNA-seq, which preserve ribosome-attached RNA on intact nuclei, allowed us to profile 2,447 mouse and 831 human enteric neurons, along with other diverse cell types from both species (*e.g.*, epithelial, stromal, and immune cells). These methods can be applied to both fresh and frozen tissue specimens, opening the way to characterizing the ENS and a range of archived frozen tissue samples. Additionally, preservation of the ER on nuclei may allow for the enrichment of nuclei with antibodies targeting specific membrane proteins, which are synthesized in the ER.

We identified all major classes of enteric neurons, spanning 24 mouse subsets and 11 human subsets, including motor, sensory, secretomotor/vasodilator and interneuron types. Mining their expression signatures allowed us to infer signaling among neurons and between neurons and non-neuronal cells, such as resident adipocytes, ICCs, immune cells, and epithelial cells. We show circadian regulation of the ENS, including core clock genes, motivating further investigation into temporal variation of ENS function, nutrient absorption, and metabolism (128). We also show differences in neuron composition across the mouse colon (*e.g.*, sensory neurons enriched in the proximal colon) suggesting that ENS function varies along the length of the GI tract. Comparison of mouse and human neurons allowed us to derive core transcriptional signatures for subsets across species, highlighting biological processes that can be modeled in mouse; for example, sensory

neurons in both species express Noggin, a gene known to support the epithelial stem cell niche (129). Taken together, these data enable the generation of testable hypotheses and experimental dissection of ENS function.

Finally, given the extensive potential for neuro-immune signaling we observe in the mouse and human ENS, we propose that neuronal dysfunction can lead to immune dysregulation, which can exacerbate inflammation and related pathologies. For example, several IBD risk genes are expressed in neurons, raising the need to further characterize the role of enteric neurons in intestinal inflammation. Intriguingly, dozens of risk genes for early-life and late-onset CNS disorders with concomitant gut dysmotility are highly expressed by enteric neurons suggesting a mechanism for gut motility dysfunction in these diseases, and that profiling the much more accessible ENS may allow us to study human disease biology. Furthermore, recent associations between the gut microbiota and extra-intestinal diseases, such as autoimmune disorders (reviewed in (130)) and cancers and cancer therapies (reviewed in (131)), suggest that immune modulation in the gut can have systemic effects. Proper immune function is thought to be necessary for CNS maintenance and repair, with immune dysregulation contributing to neurodegenerative disease (reviewed in (132)). Thus, the ENS may be a central nexus linking the gut, the immune system and the brain, and neurological dysfunction in the gut may exacerbate diseases of the CNS.

Acknowledgements

We thank Karin Pelka, Jonathan Chen, and Nir Hacohen for meaningful discussions and coordination of donor sample transport and banking. We thank Asaf Rotem and Isaac Wakiro for help with donor tissue transport. We thank Louise Trakimas and Margaret Coughlin from the Harvard Medical School Electron Microscopy Facility for help with processing nuclei samples for TEM. We thank Patricia Rogers, Chelsea Ortis, and Stephanie Saldi of the Broad Institute Flow Cytometry Core for help with FACS. We thank Leslie Gaffney and Anna Hupalowska for help with figure graphics and illustrations. We thank Fei Chen for access to microscope facilities and thank Leah Caplan for help with microscopy. We thank Vijay Kuchroo and Ana C. Anderson for support and discussions. Work was supported by the Klarman Cell Observatory (AR, RJX), HHMI (AR), the Manton Foundation (AR and RX), Broad*next*10 (AR, RJX), RC2 DK114784 (AR, RJX), DK043351, DK114784, DK117263, Helmsley Charitable Trust and Crohn's and Colitis Foundation (RJX), and the Food Allergy Science Initiative (AR, RJX). AR is a co-founder and equity holder of Celsius Therapeutics and an SAB member of ThermoFisher Scientific and Syros Pharmaceuticals. RJX is a co-founder and equity holder of Celsius Therapeutics and Jnana Therapeutics and a consultant to Novartis. This publication is part of the Human Cell Atlas - www.humancellatlas.org/publications/

Figure legends

Figure 1. Study overview.

Left: RAISIN RNA-seq. Nuclei extraction conditions were varied for the extraction of nuclei from the colons of *Sox10-Cre*; *INTACT* mice. 104 nuclei extractions were FACS enriched for labeled nuclei and transcriptomes were profiled using Smart-Seq2. Extractions yielding high quality transcriptomes were evaluated using ultra-thin section TEM, revealing intact nuclei with preserved nuclear envelope and portions of the rough endoplasmic reticulum with attached ribosomes, termed “RAISINs”. RAISINs were compatible with massively parallel droplet-based RNA-seq. Middle: an adult murine enteric neuron and glia reference map at a single cell resolution. Tagged RAISINs from the colons of three transgenic murine models were extracted, FACS enriched, and transcriptionally profiled. The age, gender, colon location and circadian phase were systematically varied to capture their effects on ENS heterogeneity. Right: human single cell atlas of the colon muscularis propria including enteric neurons and interstitial cells of Cajal (ICCs). RAISINs were extracted from the colon muscularis propria of 10 donors (cancer-proximal normal tissue; 5 female and 5 male donors ranging from 35-90 years of age) and profiled using massively parallel droplet-based RNA-seq. The transcriptomes of 163,741 single RAISINs were recovered, including 831 enteric neurons. Expression of cognate receptor-ligand pairs among cell subsets was used to model local tissue signaling. Disease risk genes for intestinal disorders were mapped onto cell subsets from both the colon muscularis and lamina propria.

Figure 2. Mouse model validation.

(a) Labeling of nuclei in the colons of ENS reporter mice. Representative images show cross-section with muscularis propria (oriented bottom) and mucosa (oriented top) (left). All models show labeled nuclei in TUBB3+ neurons; *Wnt1-Cre2*; *INTACT* mice also have labeled nuclei in the mucosa. Representative FACS plots of extracted nuclei show enriched populations (right). (b) Validation of the *Sox10-Cre* driver. Triple-transgenic mice harboring *Sox10-Cre*; *INTACT*; conditional *tdTomato* alleles were used to evaluate concordance of genetically labeled cells and TUBB3 immunofluorescence.

Figure 3. RAISIN RNA-seq captures RNA from intact nuclei and associated ribosomes.

(a) Decision tree for selection of optimal nuclei extraction conditions. (b,c) Optimization of RAISIN and INNER Cell RNA-Seq ($k = 5,236$ GFP+ sorted nuclei across all protocols). (b) Cellular composition of each nuclei extraction. The proportion of nuclei annotated as neurons, glia, or other cell subsets (triangle edges) from each extraction type (dots). Select nuclei extractions are marked in color (legend) (c) ENS signature score (y axis, $\log_2(\text{TP10K}+1)$); mean and standard error of the mean (SEM); **Materials and Methods**) and number of detected genes per nucleus (x axis, mean and SEM) for each of 36 total conditions. Dot size: percent neurons captured. Select nuclei extractions are marked in color (legend). (d) RAISIN and INNER Cell RNA-seq isolate nuclei with attached ribosomes and rough ER. Ultra-thin section transmission electron microscopy (TEM) of nuclei extractions from published methods (top) (22, 23) and with RAISIN (bottom left) and INNER Cell (bottom right) methods. (e) Higher exon:intron ratios in RAISIN and INNER Cell methods. Exon:intron ratio (y axis, $\log_2(\text{ratio})$) following snRNA-seq from each preparations in (d). All comparisons are significant (Wilcoxon test, $p\text{-value} < 10^{-10}$); boxplots: 25%, 50%, and 75% quantiles; error bars: standard deviation (SD). (f) RAISIN RNA-seq is compatible with droplet-based RNA-seq. A t-distributed stochastic neighbor embedding (t-SNE) of RAISIN RNA-seq profiles from mouse colon of 10,889 unsorted RAISINs profiled by droplet-based scRNA-seq and colored by cell type.

Figure 4. Mouse ENS atlas reveals 24 neuron and 3 glia subsets.

(a-c) High quality neuron and glia transcriptomes. Mean expression levels ($\log_2(\text{TP10K}+1)$) of hallmark genes (x axis) across cell subsets (y axis) for major cell classes (a), neuron subsets (b), or glia subsets (c). Cell subsets were profiled using either Smart-Seq2 (SS2) or droplet-based methods. **(d)** Mouse neuron reference: 24 neuron subsets profiled by RAISIN RNA-seq. t-SNE of 2,447 neuron RAISIN RNA-Seq profiles from mouse colon colored by putative neuron classes based on *post hoc* annotation (**Materials and Methods**). **(e)** Mouse glia reference: 3 glia subsets profiled by RAISIN RNA-seq. t-SNE of 2,710 glia RAISIN RNA-Seq profiles from mouse colon colored by cluster. **(f)** Neuron subsets vary by anatomical location and mouse line. Neuron subsets (columns) arranged by transcriptional similarity (dendrogram, top) and annotated with the proportion of cells isolated from each transgenic model (green pie chart) or colon segment (red/blue pie chart). Dot plot shows for select neurotransmitters or neuropeptides, the fraction of cells in each subset (dot size) expressing the synthetic enzyme (top) or receptors (bottom; genes for synthesis and receptors in **Supp. Table 3**), and the mean expression level in expressing (non-zero) cells within the subset (dot color).

Figure 5. Mouse ENS atlas reveals neuron subsets vary with circadian phase, colon location and mouse model.

(a,b) Mouse ENS gene expression is affected by circadian rhythm. Distribution of neuron gene expression levels (y axis, $\log_2(\text{TP10K}+1)$) of select genes (x axis) that are upregulated at morning (red) or evening (blue) time points in all neurons (a), or at the morning time point in PSN1s and PSN2s (b). (c) Changes in ENS expression along colon length. Mean expression levels ($\log_2(\text{TP10K}+1)$) across all neuron subsets (color bar) of significantly DE genes (columns) across colon regions (rows), arranged by location of peak expression from proximal to distal. (d) Current model of the peristaltic circuit (adapted from (133)). (e) The mechanosensitive ion channel *Piezo1* is expressed in PIMNs and PEMNs. Distribution of gene expression levels (y axis, $\log_2(\text{TP10K}+1)$) across neuron subsets (x axis) for genes in peristaltic model: *Htr4* (top), *Piezo1* (middle) and *Piezo2* (bottom). (f-h) Validation of gene expression *in situ*. Representative images of smFISH for *Calcb* and *Nos1* showing extensive *Calcb* (CGRP) expression in enteric neurons (f), *Calcb* and *Nmu* showing *Nmu* expression in putative sensory neurons, and (g) *Nog* and *Grp* showing *Noggin* expression in enteric neurons, all with TUBB3 immunostaining (h). (i) Representative image showing that *Grp+Nog+* neurons (yellow inset; *i.e.* PSN1, PSN2) do not have a history of *Sox10-Cre* expression (GFP), whereas *Grp+Nog-* neurons (red inset; *i.e.* PSN4) are labeled by the *Sox10-Cre* driver (GFP).

Figure 6. Reference map of the human colon muscularis propria reveals 11 neuron subsets with roles in immunity and disease. (a) Census of the human muscularis propria. t-SNE of 134,835 RAISIN RNA-seq profiles from the muscularis propria of cancer-proximal macroscopically normal colon resections from 10 human donors, colored by cell type, annotated *post hoc*. (b) Enteric neuron census. t-SNE of 831 RAISIN RNA-seq profiles from enteric neurons, colored by subset, annotated *post hoc*. (c) Correspondence of human and mouse enteric neurons. Percent of neurons (dot size and color) from each human subset (rows) that match each mouse neuron subset (column) according to our classifier (**Materials and Methods**). (d) Transcriptional signatures conserved between mouse and human neuron subsets. The fraction of expressing cells (dot size) and mean expression level in expressing (non-zero) cells (dot color) of selected genes (columns) conserved between mouse (top) and human (bottom) for each of five major neuron classes (rows); full list in **Supp. Table 5**. (e) ICC gene signature. Fraction of expressing cells (dot size) and mean expression level in expressing (non-zero) cells (dot color) of select ICC marker genes (columns) across human cell subsets (rows).

Figure 7. Cell-cell interactions among the enteric nervous system and other cell subsets. (a-c) Characterization of ICCs in the colon. **(a)** ICCs and not myocytes express receptors for nitric oxide. Distribution of expression levels (x axis, $\log_2(\text{TP10K}+1)$) of acetylcholine (left) and nitric oxide (right) receptors across cell subsets (y axis). **(b)** *In situ* expression of select ICC markers in the human colon. **(c)** Proposed peristaltic circuitry. NO: nitric oxide. Ach: acetylcholine. **(d,e)** Inferred cell-cell interactions between human cells in the mucosa and muscularis propria. **(d)** Statistically significant interactions. Nodes: cell subsets, annotated by type (color) and colon location (bold: muscularis). Edges connect pairs of cell subsets with a significant excess of cognate receptor-ligand pairs expressed ($p < 0.05$) relative to a null model (**Materials and Methods**). **(e)** Select receptor-ligand interactions between neurons and adipocytes, fibroblasts, and immune cell subsets. **(f,g)** Representative *in situ* validations of IL-7 expression in NOS1⁺ neurons **(f)** and IL-12 expression in CHAT⁺ neurons **(g)**.

Figure 8. Human enteric neurons express risk genes for primary enteroneuropathies, IBD, and CNS disorders with concomitant gut dysmotility. Mean expression levels (scaled $\log_2(\text{TP10K}+1)$) across cell subsets (rows) of putative risk genes (columns) implicated by GWAS for Hirschsprung's disease (HRSC), inflammatory bowel disease (IBD), autism spectrum disorders (ASD), and Parkinson's disease (PD) (**Materials and Methods**), which were identified as cell subset specific in either (**a**) the colon mucosa, or (**b**) the colon muscularis propria.

Supplementary Figure 1. snRNA-seq optimization.

(a) snRNA-seq of GFP⁺ nuclei from *Sox10-Cre; INTACT* animals. Fraction (y axis) of identified cell-types (x axis) in samples obtained from the brain (grey) and colon (black) using two previously published snRNA-seq methods (40, 134). (b-d) Quality scores across all tested parameters. Quality metrics (columns, x axes) for a range of concentrations (y axes) across detergents (b), mechanical extraction procedures (c), and buffers (d).

Supplementary Figure 2. Nuclei extracted with different protocols. Representative phase contrast images of nuclei isolated using extractions with different detergents or extraction kits (grey, **Materials and Methods**) and buffers (blue), with varying detergent concentrations and additives (marked on image). All extractions were performed with the “chop” mechanical extraction (**Materials and Methods**) unless otherwise indicated.

Supplementary Figure 3. Reproducibility and validations for the mouse ENS atlas. (a,b)

Reproducible cell subset distributions across transgenic mouse lines and individual mice. t-SNE of RAISIN RNA-seq profiles of 2,447 neurons (a) and 2,734 glia (b) colored by cell subset (left), mouse model (middle), or donor mouse (right). (c) Neuron composition in colon. Percent of all cells in the colon that are neurons (y axis) as estimated by FACS (transgene expressing nuclei vs. unlabeled nuclei) and *post-hoc* adjustment based on RAISIN RNA-seq data.

Supplementary Figure 4. Detection of *Tp2* expression in the brain, but not colon.

(a) Schematic of coronal brain section. Raphe nuclei contain serotonergic (*Tph2*⁺) neurons and served as a positive control. The pontine reticular nucleus does not contain *Tph2*--expressing neurons and served as a negative control. (b,c) Representative images of smFISH for *Tph2* in the mouse brain (b) and colon (c) of *Sox10-Cre; INTACT* (GFP) mice (*n* = 2 animals; 12 colon sections). (d,e) Representative images of smFISH for *Tph2* in the mouse brain (b) and colon (c) of wild-type C57BL/6J mice. (*n* = 2 animals; 12 colon sections). (f) Analysis of bulk RNA-seq data from several tissues of C57BL/6 mice (55). RNA expression of *Tph1* and *Tph2* from the brain, colon and small intestine. RNA expression independently analyzed in three mice per tissue is indicated 1-3.

Supplementary Figure 5. Representative *in situ* validations confirming the co-expression of marker genes in enteric neurons.

(a) Chat⁺Nos1⁺ neurons. Representative images of *Chat* and *Nos1* expression in neurons. (b-f) Grey-scale *in situ* validation showing co-expression of DAPI (blue) along with either: (b) *Piezo1* (green), *Chat* (red), and TUBB3 (white); inset: Piezo1⁺Chat⁺TUBB3⁺ PEMN; (c) *Htr4* (green), *Chat* (red), and TUBB3 (white); inset: Htr4⁺Chat⁺TUBB3⁺ PEMN; (d) *Calca* (green), *Nos1* (red), and TUBB3 (white); inset: Calca⁺Nos1⁺TUBB3⁺ PSN; (e) *Cck* (green), *Piezo2* (red), and TUBB3 (white); yellow inset: Cck⁺Piezo2⁺TUBB3⁺ PSN in muscularis propria; red inset: Cck⁺Piezo2⁺TUBB3⁺ PSN in lamina propria; or (f) *Calcb* (green), *Chat* (red), and *Sst* (white); inset: Calcb⁺Chat⁺Sst⁺ PSN.

Supplementary Figure 6. Expression profiles reveal key functions of mouse enteric neuron subsets. (a,b) Fraction of expressing cells (dot size) and the mean levels in expressing (non-zero) cells (dot color) of select genes (columns) in neuron subsets (rows). **(a)** Select marker genes, including major neurotransmitters and neuropeptides (left) and other genes (right). **(b)** Data-driven unique marker genes that are differentially expressed in each neuron subset relative to other neurons.

Supplementary Figure 7. Reproducible cell subset distributions across ten human donors and expression of key genes in human enteric neuron subsets. (a-f) Shared and donor-specific cell subsets in the human cell census. t-SNE of 134,835 RAISIN RNA-seq profiles (a,d), 831 neurons (b,e), or 6,878 glia from cancer-proximal colon resections collected from ten human donors, colored by cell subset (a-c) or patient identifier (d-f). (g-j) Removal of oxidative phosphorylation (OXPHOS) signal in human neurons improved clustering by cell subset rather than cell state. t-SNE of human enteric neurons after removal of PC1 (g, identical to c) and before removal of PC1 (h-j) colored by cell subset, PC1 score (i), or OXPHOS expression score (J).

Supplementary Figure 8. Expression of major neurotransmitters and neuropeptides by human neuron subsets. Fraction of expressing cells (dot size) and the mean expression levels in expressing (non-zero) cells (dot color) of select marker genes (columns), including major neurotransmitters and neuropeptides (a) and other genes (b), across subsets of human enteric neurons (rows). Due to low levels of CHAT expression, we used the acetylcholine transporter, SLC5A7, as a marker of cholinergic neurons.

MATERIALS AND METHODS

Human donors and tissue samples

All colon resection samples were obtained from colon cancer patients after informed consent at either the Dana Farber Cancer Institute, Boston (IRB 03-189; ORSP 3490) or Massachusetts General Hospital, Boston (IRB 02-240; ORSP 1702). Normal colon located proximal to tumor was placed into conical tubes containing Roswell Park Memorial Institute (RPMI) media supplemented with 2% human serum and placed on ice for transport to the Broad Institute, Boston. Upon arrival, the muscularis propria was dissected from the remainder of the tissue (*e.g.*, submucosa), divided into pieces (approximately 20-120 mg), which were placed into cryo-vials, frozen on dry-ice and stored at -80°C. When possible, a portion of the tissue was fixed overnight in 4% paraformaldehyde at 4°C for histology.

Mouse models

All animal work was performed under the guidelines of the Division of Comparative Medicine, in accordance with the Institutional Animal Care and Use Committees (IACUC) relevant guidelines at the Broad Institute and MIT, and consistent with the Guide for Care and Use of Laboratory Animals, National Research Council, 1996 (institutional animal welfare assurance no. A4711-01), with protocol 0122-10-16. Mice were housed under specific-pathogen-free (SPF) conditions at the Broad Institute vivarium. The following strains were used:

Strain	Jackson Laboratory (Bar Harbor, ME) catalog number	Reference	Acknowledgement
C57BL/6J	000664		

B6;CBA-Tg(Sox10-cre)1Wdr/J	025807	(135)	We thank William Richardson for donating animals to a publicly available repository.
129S4.Cg-E2f1Tg(Wnt1-cre)2Sor/J	022137	(136)	We thank Philippe Soriano for donating animals to a publicly available repository.
B6;129-Gt(ROSA)26Sortm5(CAG-Sun1/sfGFP)Nat/J	021039	(137)	We thank Jeremy Nathans for donating animals to a publicly available repository.
Tg(Uchl1-HIST2H2BE/mCherry/EGFP*)FSout/J	016981	(138)	We thank Michelle Southard-Smith for donating animals to a publicly available repository.
B6.Cg-Gt(ROSA)26Sortm9(CAG-tdTomato)Hze/J	007909	(139)	We thank Hongkui Zeng for donating animals to a publicly available repository.

METHOD DETAILS

Tissue collection for snRNA-seq

For snRNA-seq optimization, tissue was collected from 11-14 week animals. For the ENS atlas, tissue was collected from 11-14 week old and 50-52 week old mice at either 7-8am or 7-8pm. Each colon was isolated and rinsed in ice cold PBS. Next, the colon was opened longitudinally and separated into four equally-sized sections, which were frozen in a 1.5 mL tube on dry ice. For brain collection, the brain was removed, quartered and frozen in a 1.5 mL tube on dry ice. Frozen tissue was stored at -80°C until subsequent tissue processing.

Tissue collection and preparation for RNA fluorescence *in situ* hybridization and immunohistochemistry

For RNA fluorescence *in situ* hybridization (RNA FISH) and Immunohistochemistry (IHC), isolated colon was cut into four sections of equal size and processed as described (140). Briefly, tissue was fixed in 4% paraformaldehyde overnight at 4°C. Then, tissue was sequentially passed through PBS containing 7.5%, 15% and 30% (w/v) sucrose at 4°C. Tissue was then embedded in O.C.T. (23-730-571, Fisher Scientific, Hampton, NH) and stored at -80°C. Tissue was cut at 25 micron thick sections onto Superfrost Plus microscope slides (22-037-246, Fisher Scientific) using a Leica CM1950 Cryostat (Leica Biosystems Inc., Buffalo Grove, IL).

Immunofluorescence (IF)

Slides with tissue sections were washed three times in PBS for 10 minutes, blocked 1 hour in CAS-Block Histochemical Reagent (00-8120, Thermo Fisher Scientific), incubated with primary antibodies overnight at 4°C, washed three times in PBS for 10 minutes, and then incubated with secondary antibodies at for 1 hour at room temperature. Slides were then washed twice in PBS for 10 minutes and then for 10 minutes with a PBS containing DAPI (D9542, Sigma-Aldrich). Lastly, slides were mounted using Southern Biotech Fluoromount-G (010001, VWR) and sealed.

Antibodies used for IF: Rabbit anti-Tubb3 (1:1000, AB18207, Abcam), Chicken anti-mCherry (1:1000, AB356481, EMD Millipore), and Alexa Fluor 488-, 594-, and 647-conjugated secondary antibodies (Life Technologies) were used.

Single-molecule fluorescence *in situ* hybridization (smFISH)

RNAScope Multiplex Fluorescent Kit (Advanced Cell Diagnostics) was used per manufacturer's recommendations for fresh-frozen samples with the following alterations. All Wash Buffer times were increased to 5 minutes and, following final HRP-Block step, slides were washed for 10

minutes with PBS containing DAPI (Sigma-Aldrich) followed by mounting with Southern Biotech Fluoromount-G (VWR) and sealed. Probes used for smISH (Advanced Cell Diagnostics): Calca (417961), Calcb (425511), Cck (402271), Chat (408731-C2), Grp (317861-C2), Nmu (446831), Nog (467391), Nos1 (437651-C3), Piezo1 (500511), Piezo2 (400191-C3), Sst (404631-C3), Tph2 (318691-C2), ANO1 (349021-C2), CHAT (450671 and 450671-C2), GUCY1A3 (425831), IL7 (424251), IL12A (402061), KIT (606401-C3), and NOS1 (506551-C2) were used.

Combined smFISH and IF

smFISH was performed as described above, with the following changes. After the final HRP-Block step, tissue sections were incubated with primary antibodies overnight at 4°C, washed in TBST for 5 minutes, twice, and then incubated with secondary antibodies for 30 min at room temperature. Slides were then washed in TBST for 5 minutes, twice, followed by a 10 minutes wash with containing DAPI (Sigma-Aldrich) before mounting with Southern Biotech Fluoromount-G (VWR) and sealed.

Confocal microscopy and image analysis

Images were taken using a Nikon TI-E microscope with a Yokohama W1 spinning disk, 405/488/561/640 lasers, and a Plan Apo 60X/1.4 objective. Images were visualized and overlaid using FIJI (141–144). The Bio-Formats plugin (145) was used to import all images.

Nuclei Extractions

The following nuclei extractions were performed from either mouse colon or brain and subsequently processed for profiling:

Dounce homogenization: Nuclei were extracted using either dounce homogenization followed by sucrose gradient centrifugation as described (134), or using the Nuclei EZ Prep (NUC101-1KT, Sigma-Aldrich) as described (22), with the following modifications. The tissues were dounce homogenized with a 7 mL Dounce Tissue Grinder (VWR 22877-280) (20 times pestle A, 20 times pestle B) and buffer volumes were increased to 5 mL for homogenization.

Tissue grinding: Fresh-Frozen tissues were crushed into a fine powder with a mortar and pestle (89038-144 and 89038-160, VWR) over a bath of liquid nitrogen. The powder was briefly resuspended in 2 mL of liquid nitrogen for transfer to a 50 mL conical tube, where liquid nitrogen was allowed to evaporate. The tissue powder was resuspended in 5 mL of Nuclei EZ Prep reagent (NUC101-1KT, Sigma-Aldrich) or NST (NP-40, Salts and Tris; see **Supp. Table 1**) and transferred to a 7 mL Dounce Tissue Grinder. For the Nuclei EZ Prep kit, all subsequent steps were as described (146). For NST, the tissue was dounce homogenized with a 7 mL Dounce Tissue Grinder (VWR 22877-280) (20 times pestle A, 20 times pestle B), filtered through a 40 μ m strainer (Falcon), and flow-through was spun at 500g for 5 minutes at 4°C. The pellet was resuspended in 0.5-3 mL of ST (Salts: 146 mM NaCl, 1 mM CaCl₂, 21 mM MgCl₂; Tris; **Supp. Table 1**).

Chopping extraction: Fresh-frozen tissues were disaggregated in 1 mL of custom nuclear extraction buffer (see **Supp. Table 1** for all combinations used) with mild chopping by Tungsten Carbide Straight 11.5 cm Fine Scissors (14558-11, Fine Science Tools, Foster City, CA) for 10 minutes on ice. Large debris were removed with a 40 μ m strainer (Falcon). An additional 1 mL of buffer was used to wash the filter before proceeding to fluorescence-activated cell sorting (FACS).

For droplet-based RNA-Seq, nuclei were isolated as described above, but with the addition of 3 ml of ST (Salts and Tris; **Supp. Table 1**) to extracted nuclei. Nuclei were then pelleted at 500g for 5 mins at 4°C. Supernatant was discarded and the nuclei pellet was resuspended in 100-500 µL of ST buffer (Salts and Tris; **Supp. Table 1**) before filtering through a 40 µm strainer-capped round bottom tube (Falcon).

Fluorescence-activated cell sorting (FACS)

Prior to sorting, isolated nuclei and RAISINs were stained with Vybrant DyeCycle Ruby Stain (V-10309, Thermo Fisher Scientific). Sorting was performed on a MoFlo Astrios EQ Cell Sorter (Beckman Coulter) using 488nm (GFP, 513/26 filter) or 561nm (mCherry 614/20 filter), and 640nm (Vybrant DyeCycle Ruby, 671/30 filter) lasers. Single nuclei were sorted into the wells of a 96-well PCR plate containing 5 µl of TCL buffer (1031576, Qiagen) with 1% β-mercaptoethanol. The 96 well plate was sealed tightly with a Microseal F and centrifuged at 800g for 3 minutes before being frozen on dry ice. Frozen plates were stored at -80°C until whole-transcriptome amplification, library construction, sequencing, and processing.

Whole-transcriptome amplification, library construction, sequencing, and processing

Libraries from isolated single nuclei and RAISINs were generated using SMART-seq2 as described (147), with the following modifications. RNA from individual wells was first purified with Agencourt RNAClean XP beads (A63987, Beckman Coulter) prior to oligo-dT primed reverse transcription with Maxima reverse transcriptase (EP0753, Thermo Fisher Scientific) and locked TSO oligonucleotide, which was followed by 21 cycles of PCR amplification using KAPA HiFi HotStart ReadyMix (NC0295239, KAPA Biosystems). cDNA was purified twice using

Agencourt AMPure XP beads (A63881, Beckman Coulter) as described (148). The Nextera XT Library Prep kit (FC-131-1096, Illumina, San Diego, CA) with custom barcode adapters (sequences available upon request) was used for library preparation. Libraries from 384 wells (nuclei/RAISINs) with unique barcodes were combined and sequenced using a NextSeq 500 sequencer (FC-404-2005, Illumina).

Droplet-based RAISIN RNA-seq

Single RAISINs were processed through the GemCode Single Cell Platform using the GemCode Gel Bead kit (v2 chemistry), Chip and Library Kits (10X Genomics, Pleasanton, CA), following the manufacturer's protocol. RAISINs were resuspended in ST buffer (Salt and Tris; **Supp. Table 1**). An input of 7,000 RAISINs was added to each channel of a chip. The RAISINs were then partitioned into Gel Beads in Emulsion (GEMs) in the GemCode instrument, where lysis and barcoded reverse transcription of RNA occurred, followed by amplification, shearing and 5' adaptor and sample index attachment. Libraries were sequenced on an Illumina NextSeq 500.

Transmission electron microscopy (TEM)

Extracted nuclei and RAISINs were pelleted and fixed at 4°C overnight in 2.5% Glutaraldehyde and 2% Paraformaldehyde in 0.1 M sodium cacodylate buffer (pH 7.4). The pellet was then washed in 0.1M cacodylate buffer, and post-fixed with 1% Osmiumtetroxide (OsO₄) and 1.5% Potassiumferrocyanide (KFeCN₆) for 1 hour. Next, the pellet was washed in water 3 times and incubated in 1% aqueous uranyl acetate for 1 hour followed by 2 washes in water and subsequent dehydration in grades of alcohol (10 minutes each; 50%, 70%, 90%, 100%, and 100%). The pellet was then put in propyleneoxide for 1 hour and infiltrated overnight in a 1:1 mixture of

propyleneoxide and TAAB Epon (Marivac Canada Inc. St. Laurent, Canada). The following day the samples were embedded in TAAB Epon and polymerized at 60°C for 48 hours.

Ultrathin sections (about 60nm) were cut on a Reichert Ultracut-S microtome, picked up on to copper grids stained with lead citrate and examined in a JEOL 1200EX Transmission electron microscope and images were recorded with an AMT 2k CCD camera.

Processing FASTQ reads into gene expression matrices

For SMART-seq2, FASTQ files were demultiplexed and aligned to a reference transcriptome (see “Mouse and human reference transcriptomes”), and transcripts were quantified using RSEM, as previously described (149). For droplet-based scRNA-Seq, Cell Ranger v2.0 was used to demultiplex the FASTQ reads, align them to a reference transcriptome, and extract their “cell” and “UMI” barcodes. The output of each pipeline is a digital gene expression (DGE) matrix for each sample, which records the number of transcripts or UMIs for each gene that are associated with each cell barcode. DGE matrices were filtered to remove low quality cells, defined as cells with fewer than 500 detected genes. This cutoff was set to remove contaminating cells, while retaining neurons and glia, which typically have high numbers of detected genes. To account for differences in sequencing depth across cells, DGE counts were normalized by the total number of transcripts or UMIs per cell and converted to transcripts-per-10,000 (henceforth “TP10K”).

Mouse and human reference transcriptomes

For the optimization of nuclei extraction conditions, reads were aligned to the mm10 reference transcriptome. However, for the mouse and human ENS atlases, we augmented the reference

transcriptomes with introns, thus allowing pre-mRNAs to be mapped along with mature mRNAs. Both the mm10 and hg19 reference transcriptomes were modified according to the instructions provided by the 10X Genomics website (<https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/advanced/references>). Briefly, we converted the standard GTF files into pre-mRNA GTF files by changing all “transcript” feature tags to “exon” feature tags. Using these modified GTF files, we then constructed Cell Ranger compatible references using the Cell Ranger “mkref” command. These modified GTF files were used for both the Cell Ranger pipeline and for our SMART-seq2 data (*i.e.* mouse ENS atlas).

Cell clustering overview

To cluster single cells into distinct cell subsets, we followed the general procedure we have previously outlined in (150) with additional modifications. This workflow includes the following steps: the selection of variable genes, batch correction, dimensionality reduction by PCA, and clustering. In all cases, clustering was performed twice: first, to separate neurons and glia from other cells, and then, to sub-cluster the neurons and glia to obtain high-resolution clusters within each group.

Partitioning cells into neuron, glia, and “other” compartments

Cells were partitioned into neuron, glia, and non-ENS compartments based on their expression of known marker genes (see “Gene signatures”). Signature scores were calculated as the mean $\log_2(\text{TP}10\text{K}+1)$ across all genes in the signature. Each cluster was assigned to the compartment of its maximal score and all cluster assignments were inspected to ensure the accurate segregation of cells. Neurons and glia were then assembled into two separate DGE matrices for further analysis.

Variable gene selection

To identify variable genes within a sample, we first calculated the mean (μ) and the coefficient of variation (CV) of expression of each gene. Genes were then grouped into 20 equal-frequency bins (ventiles) according to their mean expression levels. LOESS regression was used to fit the relationship, $\log(\text{CV}) \sim \log(\mu)$, and the 1,500 genes with the highest residuals were evenly sampled across these expression bins. To extend this approach to multiple samples, we performed variable gene selection separately for each sample to prevent “batch” differences between samples from unduly impacting the variable gene set. A consensus list of 1,500 variable genes was then formed by selecting the genes with the greatest recovery rates across samples, with ties broken by random sampling. This consensus gene set was then pruned through the removal of all ribosomal, mitochondrial, immunoglobulin, and HLA genes, which were found to induce unwanted batch effects in some samples in downstream clustering steps.

Batch correction

We observed substantial variability between cells that had been obtained from different mice or different individuals, which likely reflects a combination of technical and biological differences. In some cases, these “batch effects” led to cells clustering first by mouse or individual, rather than by cell type or cell state. To control for these batch differences, we ran ComBat (151) with default parameters on the $\log_2(\text{TP10K}+1)$ expression matrix, allowing cells to be clustered by cell type or cell state. Importantly, these batch-corrected data were only used for the PCA and other steps relying on PCA (*e.g.* clustering, t-SNE visualization); all other analyses (*e.g.* differential expression analysis) were based on the original expression data. Note that we tested two additional methods for batch correction – one based on Canonical Correlation Analysis (152) and another on

a k -nearest neighbors (k -NN) approach (153) – but did not obtain any enhancement in performance (data not shown).

Dimensionality reduction, graph clustering, and t-SNE visualization

Cells were clustered at two stages of the analysis: first, to initially partition the cells into neuron, glia, and “other” compartments, and second, to sub-cluster neurons and glia into different subsets. In all cases, we ran low-rank PCA on the variable genes of the batch-corrected $\log_2(\text{TP10K}+1)$ expression matrix. We then applied Phenograph (154) to the k -NN graph defined using the first n PCs and k nearest neighbors, which were separately estimated for each dataset. First, to estimate n , we calculated the number of “significant” PCs using a permutation test. Because this test may underestimate the number of PCs, we conservatively increased this number (*i.e.* to 15 or 30; see table below) to ensure that most of the variability in the dataset was captured. Next, to estimate k , we considered a range of clustering solutions with varying values of k , and calculated the marker genes for each set of clusters. We selected k based on inspection of the data. When clustering data from multiple cell types, we tried to select k such that the major cell types (*e.g.* neurons, glia, and muscle) were split, without fragmenting them into several sub-clusters. When clustering neurons and glia, we tried to select a k yielding the highest granularity clusters that were still biologically distinct, determined by close examination of the marker gene lists. Finally, the Barnes-Hut t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm was run on the selected PCs with perplexity = 20 and for 10,000 iterations to produce two-dimensional embeddings of the data for visualization.

Dataset	Cell type	# Sig PCs	Used PCs	k -NN
Optimization	All cells	13	1 to 15	250 (separates neurons and glia)

Mouse atlas	All cells	16	1 to 30	250 (separates neurons and glia)
Mouse atlas	Neurons	15	1 to 30	25
Mouse atlas	Glia	7	1 to 15	250
Mouse droplet	All cells	19	1 to 30	100 (separates major cell types)
Human atlas	All cells	20	2 to 30*	100 (separates major cell types)
Human atlas	Neurons	9	1 to 15	25
Human atlas	Glia	8	1 to 15	100

* See “Clustering of human neurons”.

Clustering of human neurons

Initial clustering of the 831 human neurons revealed 15 subsets (**Supp. Fig. 7h**). However, in several cases, we noticed that a single neuron type had been split into two clusters based on the expression of oxidative phosphorylation genes, which were strongly enriched in PC1 (**Supp. Fig. 7i,j**). This could reflect differences in differentiating vs. mature neurons (155), cancer-proximal effects, or a rapid transcriptional response to tissue resection or handling. We therefore re-clustered the cells based on the other PCs (*i.e.* PCs 2 to 30), yielding 11 final subsets of human neurons (**Supp. Fig. 7c,g**).

Scoring nuclei extraction conditions

To identify optimal conditions for snRNA-seq of the ENS, we performed nuclei extractions while systematically varying the detergent (CHAPS, Digitonin, EZ, NP40, Tween), buffer (HEPES, Tricine, Tris), mechanical extraction conditions (Dounce, Grind, Chop), and additional modifiers (*e.g.* polyamines, RNase inhibitors) (**Supp. Table 1**). In total, 104 different extraction conditions

were examined. For each extraction, we profiled single nuclei transcriptomes by SMART-Seq2 and clustered the resulting RNA into neurons, glia, and “other” (*i.e.* non-ENS or low quality) clusters (see “Cell clustering overview”). To compare extractions, we calculated several quality metrics for each condition: **(1)** the proportion of recovered neurons, glia, and “other” cells, **(2)** the mean number of detected genes per cell, and **(3)** the mean ENS signature score (derived from markers of neurons and glia; see “Cell type signatures”). Conditions that yielded high-quality nuclei enriched in the ENS signature score were then identified.

Cell lineage dendrogram

As an auxiliary tool, cell subsets were organized on a dendrogram according to their transcriptional similarities (**Fig. 2b**, top). To construct this tree, we performed complete linkage clustering on the distance matrix corresponding to the mean transcriptional distances among all cell subsets, calculated using the variable genes from the $\log_2(\text{TP10K}+1)$ expression matrix. These calculations were performed using the “hclust” and “dist” functions in R with default parameters.

Enteric neuron annotation and classification

We employed the following markers and considerations in annotating enteric neurons *post hoc*.

Broad segmentation of the mouse ENS

Broadly, neurons segmented into two major divisions comprising either cholinergic or nitrergic subsets. This broad division was correlated with several other genes. For example, the glial cell line-derived neurotrophic factor (GDNF) family receptors $\alpha 1$ (Gfra1) and $\alpha 2$ (Gfra2) segregate Nos1 and Chat expressing neurons, respectively. Gfra1/2 are co-receptors for the GDNF receptor,

Ret, which is necessary for ENS formation (156, 157). Similar, Chat and Nos1 expressing subsets also differentially expressed the transcription factors (TFs), Casz1 and Etv1.

Annotating mouse excitatory motor neurons

We annotated 6 subsets of putative excitatory motor neurons (PEMNs) based on co-expression of Chat and Tac1 (158) and position within the dendrogram on one subtree (**Fig. 2b**). Subsets of PEMNs express the endogenous opioid, enkephalin (Penk), which is found in motor neurons (159), and/or the myenteric motor neuron marker, calretinin (Calb2) (160).

Annotating mouse inhibitory motor neurons

We annotated 7 subsets of putative inhibitory motor neurons (PIMNs), which have high Nos1 and Vip co-expression (161, 162), and occupy one subtree of the dendrogram (**Fig. 2b**). In total, 73% of Vip-positive neurons co-express Nos1, which is consistent with the previously reported estimate of 75% (163, 164).

Annotating mouse interneurons

Enteric interneurons (INs) relay sensory information and coordinate excitatory and inhibitory motor neuron activity, but their classification is unclear. Six potential subtypes have been previously reported: (1) descending INs that signal via Chat, 5HT and ATP, (2) descending Nos1⁺Vip⁺Grp⁺Chat⁻ INs, (3) descending Vip⁺Chat⁺Nos1⁺ INs with ATP signaling, (4) descending Chat⁺Sst⁺ INs, (5) descending Penk⁺ INs (responsive to Sst), and (6) ascending Chat⁺Penk⁺ INs with ATP signaling (165–168).

Some of these subsets (3, 5, 6) are at least partly matched as discrete clusters in our data, whereas others (1, 2, 4) are not clearly observed in our atlas. PIMN7 is a potential candidate for the descending Vip⁺Chat⁺Nos1⁺ INs with ATP signaling (**3** above), based on co-expression of Vip, Chat, Nos1, and various ATP transporters (e.g. SLC28a1, Slc28a2, Slc28a3, Slc29a1, Slc29a2, Slc29a3, Slc29a4; (169). PSN3 also express these genes, but their expression of Cck, Calca, and Calcb makes it unlikely they are interneurons. Three subsets of Chat⁺Penk⁺ putative INs (PIN1-3) may reflect either descending Penk⁺ INs (**5** above; responsive to Sst), or ascending Chat⁺Penk⁺ INs with ATP signaling (**6** above). Because all express combinations of Sst receptors, they may be descending INs. However, given the substantial number of additional receptors expressed by all of these PINs (for 5HT, VIP, GAL, GLP, prolactin, prostaglandin E2, EGF and BMP) or some of them (e.g., catecholamine synthetic enzymes), they may not be INs. Finally, there was little to no evidence for other IN subtypes: we did not detect any serotonergic (5HT) neurons (**1** above) in our sampling, consistent with previous observations (170); found no discernible cluster of Nos1⁺Vip⁺Grp⁺Chat⁻ cells; and the only Chat⁺Sst⁺ neurons we observed were the Calcb⁺ PSN4 subset, which we interpret as a sensory neuron, not INs.

Annotating mouse secretomotor and vasodilator neurons

We annotated two subsets of Glp2r⁺ putative secretomotor/vasodilator (PSVNs) in one subtree of the dendrogram (**Fig. 2b**), one Vip⁺ non-cholinergic subtype (PSVN1) and one Chat⁺ cholinergic subset (PSVN2). The PSVN2 subset expresses Gal, previously reported in neurons that innervate the epithelium and arterioles (171) and neuropeptide Y expressed in a secretomotor neurons (172). Also, some neurons in PSVN2 expresses glutamate decarboxylase 2 (Gad2), possibly forming cholinergic/GABAergic neurons.

Annotating human interneuron subtype 2

Human PIN2s express two specific markers of mouse sensory neurons, CALCB and GRP, suggesting they may be misannotated sensory neurons. Another possibility is that PIN2s correspond to multiple neuron subtypes, which cannot be resolved with the number of neurons we profiled. Consistent with this possibility, PENK and CALCB expression are mutually exclusive within this subset (3 of 34 co-positive cells; expected = 7.24; Fisher test, $p < 0.001$).

Differential expression analysis

Differential expression (DE) tests were performed using MAST (173), which fits a hurdle model to the expression of each gene, consisting of logistic regression for the zero process (*i.e.* whether the gene is expressed) and linear regression for the continuous process (*i.e.* the expression level). For the mouse atlas, this regression model included terms to capture the effects of the cell subset, age, sex, colon location, circadian phase, transgenic model, and cell complexity. For the human atlas, this regression model only included terms for cell subset and cell complexity.

For the mouse atlas, we used the regression formula, $Y_i \sim X + A + C + L + S + T + N$, where Y_i is the standardized $\log_2(\text{TP10K}+1)$ expression vector for gene i across cells, X is a variable reflecting cell subset membership (*e.g.* PSNs *vs.* non-PSNs), A is the age associated with each cell (adult *vs.* aged), C is the circadian phase for each cell (morning *vs.* evening), L is the location for each cell (segments 1-4), S is the sex for each cell (male *vs.* female), T is the transgenic model for each cell (Sox10 *vs.* Uchl1), and N is the scaled number of genes for each cell (*i.e.* cell complexity). For the human atlas, we used the regression formula, $Y_i \sim X + N$, with X and N defined as above.

Additionally, two heuristics were used to increase the speed of the tests: we required all tested genes to have a minimum fold change of 1.2 and to be expressed by at least 1% of the cells within the group of interest. In all cases, the discrete and continuous coefficients of the model were retrieved and p-values were calculated using the likelihood ratio test in MAST. Q-values were separately estimated for each cell subset comparison using the Benjamini-Hochberg correction. Unless otherwise indicated, all reported DE coefficients and q-values correspond to the discrete component of the model (*i.e.* the logistic regression).

Acquisition and scoring of gene signatures

We compiled the following lists of marker genes for enteric neurons and glia from the literature (174). These gene signatures were then combined to construct an overall “ENS” signature score (**Fig. 1c; Supp. Fig. 2**).

Neurons: Tubb3, Elavl4, Ret, Phox2b, Chrn4, Eml5, Smpd3, Tagln3, Snap25, Gpr22, Gdap111, Stmn3, Chrna3, Scg3, Syt4, Ncan, Crmp1, Adcyap1r1, Elavl3, Dlg2, Cacna2d.

Glia: Erbb3, Sox10, Fabp7, Plp1, Gas7, Nid1, Qk, Sparc, Mest, Nfia, Wwtr1, Gpm6b, Rasa3, Flrt1, Itpril1, Itga4, Lama4, Postn, Ptpz1, Pdpn, Col18a1, Nrcam.

To prevent highly expressed genes from dominating a gene signature score, we scaled each gene vector of the $\log_2(\text{TP10K}+1)$ expression matrix by its root mean squared expression across all cells (using the ‘scale’ function in R with center = FALSE). The signature score for each cell was then computed as the mean scaled expression across all genes in the signature.

Estimation of false discovery rate

Unless otherwise specified, false discovery rates were estimated with the Benjamini-Hochberg correction (175), using the “p.adjust” R function with the “fdr” method.

Matching human and mouse subsets

To map human neurons onto their mouse counterparts, we first trained a Random Forest classifier to distinguish each of 24 subsets of mouse neurons using the $\log_2(\text{TP10K}+1)$ expression matrix of the mouse variable genes that also had human orthologs (see “Variable gene selection”). The Random Forest model was built with the R “randomForest” package using default parameters with the following exception: to account for class imbalances, we down-sampled each neuron class to the minimum class size while constructing each tree (implemented using the “sampsize” argument). In total, the “out of bag” estimate of the error rate (which estimates test rather than training error) was 8.8%, indicating that we can accurately distinguish among major neuron classes. Next, to extend this model to humans, we predicted the class for each human neuron using expression data for the human orthologs of the variable genes. All class assignments were then manually examined to ensure accurate predictions for all cells. Note that we also tested an alternative approach using a variational autoencoder (VAE) (176), but did not observe a noticeable improvement in performance (data not shown).

Identifying a core transcriptional program for major neuron classes

To identify conserved transcriptional signatures for each of the five major neuron classes (*i.e.*, PEMN, PIMN, PIN, PSN, PSVN), we first mapped all mouse genes to their corresponding human orthologs (using only 1:1 orthologs), and combined both expression matrices according to these

genes. We next calculated DE orthologs within each major neuron class (see “Differential expression analysis”), then selected genes that were significantly DE in the combined dataset, the mouse dataset, and the human dataset (**Supp. Table 5**).

Using receptor-ligand pairs to infer cell-cell interactions

To identify cell-cell interactions, we mapped the FANTOM5 database of literature-supported receptor-ligand interactions (177) onto our lists of cell subset markers. Following a recent approach (CellPhoneDB (178)), we filtered this database to remove all integrins (defined using the HUGO “Integrin” gene group), which were involved in many non-specific cell-cell interactions. We further required cell subset markers to be expressed in at least 5% of all cells within the subset. For all networks, we quantified the interaction strength between two cell subsets as the number of unique receptors and ligands connecting them, resulting in an adjacency matrix summarizing all cell-cell interactions within the dataset. Statistical significance was then empirically assessed by permuting the receptors and ligands among all cell subsets, thus preserving the number of receptors and ligands encoded within each cell subset, and preserving the distribution of ligand-receptor connectivity (but possibly changing the connectivity between cell subsets, in those cases where one receptor has multiple ligands, or vice versa). After running 10,000 total permutations, p-values were computed as the number of times the edge strength in the permuted network was greater than or equal to the edge strength in the true network. To plot cell-cell interaction networks, we applied the Fruchterman-Reingold layout algorithm to a network defined using the $-\log_{10}(\text{p-value})$, using only the edges with $\text{p-value} < 0.05$. Although edge weights were used to generate the layout, they were removed from the final visualization for visual clarity (**Fig. 3i**).

Defining disease risk genes

We compiled lists of genes that have been implicated by human genetics studies as contributing to risk for the following diseases: Hirschsprung's disease (HRSC), inflammatory bowel disease (IBD), autism spectrum disorders (ASDs), and Parkinson's disease (PD). Because human genetics studies do not always pinpoint a causative risk gene, we used the literature to identify sets of genes that are particularly likely to contribute to disease risk, including: 9 HRSC-associated genes (179), 106 IBD-associated genes (180), 28 ASD-associated genes (181), and 29 PD-associated genes (182).

References

1. Yoo BB, Mazmanian SK (2017) The Enteric Network: Interactions between the Immune and Nervous Systems of the Gut. *Immunity* 46(6):910–926.
2. Furness JB, et al. (2008) The circuitry of the enteric nervous system. *Neurogastroenterology & Motility* 6(3):241–253.
3. Furness JB (2012) The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol Hepatol* 9(5):286–294.
4. Sasselli V, Pachnis V, Burns AJ (2012) The enteric nervous system. *Dev Biol* 366(1):64–73.
5. Furness JB (2012) The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol Hepatol* 9(5):286–294.
6. Bernard CE, et al. (2009) Effect of age on the enteric nervous system of the human colon. *Neurogastroenterol Motil* 21(7):746–e46.
7. Li M, et al. (2018) Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* 362(6420). doi:10.1126/science.aat7615.
8. Scheving LA (2000) Biological clocks and the digestive system. *Gastroenterology* 119(2):536–549.
9. De Giorgio R, et al. (2016) Enteric neuropathies: Yesterday, Today and Tomorrow. *Adv Exp Med Biol* 891:123–133.
10. Furness JB (2012) The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol Hepatol* 9(5):286–294.
11. Pesce M, Borrelli O, Saliakellis E, Thapar N (2018) Gastrointestinal Neuropathies: New Insights and Emerging Therapies. *Gastroenterol Clin North Am* 47(4):877–894.
12. Yoo BB, Mazmanian SK (2017) The Enteric Network: Interactions between the Immune and Nervous Systems of the Gut. *Immunity* 46(6):910–926.
13. Klose CSN, et al. (2017) The neuropeptide neuromedin U stimulates innate lymphoid cells and type 2 inflammation. *Nature* 549(7671):282–286.
14. Knights D, Lassen KG, Xavier RJ (2013) Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. *Gut* 62(10):1505–1510.
15. Chaidez V, Hansen RL, Hertz-Picciotto I (2014) Gastrointestinal problems in children with autism, developmental delays or typical development. *J Autism Dev Disord* 44(5):1117–1127.
16. Pfeiffer RF (2003) Gastrointestinal dysfunction in Parkinson’s disease. *Lancet Neurol* 2(2):107–116.

17. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
18. Zeisel A, et al. (2018) Molecular Architecture of the Mouse Nervous System. *Cell* 174(4):999–1014.e22.
19. Lasrado R, et al. (2017) Lineage-dependent spatial and functional organization of the mammalian enteric nervous system. *Science* 356(6339):722–726.
20. Rosenberg AB, et al. (2018) Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 360(6385):176–182.
21. Sathiyamurthy A, et al. (2018) Massively Parallel Single Nucleus Transcriptional Profiling Defines Spinal Cord Neurons and Their Activity during Behavior. *Cell Rep* 22(8):2216–2225.
22. Habib N, et al. (2017) Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* 14(10):955–958.
23. Habib N, et al. (2016) Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* 353(6302):925–928.
24. Lake BB, et al. (2016) Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352(6293):1586–1590.
25. Debbache J, Parfejevs V, Sommer L (2018) Cre-driver lines used for genetic fate mapping of neural crest cells in the mouse: An overview. *Genesis* 56(6-7):e23105.
26. Lewis AE, Vasudevan HN, O'Neill AK, Soriano P, Bush JO (2013) The widely used Wnt1-Cre transgene causes developmental phenotypes by ectopic activation of Wnt signaling. *Dev Biol* 379(2):229–234.
27. Matsuoka T, et al. (2005) Neural crest origins of the neck and shoulder. *Nature* 436(7049):347–355.
28. Hari L, et al. (2012) Temporal control of neural crest lineage generation by Wnt/ β -catenin signaling. *Development* 139(12):2107–2117.
29. Laranjeira C, et al. (2011) Glial cells in the mouse enteric nervous system can undergo neurogenesis in response to injury. *J Clin Invest* 121(9):3412–3424.
30. Mo A, et al. (2015) Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. *Neuron* 86(6):1369–1384.
31. Wiese CB, Fleming N, Buehler DP, Southard-Smith EM (2013) A Uchl1-Histone2BmCherry:GFP-gpi BAC transgene for imaging neuronal progenitors. *Genesis* 51(12):852–861.

32. Madisen L, et al. (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* 13(1):133–140.
33. Zeisel A, et al. (2018) Molecular Architecture of the Mouse Nervous System. *Cell* 174(4):999–1014.e22.
34. Modlin IM, et al. (2008) Gastroenteropancreatic neuroendocrine tumours. *Lancet Oncol* 9(1):61–72.
35. Habib N, et al. (2017) Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* 14(10):955–958.
36. Habib N, et al. (2016) Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* 353(6302):925–928.
37. Simon C, Lickert H, Götz M, Dimou L (2012) Sox10-iCreERT2 : a mouse line to inducibly trace the neural crest and oligodendrocyte lineage. *Genesis* 50(6):506–515.
38. Habib N, et al. (2017) Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* 14(10):955–958.
39. Habib N, et al. (2016) Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* 353(6302):925–928.
40. Habib N, et al. (2017) Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* 14(10):955–958.
41. Habib N, et al. (2016) Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* 353(6302):925–928.
42. Prunuske AJ, Ullman KS (2006) The nuclear envelope: form and reformation. *Curr Opin Cell Biol* 18(1):108–116.
43. Prunuske AJ, Ullman KS (2006) The nuclear envelope: form and reformation. *Curr Opin Cell Biol* 18(1):108–116.
44. Lasrado R, et al. (2017) Lineage-dependent spatial and functional organization of the mammalian enteric nervous system. *Science* 356(6339):722–726.
45. Haber AL, et al. (2017) A single-cell survey of the small intestinal epithelium. *Nature* 551(7680):333–339.
46. Lasrado R, et al. (2017) Lineage-dependent spatial and functional organization of the mammalian enteric nervous system. *Science* 356(6339):722–726.
47. Furness JB (2012) The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol Hepatol* 9(5):286–294.
48. Memic F, et al. (2018) Transcription and Signaling Regulators in Developing Neuronal

- Subtypes of Mouse and Human Enteric Nervous System. *Gastroenterology* 154(3):624–636.
49. Wallrapp A, et al. (2017) The neuropeptide NMU amplifies ILC2-driven allergic lung inflammation. *Nature* 549(7672):351–356.
50. Rao M, Gershon MD (2018) Enteric nervous system development: what could possibly go wrong? *Nat Rev Neurosci* 19(9):552–565.
51. Gershon MD (2009) Enteric serotonergic neurones ... finally! *J Physiol* 587(3):507.
52. Mawe GM, Hoffman JM (2013) Serotonin signalling in the gut--functions, dysfunctions and therapeutic targets. *Nat Rev Gastroenterol Hepatol* 10(8):473–486.
53. Costa M, et al. (1982) Neurons with 5-hydroxytryptamine-like immunoreactivity in the enteric nervous system: their visualization and reactions to drug treatment. *Neuroscience* 7(2):351–363.
54. Costa M, et al. (1996) Neurochemical classification of myenteric neurons in the guinea-pig ileum. *Neuroscience* 75(3):949–967.
55. Söllner JF, et al. (2017) An RNA-Seq atlas of gene expression in mouse and rat normal tissues. *Sci Data* 4:170185.
56. Zeisel A, et al. (2018) Molecular Architecture of the Mouse Nervous System. *Cell* 174(4):999–1014.e22.
57. Kim H, et al. (2018) Generation of a highly efficient and tissue-specific tryptophan hydroxylase 1 knockout mouse model. *Sci Rep* 8(1):17642.
58. Betancur P, Bronner-Fraser M, Sauka-Spengler T (2010) Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. *Proc Natl Acad Sci U S A* 107(8):3570–3575.
59. Matsuoka T, et al. (2005) Neural crest origins of the neck and shoulder. *Nature* 436(7049):347–355.
60. Herrero A, Duhart JM, Ceriani MF (2017) Neuronal and Glial Clocks Underlying Structural Remodeling of Pacemaker Neurons in. *Front Physiol* 8:918.
61. Assas BM, Pennock JI, Miyan JA (2014) Calcitonin gene-related peptide is a key neurotransmitter in the neuro-immune axis. *Front Neurosci* 8:23.
62. Junttila IS (2018) Tuning the Cytokine Responses: An Update on Interleukin (IL)-4 and IL-13 Receptor Complexes. *Front Immunol* 9:888.
63. Yan J, Wang H, Liu Y, Shao C (2008) Analysis of gene regulatory networks in the mammalian circadian rhythm. *PLoS Comput Biol* 4(10):e1000193.

64. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
65. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
66. Gershon MD, Tack J (2007) The serotonin signaling system: from basic understanding to drug development for functional GI disorders. *Gastroenterology* 132(1):397–414.
67. Mawe GM, Hoffman JM (2013) Serotonin signalling in the gut--functions, dysfunctions and therapeutic targets. *Nat Rev Gastroenterol Hepatol* 10(8):473–486.
68. Bulbring E, Lin RC, Schofield G (1958) An investigation of the peristaltic reflex in relation to anatomical observations. *Q J Exp Physiol Cogn Med Sci* 43(1):26–37.
69. Bulbring E, Lin RC (1958) The effect of intraluminal application of 5-hydroxytryptamine and 5-hydroxytryptophan on peristalsis; the local production of 5-HT and its release in relation to intraluminal pressure and propulsive activity. *J Physiol* 140(3):381–407.
70. Ginzel KH (1959) Are mucosal nerve fibres essential for the peristaltic reflex? *Nature* 184(Suppl 16):1235–1236.
71. Diamant ML, Kosterlitz HW, McKENZIE J (1961) Role of the mucous membrane in the peristaltic reflex in the isolated ileum of the guinea pig. *Nature* 190:1205–1206.
72. Tsuji S, Anglade P, Ozaki T, Sazi T, Yokoyama S (1992) Peristaltic movement evoked in intestinal tube devoid of mucosa and submucosa. *Jpn J Physiol* 42(3):363–375.
73. Keating DJ, Spencer NJ (2010) Release of 5-hydroxytryptamine from the mucosa is not required for the generation or propagation of colonic migrating motor complexes. *Gastroenterology* 138(2):659–70 670.e1–2.
74. Spencer NJ, et al. (2011) Mechanisms underlying distension-evoked peristalsis in guinea pig distal colon: is there a role for enterochromaffin cells? *Am J Physiol Gastrointest Liver Physiol* 301(3):G519–27.
75. Bertrand PP (2006) Real-time measurement of serotonin release and motility in guinea pig ileum. *J Physiol* 577(Pt 2):689–704.
76. Li Z, et al. (2011) Essential roles of enteric neuronal serotonin in gastrointestinal motility and the development/survival of enteric dopaminergic neurons. *J Neurosci* 31(24):8998–9009.
77. Kim H, et al. (2018) Generation of a highly efficient and tissue-specific tryptophan hydroxylase 1 knockout mouse model. *Sci Rep* 8(1):17642.
78. Alcaïno C, Farrugia G, Beyder A (2017) Mechanosensitive Piezo Channels in the Gastrointestinal Tract. *Curr Top Membr* 79:219–244.

79. Mazzuoli-Weber G, et al. (2019) Piezo proteins: incidence and abundance in the enteric nervous system. Is there a link with mechanosensitivity? *Cell Tissue Res* 375(3):605–618.
80. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
81. Mawe GM, Hoffman JM (2013) Serotonin signalling in the gut--functions, dysfunctions and therapeutic targets. *Nat Rev Gastroenterol Hepatol* 10(8):473–486.
82. Liu M-T, Kuan Y-H, Wang J, Hen R, Gershon MD (2009) 5-HT₄ receptor-mediated neuroprotection and neurogenesis in the enteric nervous system of adult mice. *J Neurosci* 29(31):9683–9699.
83. Pan H, Galligan JJ (1994) 5-HT_{1A} and 5-HT₄ receptors mediate inhibition and facilitation of fast synaptic transmission in enteric neurons. *Am J Physiol* 266(2 Pt 1):G230–8.
84. Cottrell GS, et al. (2012) Localization of calcitonin receptor-like receptor (CLR) and receptor activity-modifying protein 1 (RAMP1) in human gastrointestinal tract. *Peptides* 35(2):202–211.
85. Haber AL, et al. (2017) A single-cell survey of the small intestinal epithelium. *Nature* 551(7680):333–339.
86. Furness JB, Jones C, Nurgali K, Clerc N (2004) Intrinsic primary afferent neurons and nerve circuits within the intestine. *Prog Neurobiol* 72(2):143–164.
87. Yu Y-B, et al. (2012) Brain-derived neurotrophic factor contributes to abdominal pain in irritable bowel syndrome. *Gut* 61(5):685–694.
88. Szurszewski JH, Ermilov LG, Miller SM (2002) Prevertebral ganglia and intestinofugal afferent neurones. *Gut* 51 Suppl 1:i6–10.
89. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
90. Corleto VD (2010) Somatostatin and the gastrointestinal tract. *Curr Opin Endocrinol Diabetes Obes* 17(1):63–68.
91. Haramis A-PG, et al. (2004) De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. *Science* 303(5664):1684–1686.
92. Klose CSN, et al. (2017) The neuropeptide neuromedin U stimulates innate lymphoid cells and type 2 inflammation. *Nature* 549(7671):282–286.
93. Klose CSN, et al. (2017) The neuropeptide neuromedin U stimulates innate lymphoid cells and type 2 inflammation. *Nature* 549(7671):282–286.
94. Satoh-Takayama N, et al. (2010) IL-7 and IL-15 independently program the differentiation

- of intestinal CD3-NKp46+ cell subsets from Id2-dependent precursors. *J Exp Med* 207(2):273–280.
95. Degan S, Lopez GY, Kevill K, Sunday ME (2008) Gastrin-releasing peptide, immune responses, and lung disease. *Ann N Y Acad Sci* 1144:136–147.
96. Furness JB (2000) Types of neurons in the enteric nervous system. *J Auton Nerv Syst* 81(1-3):87–96.
97. Drucker DJ, Yusta B (2014) Physiology and pharmacology of the enteroendocrine hormone glucagon-like peptide-2. *Annu Rev Physiol* 76:561–583.
98. Afroze S, et al. (2013) The physiological roles of secretin and its receptor. *Ann Transl Med* 1(3):29.
99. Furness JB (2000) Types of neurons in the enteric nervous system. *J Auton Nerv Syst* 81(1-3):87–96.
100. Artis D, Spits H (2015) The biology of innate lymphoid cells. *Nature* 517(7534):293–301.
101. Cianferoni A, Spergel J (2014) The importance of TSLP in allergic disease and its role as a potential therapeutic target. *Expert Rev Clin Immunol* 10(11):1463–1474.
102. Wallrapp A, et al. (2017) The neuropeptide NMU amplifies ILC2-driven allergic lung inflammation. *Nature* 549(7672):351–356.
103. Black SAG, Rylett RJ (2012) Choline transporter CHT regulation and function in cholinergic neurons. *Cent Nerv Syst Agents Med Chem* 12(2):114–121.
104. Legay C (2000) Why so many forms of acetylcholinesterase? *Microsc Res Tech* 49(1):56–72.
105. Al-Shboul OA (2013) The importance of interstitial cells of cajal in the gastrointestinal tract. *Saudi J Gastroenterol* 19(1):3–15.
106. Gomez-Pinilla PJ, et al. (2009) Ano1 is a selective marker of interstitial cells of Cajal in the human and mouse gastrointestinal tract. *Am J Physiol Gastrointest Liver Physiol* 296(6):G1370–81.
107. Al-Shboul OA (2013) The importance of interstitial cells of cajal in the gastrointestinal tract. *Saudi J Gastroenterol* 19(1):3–15.
108. König P, Groneberg D, Jäger R, Friebe A (2011) NO-sensitive guanylyl cyclase is expressed in pericytes but absent from endothelial cells in the murine lung. *BMC Pharmacol* 11(S1). doi:10.1186/1471-2210-11-s1-p38.
109. Groneberg D, König P, Koesling D, Friebe A (2011) Nitric oxide-sensitive guanylyl

- cyclase is dispensable for nitrergic signaling and gut motility in mouse intestinal smooth muscle. *Gastroenterology* 140(5):1608–1617.
110. Smillie CS, et al. (2018) Rewiring of the cellular and inter-cellular landscape of the human colon during ulcerative colitis. *bioRxiv*:455451.
 111. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
 112. Furness JB (2012) The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol Hepatol* 9(5):286–294.
 113. Lima WG, Marques-Oliveira GH, da Silva TM, Chaves VE (2017) Role of calcitonin gene-related peptide in energy metabolism. *Endocrine* 58(1):3–13.
 114. Loh K, Herzog H, Shi Y-C (2015) Regulation of energy homeostasis by the NPY system. *Trends Endocrinol Metab* 26(3):125–135.
 115. Yoo BB, Mazmanian SK (2017) The Enteric Network: Interactions between the Immune and Nervous Systems of the Gut. *Immunity* 46(6):910–926.
 116. Athie-Morales V, Smits HH, Cantrell DA, Hilkens CMU (2004) Sustained IL-12 signaling is required for Th1 development. *J Immunol* 172(1):61–69.
 117. Klose CSN, et al. (2017) The neuropeptide neuromedin U stimulates innate lymphoid cells and type 2 inflammation. *Nature* 549(7671):282–286.
 118. Chang D, et al. (2017) A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat Genet* 49(10):1511–1516.
 119. Sanders SJ, et al. (2015) Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* 87(6):1215–1233.
 120. Bondurand N, Southard-Smith EM (2016) Mouse models of Hirschsprung disease and other developmental disorders of the enteric nervous system: Old and new players. *Dev Biol* 417(2):139–157.
 121. Klose CSN, et al. (2017) The neuropeptide neuromedin U stimulates innate lymphoid cells and type 2 inflammation. *Nature* 549(7671):282–286.
 122. Pla P, Larue L (2003) Involvement of endothelin receptors in normal and pathological development of neural crest cells. *Int J Dev Biol* 47(5):315–325.
 123. Ischia J, Patel O, Shulkes A, Baldwin GS (2009) Gastrin-releasing peptide: different forms, different functions. *Biofactors* 35(1):69–75.
 124. Tomas-Roca L, et al. (2015) De novo mutations in PLXND1 and REV3L cause Möbius syndrome. *Nat Commun* 6:7199.

125. Gross KJ, Pothoulakis C (2007) Role of neuropeptides in inflammatory bowel disease. *Inflamm Bowel Dis* 13(7):918–932.
126. Zeisel A, et al. (2018) Molecular Architecture of the Mouse Nervous System. *Cell* 174(4):999–1014.e22.
127. Lasrado R, et al. (2017) Lineage-dependent spatial and functional organization of the mammalian enteric nervous system. *Science* 356(6339):722–726.
128. Cagampang FR, Bruce KD (2012) The role of the circadian clock system in nutrition and metabolism. *Br J Nutr* 108(3):381–392.
129. Haramis A-PG, et al. (2004) De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. *Science* 303(5664):1684–1686.
130. Vatanen T, et al. (2016) Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell* 165(6):1551.
131. Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA (2018) The Influence of the Gut Microbiome on Cancer, Immunity, and Cancer Immunotherapy. *Cancer Cell* 33(4):570–580.
132. Deczkowska A, Schwartz M (2018) Targeting neuro-immune communication in neurodegeneration: Challenges and opportunities. *J Exp Med* 215(11):2702–2704.
133. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
134. Habib N, et al. (2016) Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* 353(6302):925–928.
135. Matsuoka T, et al. (2005) Neural crest origins of the neck and shoulder. *Nature* 436(7049):347–355.
136. Lewis AE, Vasudevan HN, O'Neill AK, Soriano P, Bush JO (2013) The widely used Wnt1-Cre transgene causes developmental phenotypes by ectopic activation of Wnt signaling. *Dev Biol* 379(2):229–234.
137. Mo A, et al. (2015) Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. *Neuron* 86(6):1369–1384.
138. Wiese CB, Fleming N, Buehler DP, Southard-Smith EM (2013) A Uchl1-Histone2BmCherry:GFP-gpi BAC transgene for imaging neuronal progenitors. *Genesis* 51(12):852–861.
139. Madisen L, et al. (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* 13(1):133–140.

140. Drokhyansky E, et al. (2017) The brain parenchyma has a type I interferon response that can limit virus spread. *Proc Natl Acad Sci U S A* 114(1):E95–E104.
141. Pietzsch T, Preibisch S, Tomancák P, Saalfeld S (2012) ImgLib2--generic image processing in Java. *Bioinformatics* 28(22):3009–3011.
142. Schindelin J, et al. (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9(7):676–682.
143. Schindelin J, Rueden CT, Hiner MC, Eliceiri KW (2015) The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev* 82(7-8):518–529.
144. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.
145. Linkert M, et al. (2010) Metadata matters: access to image data in the real world. *J Cell Biol* 189(5):777–782.
146. Habib N, et al. (2017) Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* 14(10):955–958.
147. Shekhar K, et al. (2016) Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* 166(5):1308–1323.e30.
148. Shekhar K, et al. (2016) Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* 166(5):1308–1323.e30.
149. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.
150. Haber AL, et al. (2017) A single-cell survey of the small intestinal epithelium. *Nature* 551(7680):333–339.
151. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1):118–127.
152. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 36(5):411.
153. Shekhar K, et al. (2016) Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* 166(5):1308–1323.e30.
154. Levine JH, et al. (2015) Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* 162(1):184–197.
155. Shekhar K, et al. (2016) Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* 166(5):1308–1323.e30.

156. Schuchardt A, D'Agati V, Pachnis V, Costantini F (1996) Renal agenesis and hypodysplasia in ret-k- mutant mice result from defects in ureteric bud development. *Development* 122(6):1919–1929.
157. Rossi J, et al. (1999) Retarded growth and deficits in the enteric and parasympathetic nervous system in mice lacking GFR alpha2, a functional neurturin receptor. *Neuron* 22(2):243–252.
158. Brookes SJ, Steele PA, Costa M (1991) Identification and immunohistochemistry of cholinergic and non-cholinergic circular muscle motor neurons in the guinea-pig small intestine. *Neuroscience* 42(3):863–878.
159. Brookes SJ, Steele PA, Costa M (1991) Identification and immunohistochemistry of cholinergic and non-cholinergic circular muscle motor neurons in the guinea-pig small intestine. *Neuroscience* 42(3):863–878.
160. Brookes SJ, Steele PA, Costa M (1991) Calretinin immunoreactivity in cholinergic motor neurones, interneurons and vasomotor neurones in the guinea-pig small intestine. *Cell Tissue Res* 263(3):471–481.
161. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
162. Sang Q, Young HM (1996) Chemical coding of neurons in the myenteric plexus and external muscle of the small and large intestine of the mouse. *Cell Tissue Res* 284(1):39–53.
163. Sang Q, Young HM (1996) Chemical coding of neurons in the myenteric plexus and external muscle of the small and large intestine of the mouse. *Cell Tissue Res* 284(1):39–53.
164. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
165. Grider JR (2003) Neurotransmitters mediating the intestinal peristaltic reflex in the mouse. *J Pharmacol Exp Ther* 307(2):460–467.
166. Young HM, Furness JB, Povey JM (1995) Analysis of connections between nitric oxide synthase neurons in the myenteric plexus of the guinea-pig small intestine. *J Neurocytol* 24(4):257–263.
167. Furness JB (2012) The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol Hepatol* 9(5):286–294.
168. Grider JR (1994) Interplay of somatostatin, opioid, and GABA neurons in the regulation of the peristaltic reflex. *Am J Physiol* 267(4 Pt 1):G696–701.
169. Brookes SJ, Steele PA, Costa M (1991) Identification and immunohistochemistry of cholinergic and non-cholinergic circular muscle motor neurons in the guinea-pig small intestine. *Neuroscience* 42(3):863–878.

170. Sang Q, Young HM (1996) Chemical coding of neurons in the myenteric plexus and external muscle of the small and large intestine of the mouse. *Cell Tissue Res* 284(1):39–53.
171. Furness JB, Costa M, Rökaeus A, McDonald TJ, Brooks B (1987) Galanin-immunoreactive neurons in the guinea-pig small intestine: their projections and relationships to other enteric neurons. *Cell Tissue Res* 250(3):607–615.
172. Furness JB (2012) The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol Hepatol* 9(5):286–294.
173. Finak G, et al. (2015) MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* 16:278.
174. Lasrado R, et al. (2017) Lineage-dependent spatial and functional organization of the mammalian enteric nervous system. *Science* 356(6339):722–726.
175. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Series B Stat Methodol* 57(1):289–300.
176. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N (2018) Deep generative modeling for single-cell transcriptomics. *Nat Methods* 15(12):1053–1058.
177. Ramilowski JA, et al. (2015) A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat Commun* 6:7866.
178. Vento-Tormo R, et al. (2018) Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* 563(7731):347–353.
179. Bondurand N, Southard-Smith EM (2016) Mouse models of Hirschsprung disease and other developmental disorders of the enteric nervous system: Old and new players. *Dev Biol* 417(2):139–157.
180. Smillie CS, et al. (2018) Rewiring of the cellular and inter-cellular landscape of the human colon during ulcerative colitis. doi:10.1101/455451.
181. Sanders SJ, et al. (2015) Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* 87(6):1215–1233.
182. Chang D, et al. (2017) A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat Genet* 49(10):1511–1516.

Figure 1

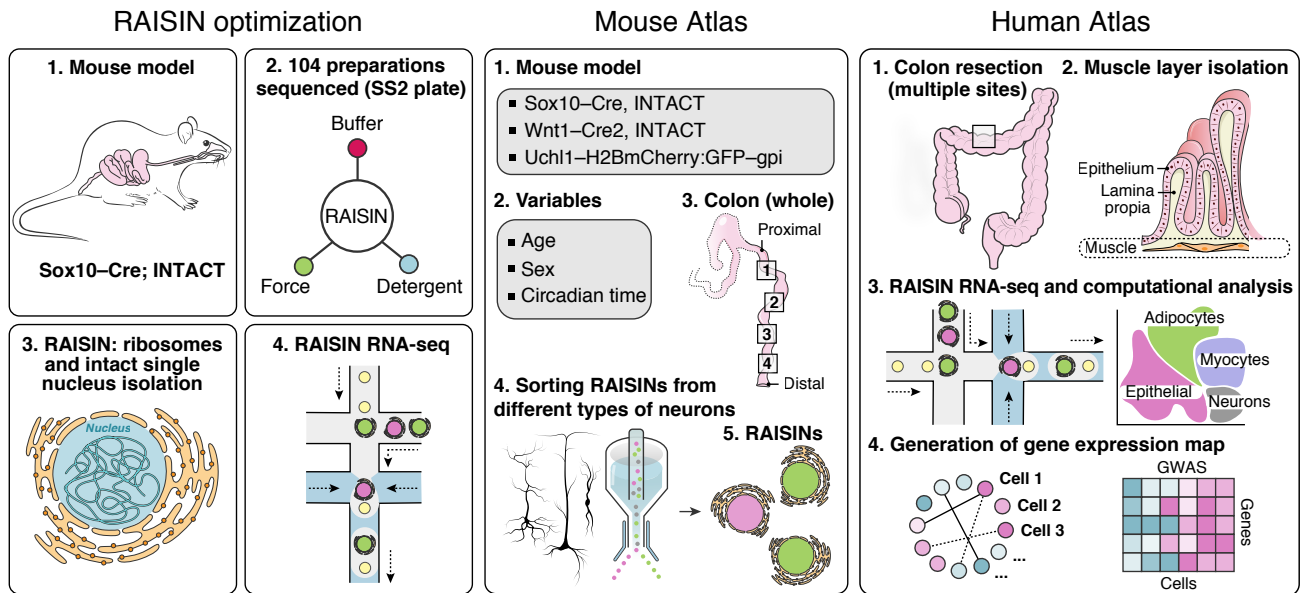
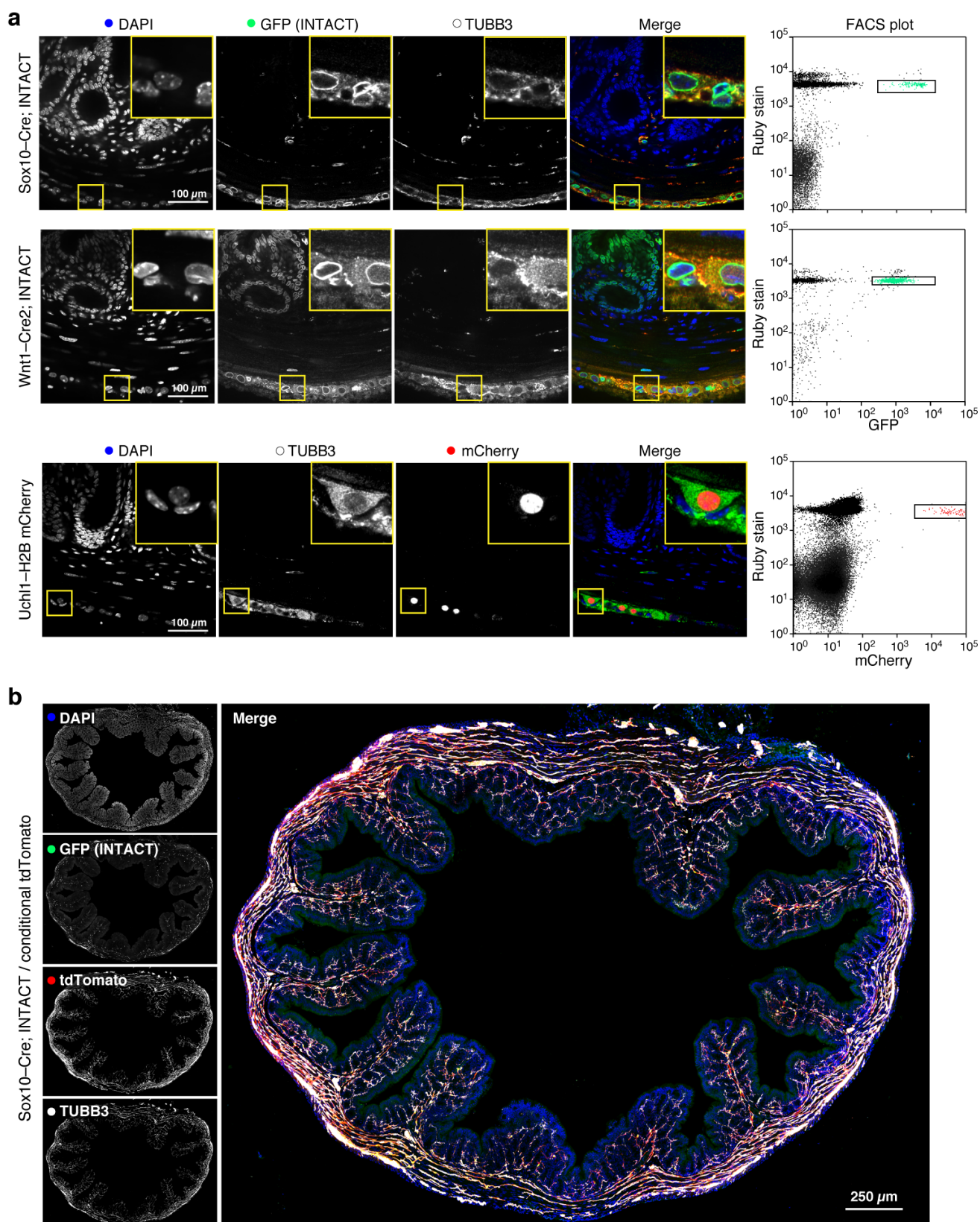
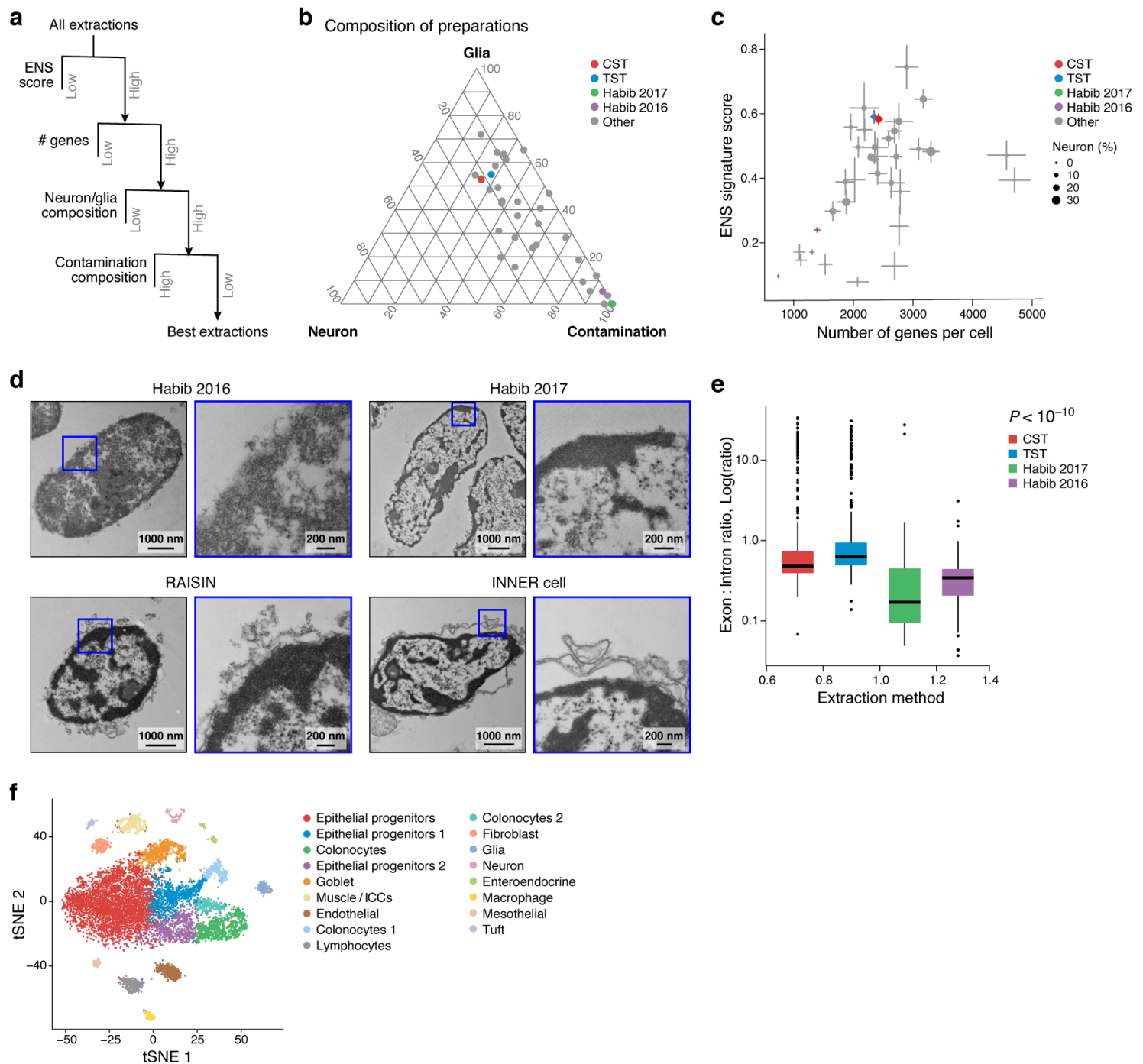


Figure 2



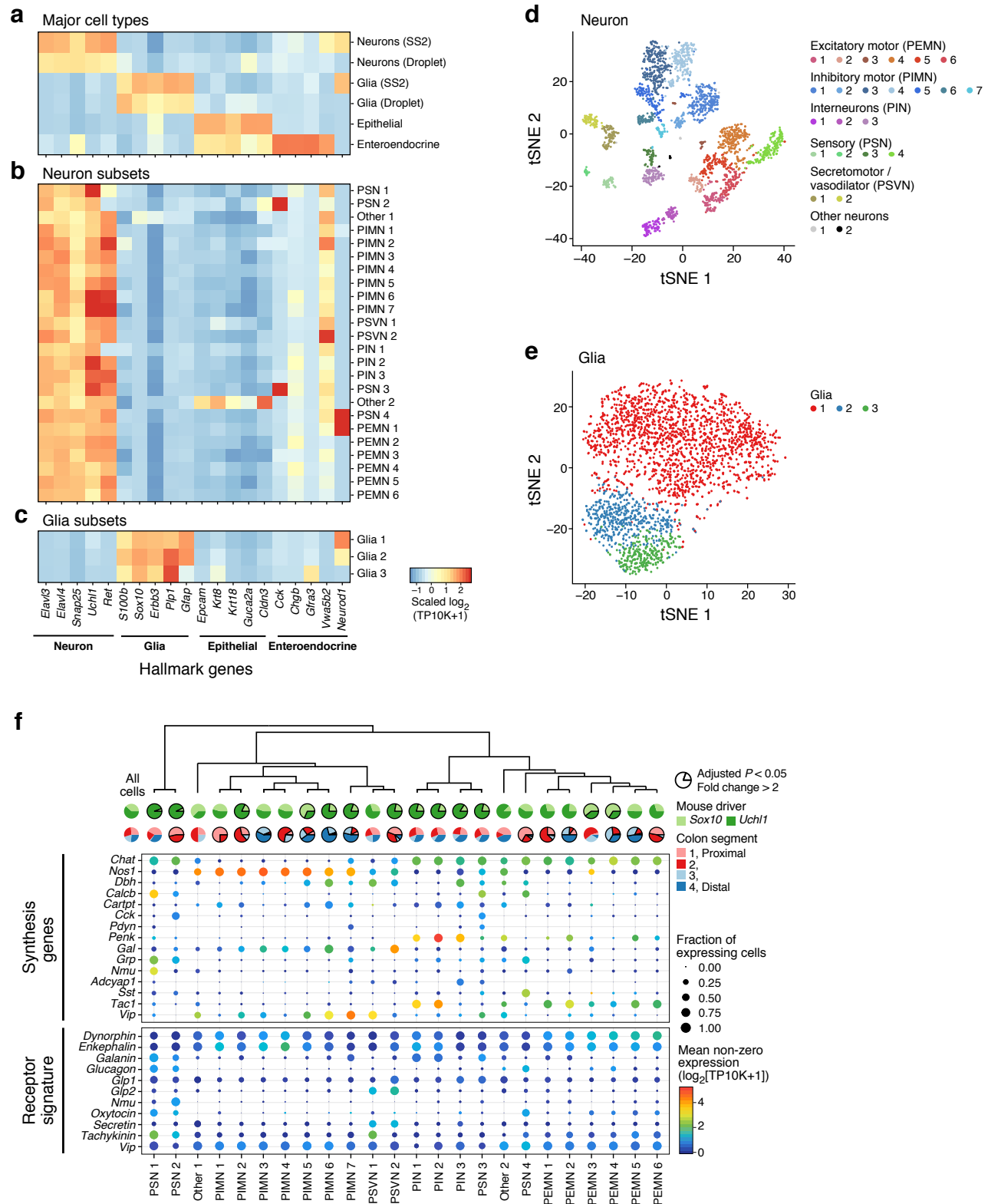
Resolution limited due to upload size. Higher resolution images available upon request

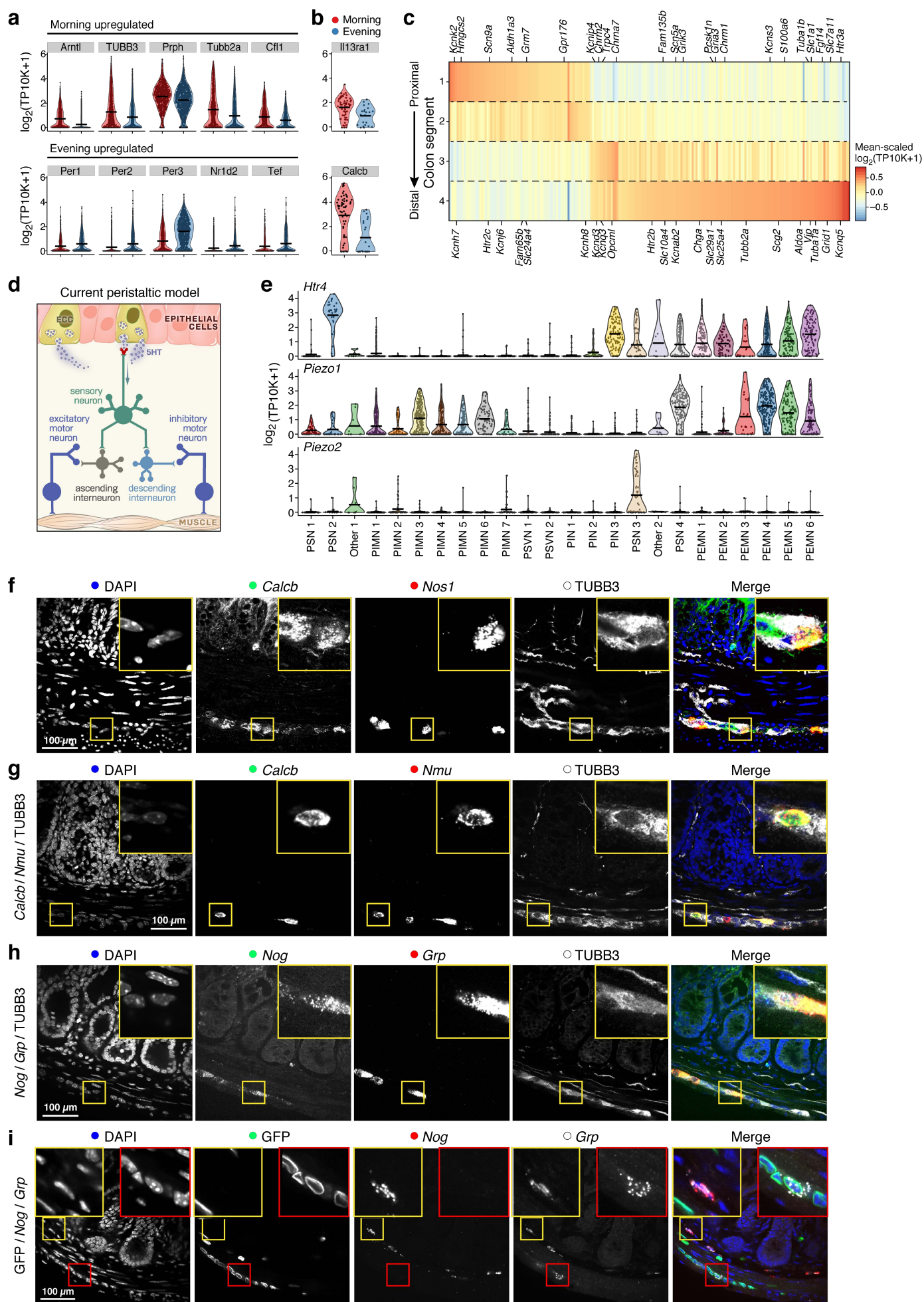
Figure 3



Resolution limited due to upload size. Higher resolution images available upon request

Figure 4





Resolution limited due to upload size. Higher resolution images available upon request

Figure 6

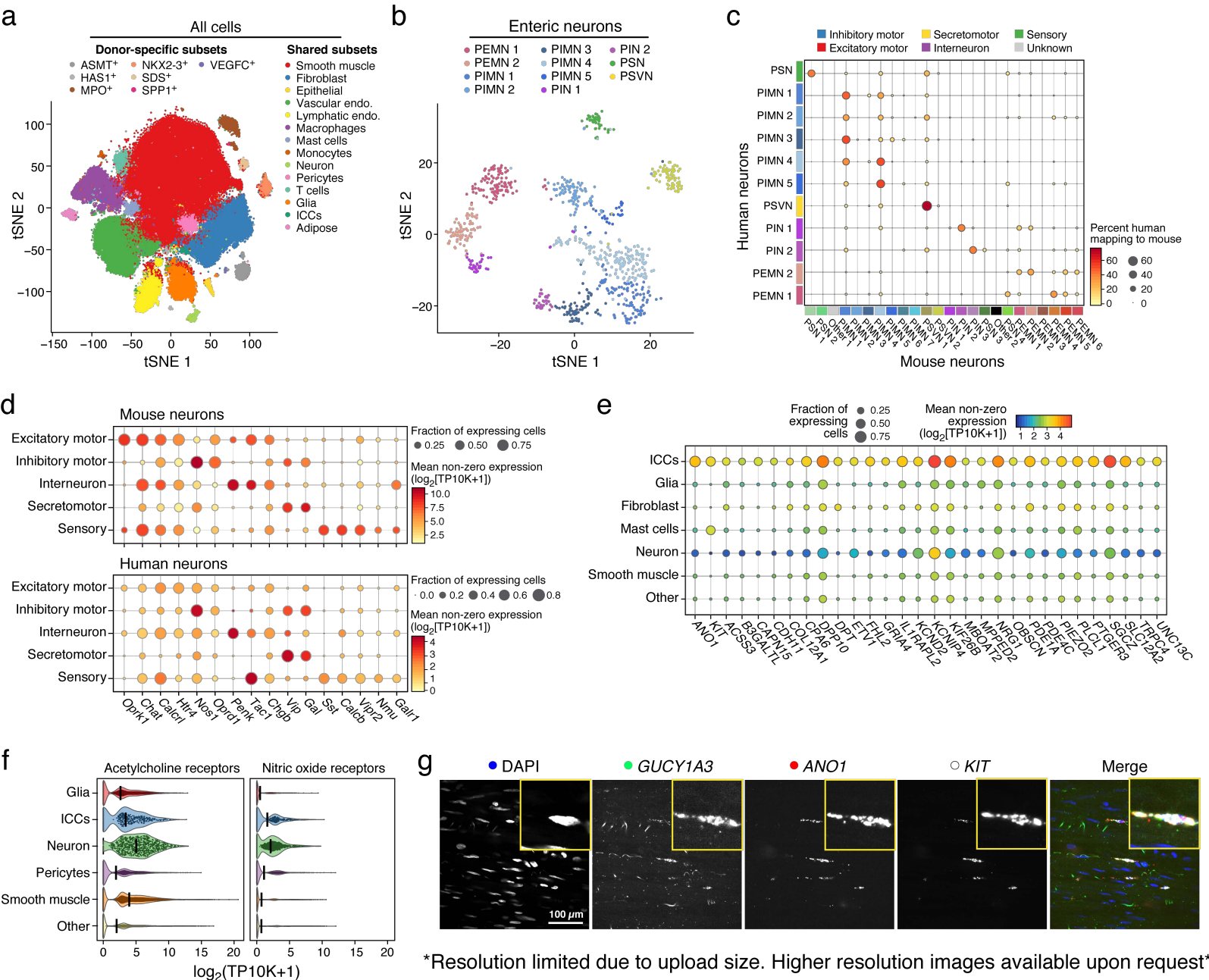
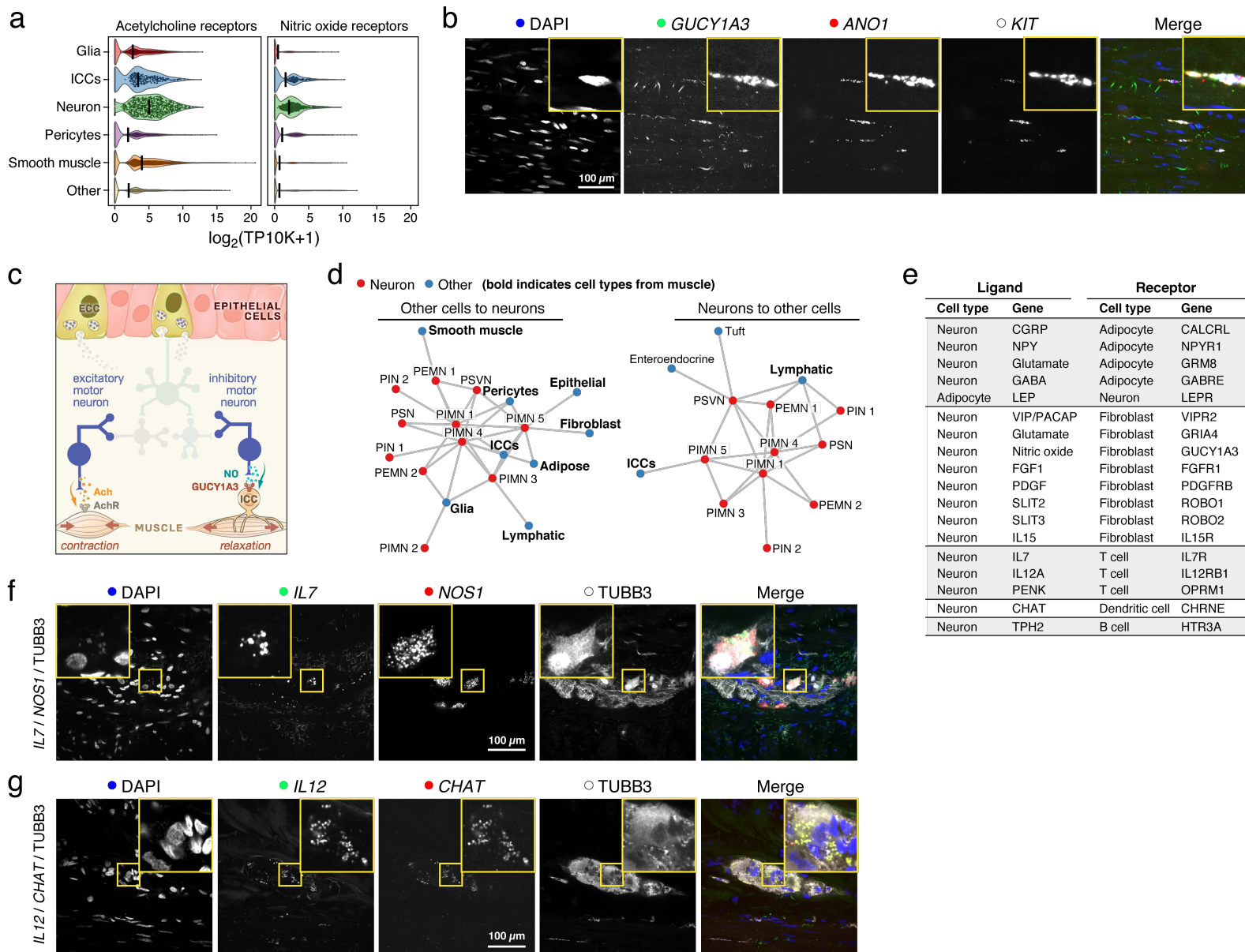
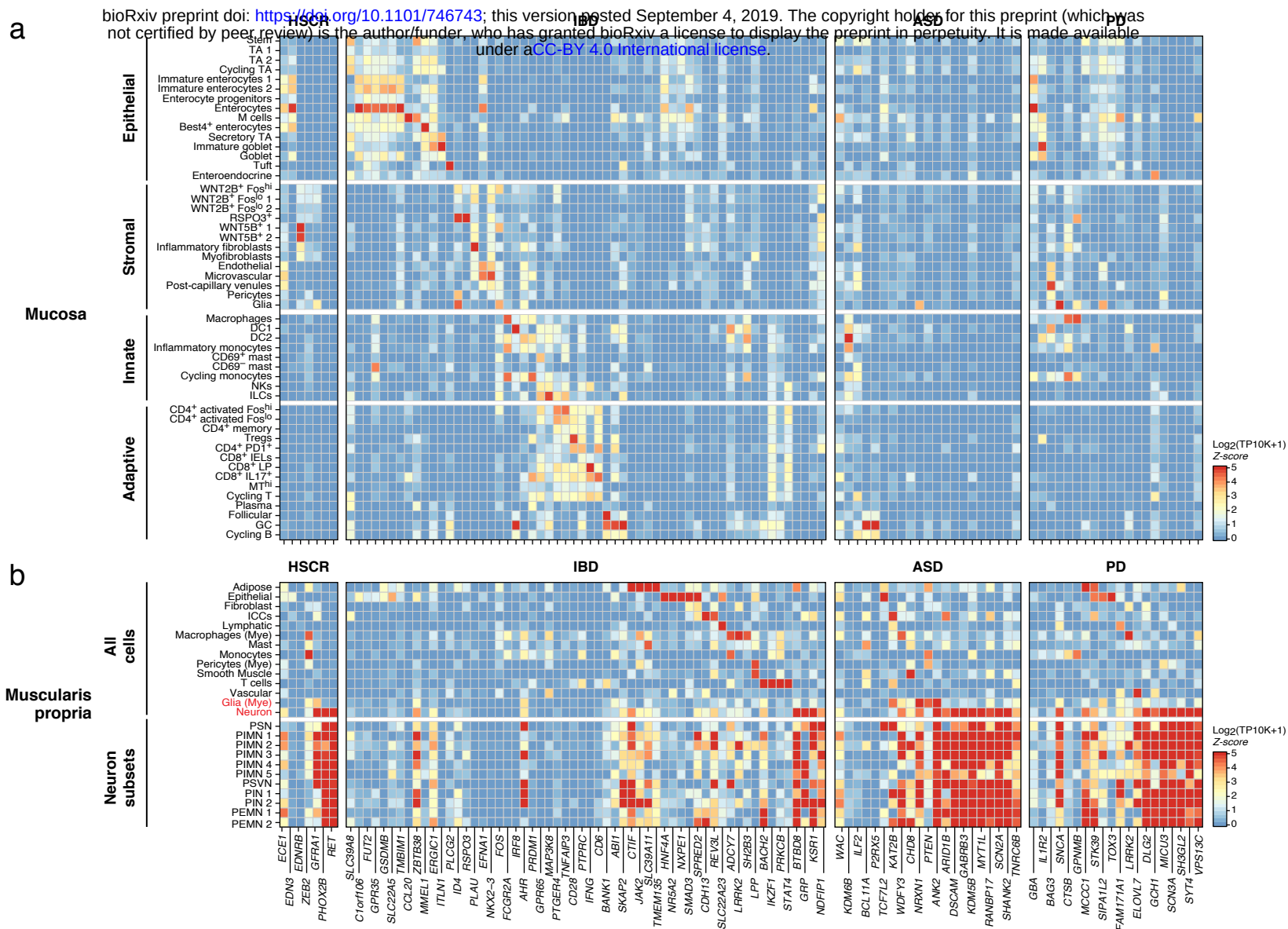


Figure 7

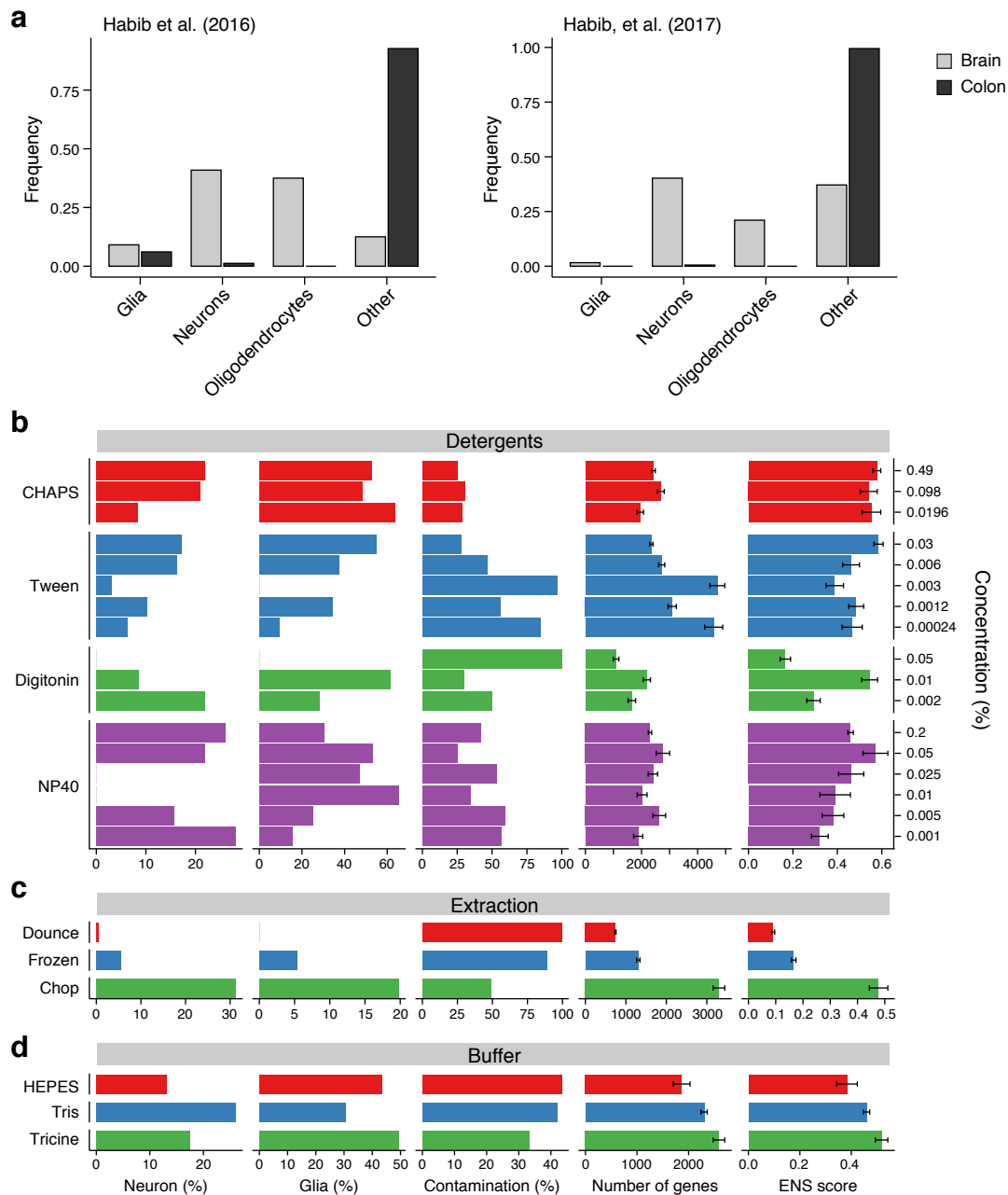


Resolution limited due to upload size. Higher resolution images available upon request

Figure 8



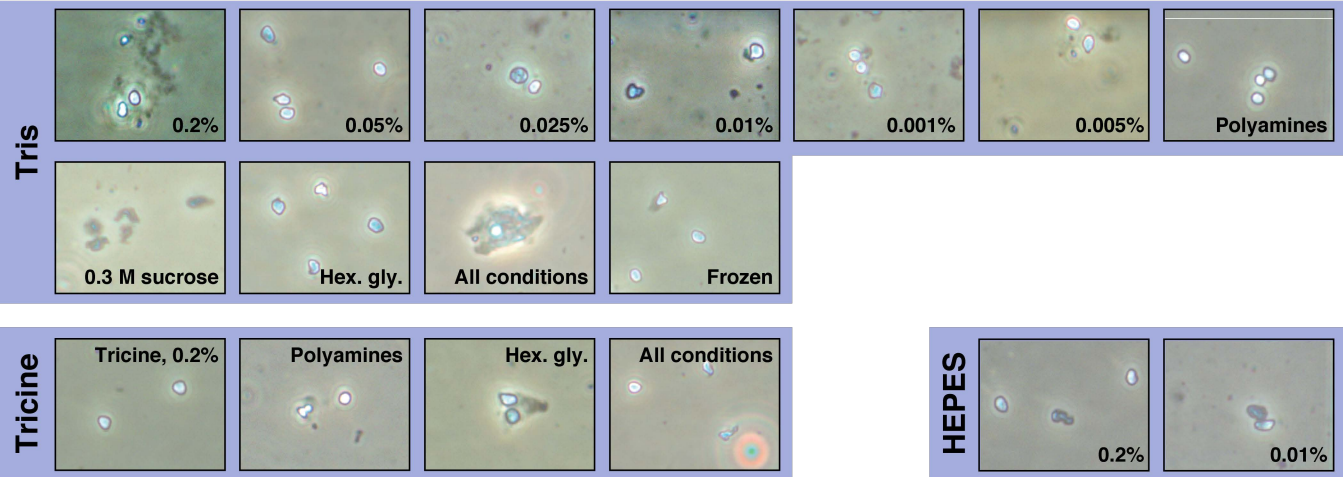
Supplementary Figure 1



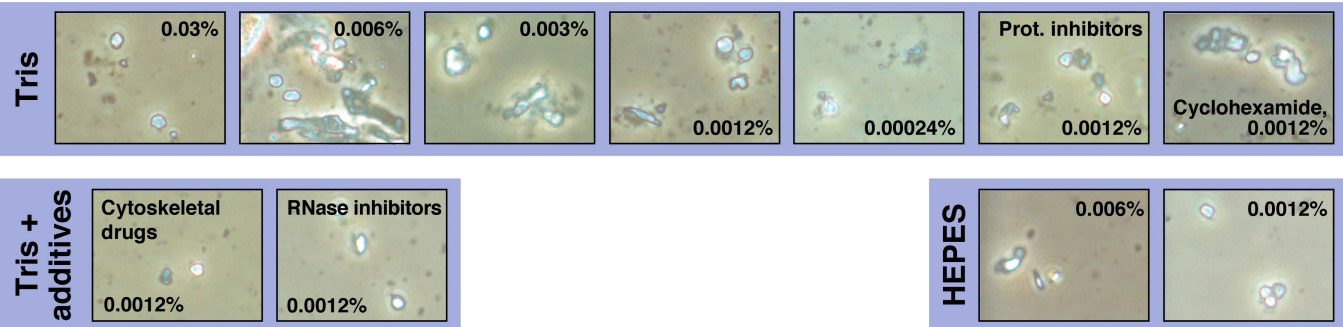
Supplementary Figure 2

■ Detergent ■ Buffer — 25 μ m

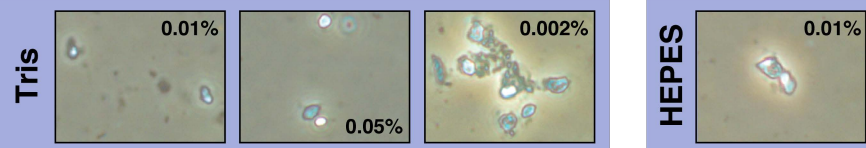
NP40



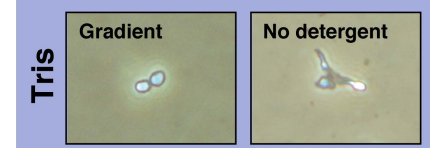
Tween 20



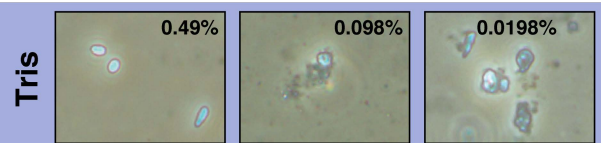
Digitonin



Other



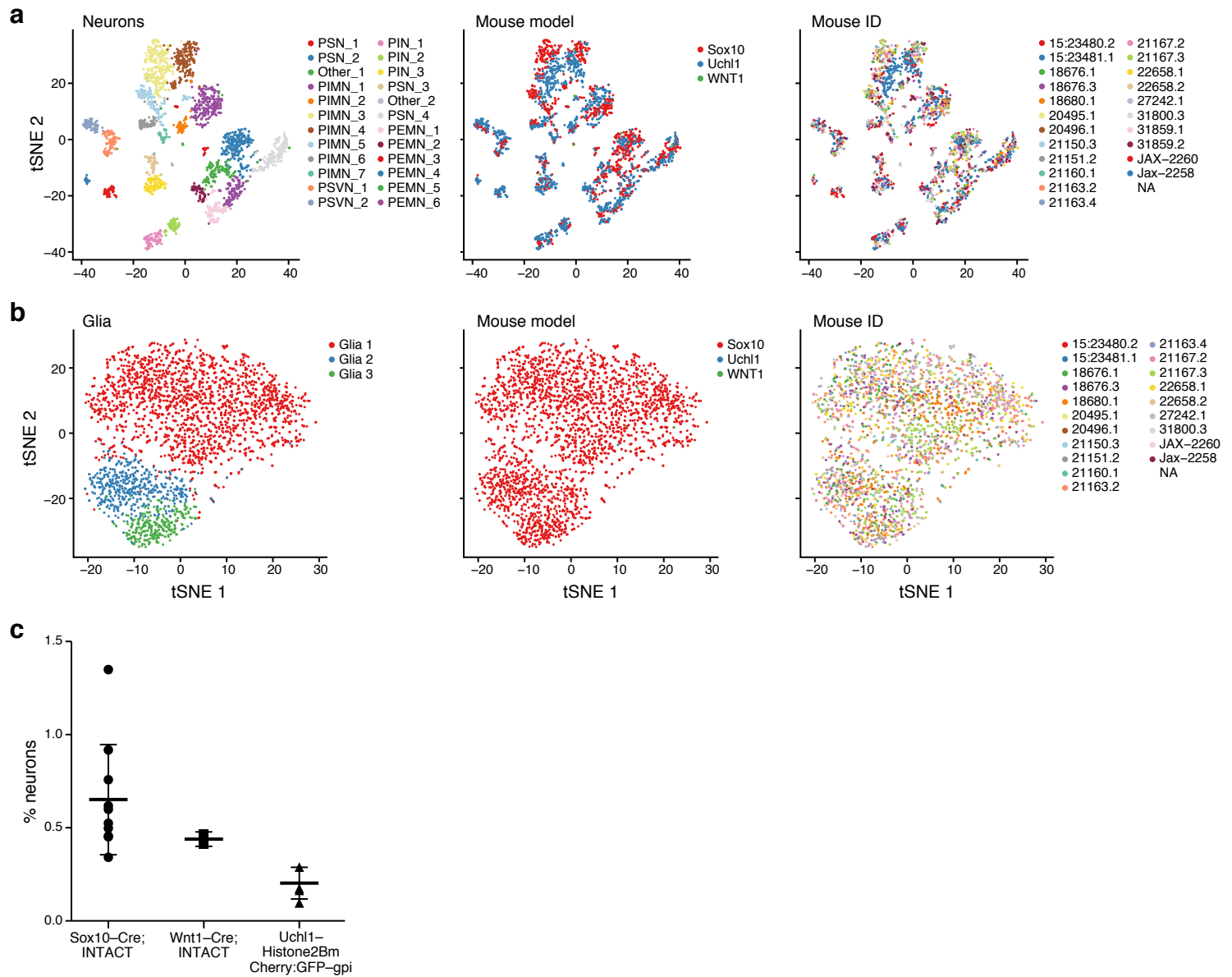
CHAPS

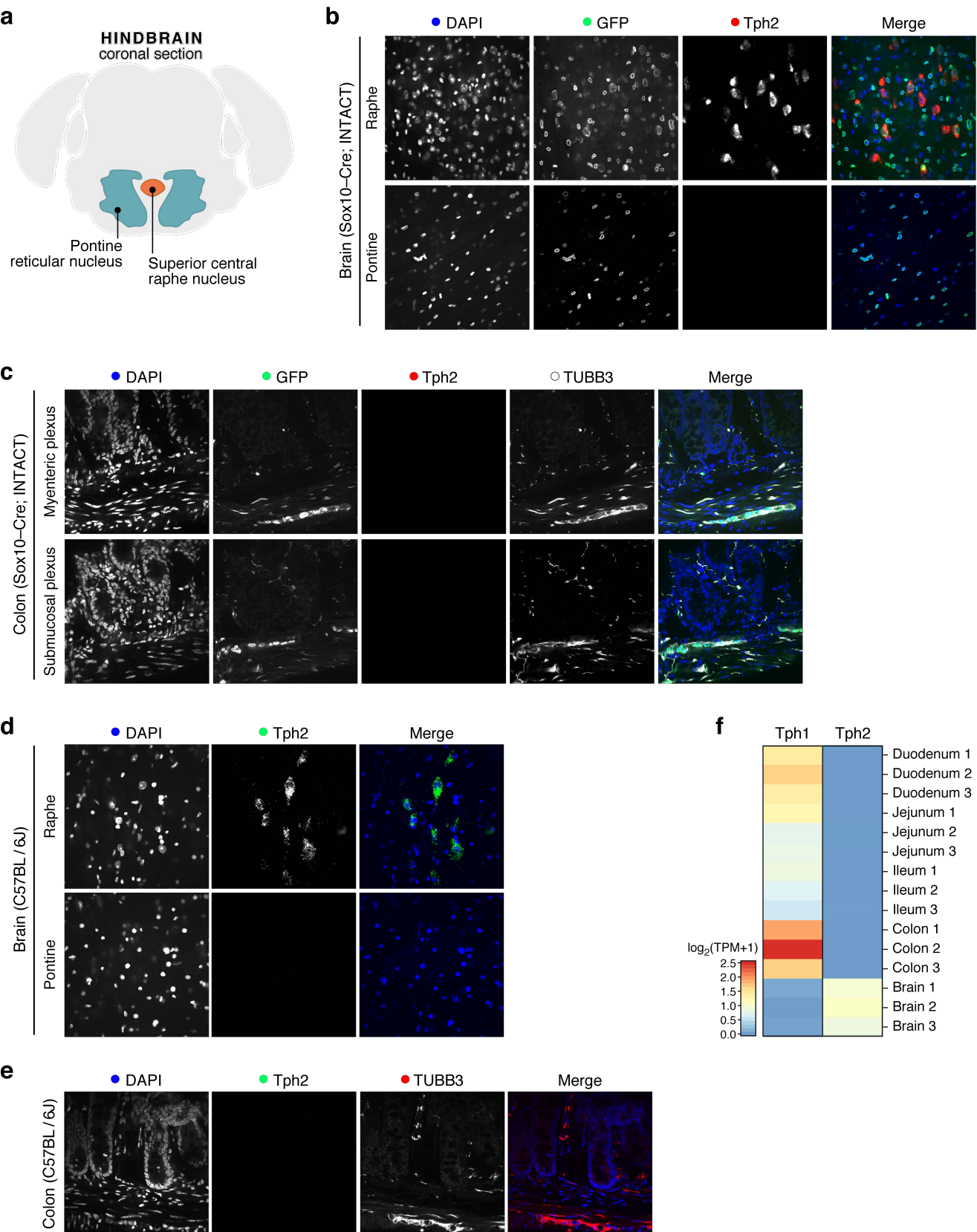


EZ



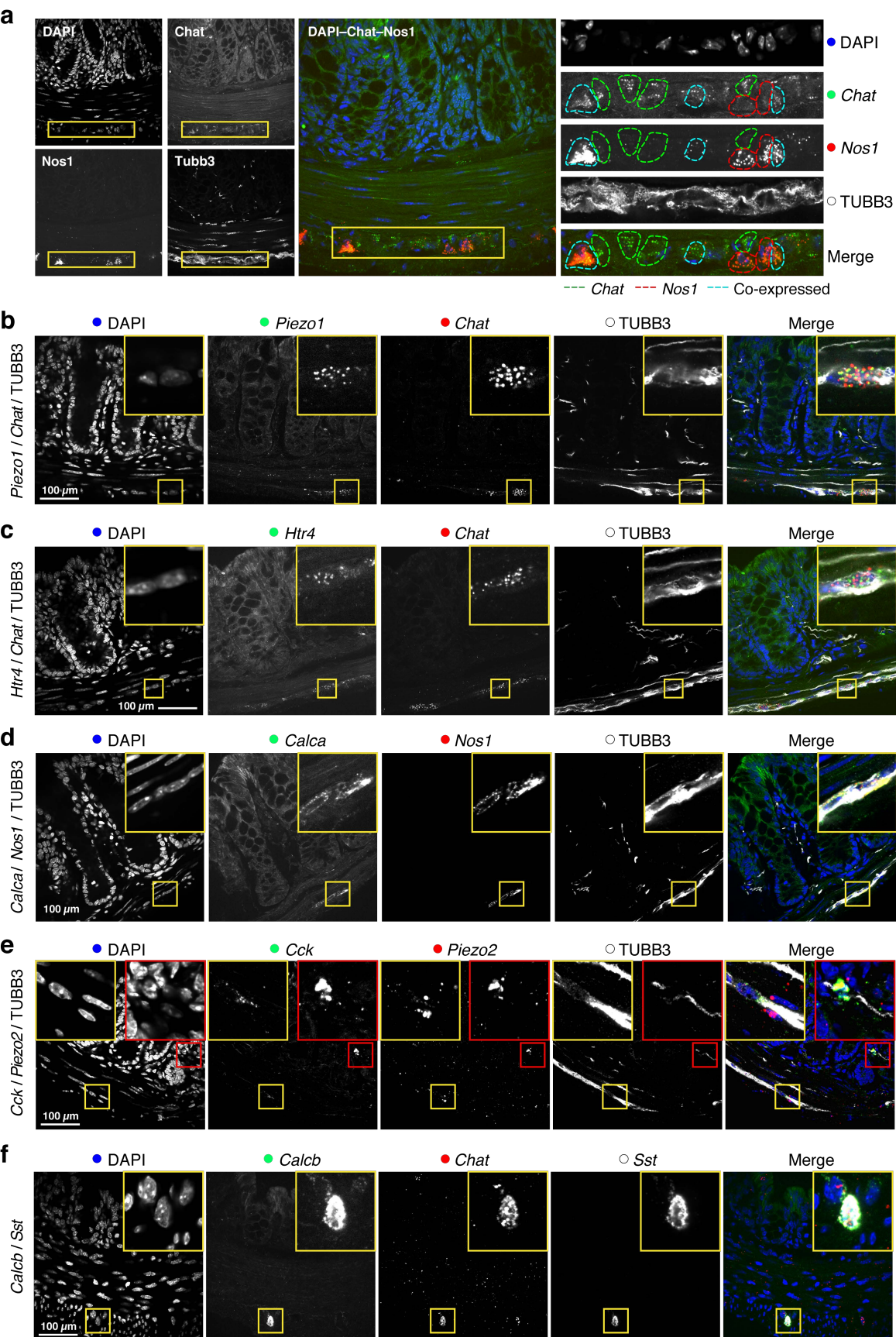
Supplementary Figure 3





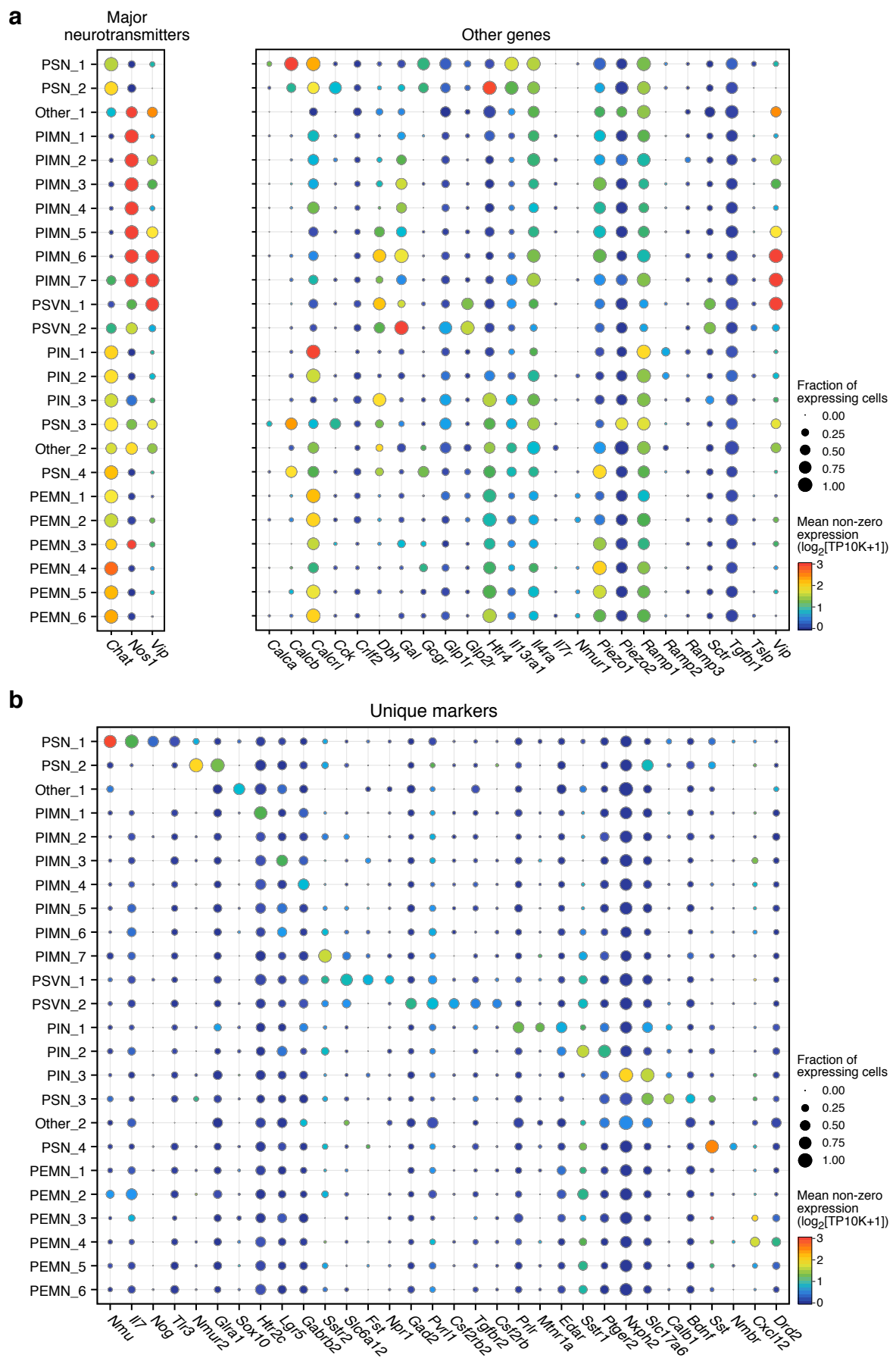
Resolution limited due to upload size. Higher resolution images available upon request

Supplementary Figure 5

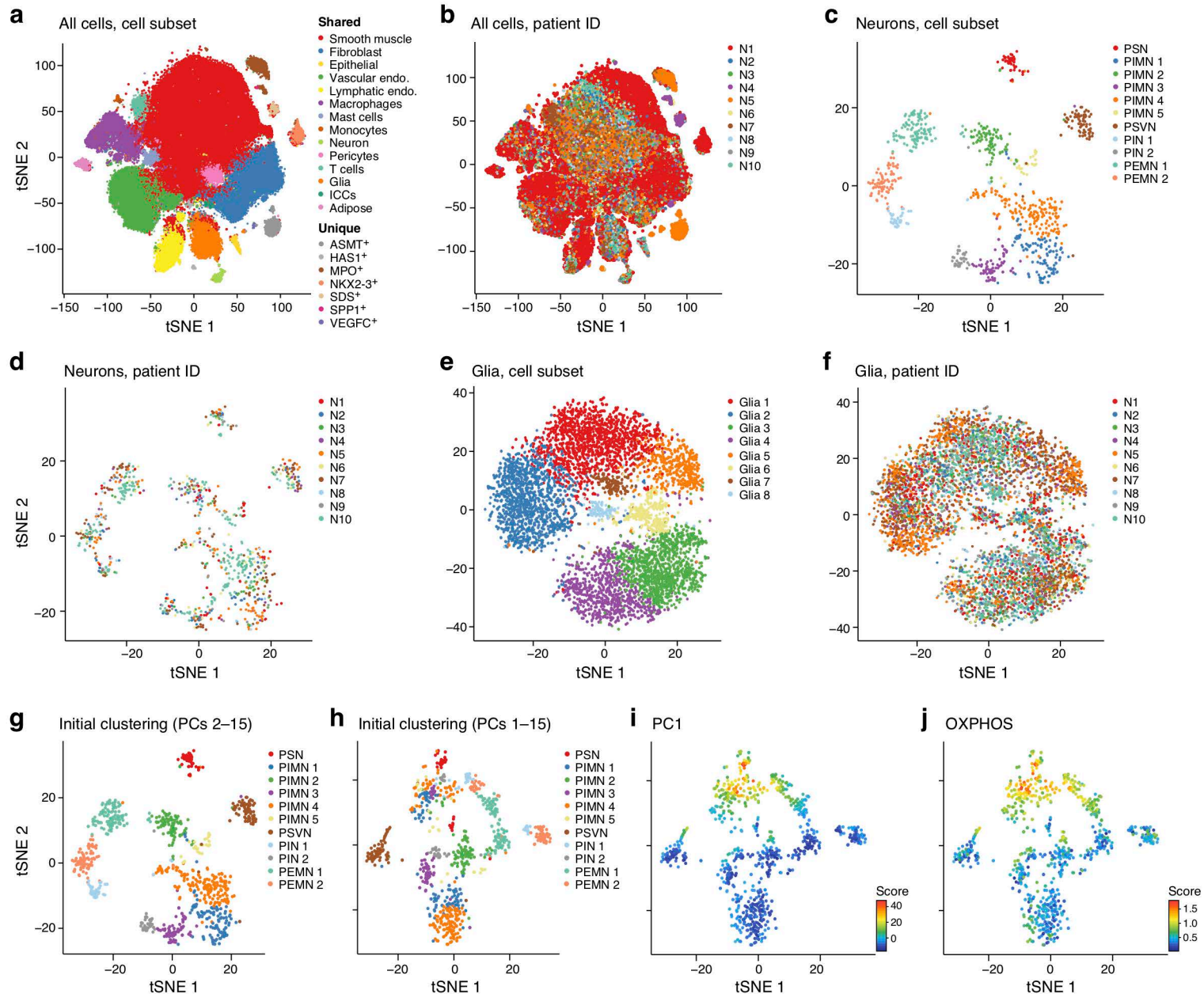


Resolution limited due to upload size. Higher resolution images available upon request

Supplementary Figure 6



Supplementary Figure 7



Supplementary Figure 8

