Ret loss-of-function decreases neural crest progenitor proliferation and restricts developmental fate potential during enteric nervous system development

Elizabeth Vincent [1,*], Sumantra Chatterjee [2,3,*], Gabrielle H. Cannon [1], Dallas Auer [1], Holly Ross [1], Aravinda Chakravarti [2,3,#], Loyal A. Goff [1,4,5,#]

Affiliations

1. McKusick-Nathans Department of Genetic Medicine, Johns Hopkins University School of Medicine
2. Center for Human Genetics and Genomics, NYU Grossman School of Medicine
3. Department of Neuroscience and Physiology, NYU Grossman School of Medicine
4. Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine
5. Kavli Neurodiscovery Institute, Johns Hopkins University

* Authors contributed equally

\# Co-corresponding authors

Abstract

The receptor tyrosine kinase gene *RET* plays a critical role in the fate specification of enteric neural crest-derived cells (ENCDCs) during enteric nervous system (ENS) development. Pathogenic *RET* loss of function (LoF) alleles are associated with Hirschsprung disease (HSCR), which is marked by aganglionosis of the gastrointestinal (GI) tract. ENCDCs invade the developing GI tract, proliferate, migrate caudally, and differentiate into all of the major ENS cell types. Although the major phenotypic consequences and the underlying transcriptional changes from Ret LoF in the developing ENS have been described, its cell type and state-specific effects are unknown. Consequently, we performed single-cell RNA sequencing (scRNA-seq) on an enriched population of ENCDCs isolated from the developing GI tract of *Ret* null heterozygous and homozygous mouse embryos at embryonic day (E)12.5 and E14.5. We demonstrate four significant findings: (1) *Ret*-expressing ENCDCs are a heterogeneous population composed of ENS progenitors as well as glial and neuronal committed cells; (2) neurons committed to a predominantly inhibitory motor neuron developmental trajectory are not produced under Ret LoF, leaving behind a mostly excitatory motor neuron developmental program; (3) HSCR-associated and *Ret* gene regulatory network genes exhibit distinct expression patterns across *Ret*-expressing ENCDC with their expression impacted by *Ret* LoF; and (4) *Ret* deficiency leads to precocious differentiation and reduction in the number of proliferating ENS precursors. Our results support a model in which Ret contributes to multiple distinct cellular phenotypes associated with the proper development of the ENS, including the

specification of inhibitory neuron subtypes, cell cycle dynamics of ENS progenitors, and the developmental timing of neuronal and glial commitment.

Summary Statement

Ret LoF affects proper development of the mouse ENS through multiple distinct cellular phenotypes including restriction of neuronal fate potential, disruption of ENCDC migration, and modulation of progenitor proliferation rate.

Introduction

The enteric nervous system (ENS) is primarily derived from migrating vagal and sacral neural crest cells (Kulkarni et al. 2020; Le Douarin and Smith 1988; Kulkarni et al. 2018; Durbec et al. 1996; Kapur et al. 1992) that colonize the developing gut tube between embryonic day (E)9.5 to E14.5 in mice (Young et al. 1998) and Carnegie stage (CS) 14 to CS22 in humans (Goldstein et al. 2013), leading to the neuronal networks of the myenteric and submucosal plexuses of the gut. The ENS has essential roles in propulsive and mixing motility, absorption, immunity, and vascular function (Furness 2012) and disruption of proper ENS development or organization underlies several disorders of the gastrointestinal system (Heuckeroth 2018). Previous studies of ENS development have demonstrated that invading enteric neural crest-derived cells (ENCDCs) are predominantly neuronal precursors early in mouse development (E10.5), but as development proceeds they differentiate into either glia, expressing classical glial markers including *Gfap, Plp1,* and *S100b* (Hao and Young 2009; Kang et al. 2021), or neurons (Hao and Young 2009). Single-cell transcriptional profiling of mouse ENCDCs during mid and late (E12.5, E15.5 and E18.5) developmental time points reflects this heterogeneity and temporal change in cell fate (Lasrado et al. 2017; Morarach et al. 2021, 2017; Elmentaite et al. 2021).

One of the key factors involved in the migration of ENCDCs into the developing gut, and their differentiation into enteric neurons and glia, is the receptor tyrosine kinase RET (Natarajan et al. 2002; Lasrado et al. 2017). RET expression in ENCDCs promotes proliferation and cell survival via the MAPK and PI3K/AKT pathways (Mulligan 2014). Consequently, gain-of-function (GoF) mutations in *RET* have

59 been associated with various cancers including medullary thyroid carcinoma and multiple endocrine
60 neoplasias type 2A and 2B (Chatterjee et al. 2016b). Conversely, *RET* LoF due to both coding and non-
61 coding regulatory variants leads to Hirschsprung disease (HSCR; congenital intestinal aganglionosis), an
62 uncommon (~1/5,000 live births) developmental disorder of the ENS characterized by absence of enteric
63 ganglia along variable lengths of the distal bowel (Tilghman et al. 2019; Chatterjee et al. 2016a; Kapoor
64 et al. 2015; Chatterjee et al. 2021; Alves et al. 2013). In the developing mouse gut, *Ret* deficiency leads
65 to significant transcriptional changes in transcription factors, signaling molecules, and genes involved in
66 specific transport and biosynthesis processes, leading to aganglionosis (Heanue and Pachnis 2006;
67 Chatterjee et al. 2019). Tissue-level bulk transcriptomic studies in *Ret*-deficient developing mouse gut
68 (Heanue and Pachnis 2006; Chatterjee et al. 2019)) have uncovered significant expression changes in
69 multiple transcription factors critical to neural crest migration and ENS differentiation including *Sox10,*
70 *Pax3* and *Phox2b* (Southard-Smith et al. 1998; Lang et al. 2000; Pattyn et al. 1999), as well as genes
71 harboring mutations in HSCR patients such as *Ednrb* and *Gdnf* (Heanue and Pachnis 2006; Chatterjee
72 et al. 2019).

73 Development and fate specification in the ENS occurs in conjunction with a prominent migration
74 along the forming gut tube (Kang et al. 2021; Nagy and Goldstein 2017). As such, temporal and spatial
75 influences affect ENCDC competence and can modulate fate decisions during development (Kang et al.
76 2021; Obermayr et al. 2013). While *Ret* is well known to play a significant role in the proper establishment
77 of the ENS, the cell-type-specific effects of *Ret* LoF on normal ENS development have not yet been fully
78 elucidated and could contribute to our understanding of normal ENS development.

79 To identify the cell-type-specific contributions of *Ret* to the proper organization and development
80 of the ENS, we examined the cellular heterogeneity of the ENS during mouse gut development in the
81 presence and absence of *Ret*. We profiled the transcriptome of 1,003 enteric neural crest-derived single
82 cells in heterozygous *Ret* null and prenatally lethal homozygous *Ret* null mice at E12.5 and E14.5. The
83 *Ret* heterozygous mice do not exhibit aganglionosis in the developing gut or any other known phenotypes
84 associated with *Ret* LOF and hence serves as an suitable control for our study which focusses on the

85 cellular dysregulation of enteric neurons and glia; the primary cell affected cell type due to *Ret* LoF.
86 ENCDCs and their progeny were isolated using a cyan fluorescent protein (CFP) reporter driven by the
87 *Ret* promoter for enrichment (Uesaka et al. 2008). Using a combination of differential gene expression
88 analysis, gene co-expression pattern analysis, RNA velocity, and cell cycle state predictions, we identify
89 several key developmental phenotypes associated with the loss of *Ret.* We find the specific effects of
90 *Ret* LoF vary across different cellular states in ENCDCs including effects on progenitor proliferation,
91 migration, differentiation, and fate specification. We confirm a previously described bifurcation in enteric
92 neuronal developmental fate specification (Morarach et al. 2021; Elmentaite et al. 2021) at an even earlier
93 time point. We further demonstrate that *Ret* expression is required for commitment to one of these two
94 neuronal fates and subsequently, the establishment of cell types from a predominantly inhibitory neuron
95 developmental trajectory. Additionally, we show that homozygous *Ret* LoF leads to precocious
96 differentiation of ENCDCs and a significant reduction in ENCDC progenitor proliferation rate. Each of
97 these observed phenotypes may independently contribute to the deficits in proper ENS cell numbers,
98 development, organization, and function observed in HSCR. These results highlight the diverse cell-type-
99 specific roles for *Ret*, a key developmental gene, in the proper establishment of the ENS with implications
100 for understanding HSCR-related phenotypes, and provide insights into the unique nature of ENS
101 development.

102 **Results**

103 To assess cell-type-specific roles for *Ret* in the developing ENS, we used a previously described
104 mouse line harboring a *Ret*-null fluorescent reporter allele (Ret^{CFP}) (Uesaka et al. 2008). All ENCDCs
105 express *Ret* during early development (Natarajan et al. 2002) allowing for their isolation via this reporter
106 system. We sorted *Ret* expressing ENCDC via FACS to decipher the cell autonomous effect of *Ret* LoF.
107 *Ret*-null ($Ret^{CFP/CFP}$) mouse embryos exhibit total colonic aganglionosis and kidney agenesis, and are
108 perinatal lethal (Uesaka et al. 2008). Importantly, no defects in ENS formation have been observed in
109 *Ret*-heterozygous ($Ret^{CFP/+}$) mice (Uesaka et al. 2008) nor any significant difference in the genome wide

110 expression observed when compared to wildtype GI tract at the same developmental times (Kapoor et
111 al. 2017). Comparison of *Ret*$^{CFP/CFP}$ to *Ret*$^{CFP/+}$ mice can therefore elucidate the mechanisms by which
112 *Ret* LoF affects the proper formation of the ENS.

113 To investigate transcriptional changes of ENCDCs, we performed single-cell RNA-sequencing
114 (scRNA-seq) on FACS-enriched populations of ENCDCs in the developing murine gut. Briefly, CFP-
115 positive single cells were collected from the dissociated stomach and intestine of male and female E12.5
116 and E14.5 *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ embryos and subjected to a modified Smart-seq2 scRNA-Seq library
117 preparation (Picelli et al. 2014; Chevée et al. 2018) (see Methods). Libraries from 95 cells for each of two
118 biological replicates were sequenced from all age, sex, and genotype combinations, with the exception
119 of E14.5 *Ret*$^{CFP/CFP}$ embryos. Consistent with a previously described reduction in the number of CFP^{+}
120 cells seen in a *Ret* hypomorphic mouse model (Uesaka et al. 2008), we observed reduction in the cell
121 capture rate in the E14.5 *Ret*$^{CFP/CFP}$ mice as compared to other ages and genotypes. We also observed
122 a pronounced sex bias in the number of CFP^{+} cells recovered from E14.5 *Ret*$^{CFP/CFP}$ embryos, with fewer
123 viable cells collected from males, consistent with the established sex bias observed in HSCR (Amiel et
124 al. 2008) and described in other mouse models of HSCR (Dang et al. 2011; Uesaka et al. 2008). This
125 reduced number of Ret$^{CFP/CFP}$ cells as a function of developmental age and sex, highlights a specific
126 challenge of this mouse model as the number of viable CFP^{+} cells rapidly diminishes in these conditions.
127 As such, only 22 cells from one male E14.5 *Ret*$^{CFP/CFP}$ embryo and 190 cells from four female E14.5
128 *Ret*$^{CFP/CFP}$ embryos were recovered (Table 1). Full-length cDNA scRNA-Seq libraries from 1,341
129 individual cells were sequenced to an average depth of 9.91×10^{5} (standard deviation 7.43×10^{5}) paired-
130 end reads per cell. Low quality cells were removed using standard metrics (see Methods) leaving 1,003
131 cells meeting all quality control parameters (Figure 1A; Table 1) with an average read depth of 1.16×10^{6}
132 (standard deviation 6.83×10^{5}) paired-end reads per cell. Relative FPKM values were converted to
133 estimated RNA copies per cell (CPC) using the CENSUS algorithm (Qiu et al. 2017). We identified a set
134 of 14,319 expressed genes with greater than zero CPC in at least 20 cells and a standard deviation
135 greater than zero across all cells. To visualize the transcriptional relationships between cells, we identified

136 a set of 419 high-variance genes to use as input to the uniform manifold approximation and projection
137 (UMAP) dimensionality reduction tool to establish a biologically relevant two-dimensional embedding of
138 the cells (Figure 1B, Figure S1).

139 ***Annotation of cell types and states of the developing ENS***

140 To quantify cell type and cell state differences with respect to genotype, we first used a Gaussian-
141 based clustering algorithm (Scrucca et al. 2016) to identify 9 clusters (Figure 1C) and calculated gene-
142 wise specificity scores to identify gene expression profiles for each cluster (Figure S2A). High specificity
143 scores relative to mean expression identified clusters of actively cycling progenitors (cluster 2, Hist1h
144 family genes; cluster 3, *Aurka/b*, *Cdk1*), a neurogenic commitment window (cluster 4, *Neurod4*; cluster
145 5, *Pcp2, Hes5*, cluster 6, *Hes6*, *Btg2*), maturing neurons (cluster 7, *Etv1*, *Vip*, *Gal, Cartpt*; cluster 8,
146 *Slc18a3 (ChAT), Npy, Snap25, Gfra2*; cluster 9, *Th, Dbh, Ddc*), as well as a population of maturing enteric
147 glia (cluster 1, *Plp1*, *Myl9*) (Figure S2B, Table 2). The neural progenitor cell markers p75 (*Ngfr*) and the
148 pan neuronal marker TuJ1 (*Tubb3*), which have traditionally been used to label mature enteric neurons
149 along with *Ret* (Hao and Young 2009; Young et al. 1999), have a distinct, largely non-overlapping
150 expression profile in *Ret*-expressing cells at these time points (Figure S2C). Consistent with this, we
151 observe *Ngfr* and *Tubb3* expression to be restricted to non-overlapping populations of cells at either end
152 of our learned manifold.

153 To annotate individual cells as either progenitors, developing glia, or developing neurons we next
154 evaluated the use of canonical marker gene expression for mature ENS cell types. However, during
155 development, the majority of cells may not yet express markers of mature, terminally differentiated cell
156 types. To address this limitation, we chose to assign cell types using a combination of transfer learning
157 and marker gene expression, where available, to predict the most likely annotation probabilities based
158 on an existing cellular atlas of adult mouse enteric neurons and glia (Zeisel et al. 2018). To obtain
159 transcriptional signatures of mature ENS cell types, we learned patterns of gene co-expression on the
160 annotated set of enteric neurons and glia from the public dataset using single-cell Coordinated Gene

161 Activity in Pattern Sets (scCoGAPS) (Stein-O'Brien et al. 2019) and trained a random forest (RF) model
162 on those learned patterns to predict cell type (see Methods). We next projected our single-cell data into
163 the latent spaces defined by these patterns using projectR (Stein-O'Brien et al. 2019; Sharma et al. 2019)
164 (Figure S3) and employed the RF model to classify cells in our dataset based on their projected pattern
165 weights, thereby agnostically annotating the cells in our data as glia or neuron (Figure S4A-C). We next
166 calculated cell type proportions by cluster (Figure S4D) and for clusters in which at least 90% of cells
167 shared an annotation, we transferred this majority annotation to all members of a given cluster. As a
168 result, clusters 1-5 (1 = 99% glia, 2 = 98%, 3 = 92%, 4 = 95%, 5 = 94%) were annotated as glia and
169 clusters 7 and 8 (7 = 98% neuron, 8 = 90%) were annotated as neurons. Annotations for clusters 6 (58%
170 glia) and 9 (63% neuron) were not altered.

171 Our single-cell gene expression profiling is a cross-sectional snapshot in developmental
172 progression for a population of cells that have not yet fully committed to mature cell type identities. To
173 infer the direction of state transitions for our ENCDC population, we next calculated RNA velocity
174 estimates (La Manno et al. 2018) and observed that the learned velocity fields comprise diverging
175 trajectories from a common origin (Figure 1F). Cells with high projection weights in patterns associated
176 with proliferating ENCDCs (pattern 40: *Brca1/2, Cdc6, Pola1/2*; pattern 41: *Ccna1/b1, Cenpe, Kif24*;
177 pattern 42: *Aurka/b, Bub1, Top2a*; Figure S3, Tables 3, 4, and 5) exhibit an ellipsoid pattern of velocities,
178 consistent with transition through the cell cycle. Extending from this progenitor population we observed
179 an population of cells expressing canonical markers consistent with progression towards a neuronal fate
180 (cluster 6: *Btg2*, *Hes6* Figure S2B) that further branches into two distinct trajectories of neuronal
181 specification (Figure 1F), consistent with previous observations of postmitotic branching of cellular fate
182 specification for neural crest-derived neurons in later gut development (Morarach et al. 2021). We infer
183 through shared marker gene expression that cluster 7 in our data corresponds to a previously described
184 developmental branch (branch A) (*Etv1*$^+$, *Nos1*$^+$, *Vip*$^+$, *Gal*$^+$) and cluster 8 corresponds to branch B
185 (*Bnc2*$^+$, *Dlx5*$^+$, *Mgat4c*$^+$, *Ndufa4l2*$^+$) (Figure S6A). These data highlight that *Ret*-expressing cells isolated
186 from the GI tract encompass two broad fates during development (glial-commitment and neuronal-

187 commitment) and confirm that neuronally-committed ENCDCs diverge post-mitotically into two distinct
188 developmental trajectories, outlining the diversity of cellular fate trajectories accessible to ENCDCs within
189 the developing mouse gut.

190 ***Ret LoF induces distinct transcriptional responses in different ENS cell states***

191 We first asked whether loss of *Ret* changes the proportion of ENCDC subtypes in our dataset. To
192 assess this, we tested for genotype bias in cluster composition by comparing the proportion of $Ret^{CFP/CFP}$
193 cells against the naïve (overall) proportion of 34%. Using a chi-squared test for equal proportions, we
194 observed significant deviations in seven of the nine identified clusters after Bonferroni correction for
195 multiple tests: clusters 1, 8, and 9 were enriched for $Ret^{CFP/CFP}$ cells (cluster 1: 83%, Bonferroni adjusted
196 $p=9.86x10^{-18}$; cluster 8: 57%, $p=5.64x10^{-4}$; cluster 9: 97%, $p=1.24x10^{-11}$), clusters 4-7 were depleted for
197 $Ret^{CFP/CFP}$ cells (cluster 4: 9%, Bonferroni adjusted $p=3.48x10^{-5}$; cluster 5: 20%, $p=1.00x10^{-3}$; cluster 6:
198 33%, $p=3.44x10^{-3}$; cluster 7: 2%, $p=6.33x10^{-5}$), while clusters 2 and 3 were relatively invariant (cluster 2:
199 42% $Ret^{CFP/CFP}$, Bonferroni adjusted p=0.196; cluster 3: 35%, p=1.0) (Figure 1D). Clusters 6, 7, and 8 are
200 neuronal-committed clusters and differ significantly from the naive proportion. Clusters 6 and 7 are
201 enriched for $Ret^{CFP/+}$ cells and 8 is enriched for $Ret^{CFP/CFP}$ cells, indicating strong genotypic effects
202 associated with neuronal fate specification. Three glial clusters, clusters 1, 4, and 5, differ significantly
203 from the naive proportion as well, with cluster 1 enriched and clusters 4 and 5 depleted for $Ret^{CFP/CFP}$
204 cells. Cluster-specific enrichment of genotype within the neuronal and glial populations suggests that
205 genotypic effects of *Ret* LoF modulates cell fate commitment with divergent results across ENCDC cell
206 states.

207 To identify the transcriptional effects of *Ret* LoF on the developing ENS, we tested 14,319 genes
208 and identified 520 as differentially expressed with respect to genotype across all cells (monocle likelihood
209 ratio test; Benjamini-Hochberg corrected q-value < 0.05, Figure 2A, Table 6A). Hierarchical clustering
210 based on expression of these 520 genes revealed that cells cluster first by cell type then by genotype,
211 again suggesting that the transcriptional changes observed are largely cell-type-specific. Gene ontology

212 analysis revealed enrichment for several relevant processes, including gliogenesis, cell fate commitment,
213 and axon development (hypergeometric test q-value < 0.05, Figure 2B, Table 7A). Significantly
214 upregulated genes expressed in *Ret*$^{CFP/CFP}$ progenitor/glia includes several associated with ENS glial
215 maturation (*Fabp7*, *Plp1*, *Myl9*), indicating that these progenitor/glia cells may be on average more
216 mature than their *Ret*$^{CFP/+}$ counterparts. A second cluster of genes is more highly expressed in *Ret*$^{CFP/+}$
217 neurons and includes genes with specific expression in the neurogenic branch A (*Cartpt, Vip, Nos1,*
218 *Etv1*), consistent with our observation that *Ret*$^{CFP/+}$ cells are significantly overrepresented in branch A
219 (cluster 7).

220 We next sought to map the transcriptional signatures learned from our ENCDC dataset to specific
221 cellular features by using coordinated sets of co-regulated genes. To this end, we applied scCoGAPS to
222 learn 25 patterns of gene co-expression from our ENCDC scRNA-Seq dataset (Figure S5, Table 8). We
223 calculated the Pearson correlation for pattern weights by cell against both known and learned cellular
224 annotations (Figure 2C, Table 9). This revealed patterns that correlated with technical (read depth and
225 alignment rates), experimental (genotype, age, and sex), as well as biological (cell type and cluster)
226 variables. We identified patterns correlated with the acquisition of different neuronal identities (Pattern
227 20: *Elavl4* (*HuD*)*, Nrxn3, Tubb3, Nnat, Syt1*; Pattern 21: *Nos1, Tubb2a, Mapt, Sncg, Ntng1*; Pattern 22:
228 *Onecut2, Gabrg2, Gabra3, Slc18a3* (*ChAT*)*, Gfra2, Elavl2* (*HuB*)*, Elavl3* (*HuC*)) (Figure S5; Table 10), a
229 transient neurogenic window (Pattern 19: *Pou4f1, Gadd45g, Neurod4, Hes6, Btg1/2, Sox11, Phox2b*),
230 as well as glial identity (Pattern 12: *S100b, Plp1, Fabp7*). In addition, we also identified patterns well
231 correlated with specific genotype-biased clusters of cells including Pattern 18 (*Ret, Gal, Ass1, Etv1*) and
232 Pattern 21 specifically used by Branch A neurons in cluster 7, and Pattern 12, specifically used by cells
233 in cluster 1 undergoing glial maturation and significantly enriched for *Ret*$^{CFP/CFP}$ cells.

234 Previous in vivo studies in mice have shown a severe migratory delay beginning as early as E9.5
235 in the absence of *Ret (Natarajan et al. 2002; Uesaka et al. 2008, 2015)*. Consistent with this observation,
236 we identified four gene expression patterns (Patterns 2, 18, 19, and 25), well correlated with genotype,
237 where *Hox* gene family members were identified as pattern markers (Stein-O'Brien et al. 2017) (Table

238 10). Pattern 25 markers include the more distal *Hoxa/c9* and *Hoxa/c10*, and the separation of these cells
239 from the main manifold of ENCDCs, combined with the specific expression of marker genes such as
240 *Maob* and *Pcdh10*, suggests that the handful of cells specifically using Pattern 25 are likely caudal
241 Schwann cell precursors. The enrichment for Hox genes within the remaining three pattern marker gene
242 lists with strong correlation to genotype, suggests that Hox gene expression in ENCDCs may be a
243 discriminating feature between the two genotypes. This was confirmed by the identification of seven *Hox*
244 genes, (*Hoxa2*, *b2*, *d3*, and the more caudally-expressed *Hoxa4*, *a5*, *b5*, and *b6*), as well as several
245 anterior-posterior pattern specification genes (*Dll1, Gpc3, Hes1, Lef1, Nrarp, Sema3c,* and *Tbx3*),
246 significantly differentially expressed (monocle likelihood ratio test; $q<0.05$) with respect to genotype
247 (Figure 2A-B). With the exception of *Gpc3, Lef1,* and *Sema3c*, all of the differentially expressed genes
248 associated with anterior-posterior positioning show increased expression in $Ret^{CFP/+}$ cells (Figure 2E).
249 Consistent with this and other previous work at E12.5 (Hirst et al. 2018), we observe that ENCDCs in the
250 absence of *Ret* have not left the extrinsically innervating nerve fibers of the stomach by E12.5 nor by
251 E14.5 (Figure 2D), in stark contrast to $Ret^{CFP/+}$ cells which have formed a dense network throughout the
252 stomach and all but the end of the hindgut by E12.5. As the *Hox* gene family is well-known to contribute
253 to axial patterning, this differential expression of *Hox* genes may therefore be a consequence of
254 significant rostro-caudal positional differences in ENCDCs in the presence and absence of *Ret*.

255 ***Ret is required for commitment to the branch A neuronal developmental trajectory***

256 Recent work in E15.5 mouse gut shows that after neuronal commitment, cells of the developing
257 ENS branch into two main trajectories, referred here as branch A (*Etv1+, Gal+, Nos1+*, *Vip+*) and branch
258 B (*Bnc2+*, *Dlx5+*, *Mgat4c+*, *Ndufa4l2+*) (Morarach et al. 2021). Our UMAP embedding reveals this
259 branching behavior can be observed as early as E12.5 and that the choice of branch is *Ret*-dependent
260 (Figure 3A, Figure S6A). As observed above, cluster 7, corresponding to branch A, is almost exclusively
261 composed of $Ret^{CFP/+}$ cells. Conversely, while cluster 8 (branch B) is significantly enriched for $Ret^{CFP/CFP}$
262 cells (Figure 1D), both genotypes are represented along the branch B trajectory. This suggests that there

263 is a significant bias in the choice of developmental trajectory mediated in part by *Ret* expression in
264 ENCDCs.

265 We next sought to identify branch-specific differences in gene expression with respect to
266 genotype. When we tested cluster 8 (branch B), we identified 12 genes differentially expressed with
267 respect to genotype (monocle likelihood ratio test, q<0.05, Figure 3B, Table 6B), four of which are more
268 highly expressed in $Ret^{CFP/+}$ and eight more highly expressed in $Ret^{CFP/CFP}$. Gene ontology analysis
269 indicates that genes upregulated in $Ret^{CFP/CFP}$ are enriched (hypergeometric test, q<0.05, Table 7B) for
270 dopamine biosynthetic processes ($q<4.52x10^{-6}$, *Agtr2*, *Ddc*, *Th*), catecholamine biosynthetic processes
271 ($q<6.44x10^{-6}$, *Agtr2*, *Ddc*, *Th*), and regulation of corticosteroid hormone secretion ($q<3.35x10^{-4}$, *Agtr2*,
272 *Tac1*). This suggests that, while both genotypes exhibit generally similar transcriptional profiles within
273 branch B, there may still be functional consequences of *Ret* LoF within this trajectory.

274 To examine differences between branch A and B trajectories and identify branch-specific markers,
275 we tested for differential expression irrespective of genotype to identify 111 genes (monocle likelihood
276 ratio test, q<0.05, Figure 3C, Table 6C). Gene ontology enrichment analysis (Table 7C) reveals that
277 genes with higher expression in cluster 7 are enriched for regulation of neurological system processes
278 (hypergeometric test, $q=1.91x10^{-5}$, *Cartpt, Ednrb, Fgfr1, Nos1, Nrg1, Vip*) and axonogenesis
279 (hypergeometric test, $q=6.11x10^{-5}$, *Ablim1, Auts2, Cdh11, Etv1, Lrrc4c, Lrrn3, Nrg1, Ntng1, Ret*), while
280 genes more highly expressed in cluster 8 are enriched for catecholamine biosynthetic processes
281 (hypergeometric test, $q=1.15x10^{-8}$, *Agtr2, Dbh, Ddc, Gata3, Gch1, Th*), regulation of neurotransmitter
282 levels (hypergeometric test, $q=2.00x10^{-6}$*, Chrna7, Dbh*, *Ddc, Gch1, Nrxn1, Prima1, Slc18a1, Sytl4, Th*),
283 and dopamine biosynthetic processes (hypergeometric test, $q=2.36x10^{-6}$*, Agtr2, Dbh*, *Ddc, Gch1, Th*).
284 The increased expression of genes in this latter geneset in $Ret^{CFP/CFP}$ cells in cluster 8 relative to $Ret^{CFP/+}$
285 cells (Figure 3B) suggests that differentiation and/or maturation rate of cells in branch B may be affected
286 by loss of *Ret*. Additionally, several established markers of branches A and B (Morarach et al. 2021) are
287 differentially expressed in accordance with expectations (*Etv1, Gal, Nos1,* and *Vip are* higher in cluster
288 7/branch A; *Bnc2, Dlx5,* and *Ndufa4l2* are higher in cluster 8/branch B).

289 Both Ret and CFP are downregulated along the branch B trajectory in addition to Gfra1 (Figure
290 3D), which encodes the canonical RET co-receptor, and Ednrb, which is known to be epistatic with Ret
291 (McCallion et al. 2003; Chatterjee and Chakravarti 2019; Carrasquillo et al. 2002). Interestingly,
292 expression of these same genes is maintained in $Ret^{CFP/+}$ cells along the branch A trajectory, and each
293 exhibits a biased expression for cells along this trajectory (Figure 3D). The exclusion of $Ret^{CFP/CFP}$ cells,
294 and the branch-A-specific persistence of Ret and its canonical co-receptor indicates that commitment to
295 the branch A developmental trajectory, and subsequently, establishment of the enteric neuronal subtypes
296 derived from cells adopting this trajectory, requires continued Ret expression. The downregulation of both
297 Ret and CFP expression along the branch B trajectory, which contains numerous cells of both genotypes
298 (Figure 3A&D), suggests that this trajectory may not require persistent Ret expression. While prior
299 immunohistochemical studies found Ret protein expression is down-regulated in emerging glia and
300 maintained in emerging neurons (Young et al., 1999), our scRNA-Seq data indicate that Ret exhibits
301 differential expression between the two branches of differentiating neurons with Ret expression
302 maintained in branch A and down-regulated in branch B, consistent with other recent single cell studies
303 (Elmentaite et al., 2021; Morarach et al., 2021). In aggregate, these results suggest that in the $Ret^{CFP/CFP}$
304 developing ENS, since branch A is an inaccessible developmental trajectory, and a greater portion of
305 neuronally committed ENCDCs may engage the branch B trajectory.

306 To validate Branch A inaccessibility in $Ret^{CFP/CFP}$ ENCDCs, we selected *Cartpt* (branch A/cluster
307 7) and *Gfra2* (branch B/cluster 8) from the list of differentially expressed genes as markers for *in vivo*
308 validations due to their high expression and high specificity in each branch (Figure 3D). We performed
309 whole-mount immunohistochemistry (IHC) on isolated gut (esophagus, stomach, and intestine) of E12.5
310 and E14.5, $Ret^{CFP/+}$ and $Ret^{CFP/CFP}$ embryos to identify and localize cells from each branch. We probed
311 for CART, the product of *Cartpt*, and CFP to localize branch A cells. In the $Ret^{CFP/+}$ samples at E12.5 we
312 observe co-expression of CART and CFP in the stomach (Figure 3E) but not in the esophagus (data not
313 shown). By E14.5, we observe co-expression of CART and CFP extending all the way to the hindgut of
314 $Ret^{CFP/+}$ samples (Figure 3F). Consistent with our scRNA-seq data, we do not observe any CART

315 expression in the stomach of *Ret*$^{CFP/CFP}$ samples at either age (Figure 3E-F) nor in the esophagus (data
316 not shown). We further validated the lack of branch A cells within the intestine of *Ret*$^{CFP/CFP}$ cells using
317 single-molecule fluorescent in situ hybridization (RNAscope) to assess mRNA expression of the branch
318 A marker genes *Cartpt* and *Gal* (Figure S6B-C). These results demonstrate that cells associated with the
319 branch A developmental trajectory are not produced within the GI tract in the absence of *Ret*.

320 To determine the spatial distribution of branch B cells we exploited the specific expression of
321 *Gfra2*, which encodes a non-canonical RET co-receptor, in branch B neurons, which is correlated with a
322 decrease in expression of *Gfra1* (Figure 3D). When we probed for GFRA2 and CFP via IHC we did not
323 observe GFRA2 expression in the E12.5 stomach of *Ret*$^{CFP/+}$ samples (Figure 3G). We did, however,
324 observe co-expression of GFRA2 and CFP restricted to cells along the extrinsically innervating nerve
325 fibers of the stomach of *Ret*$^{CFP/CFP}$ mice at this age (Figure 3G), confirming that branch B neuronal
326 development trajectory is indeed accessible in the absence of *Ret*. At E12.5 we did not observe GFRA2
327 outside the extrinsically innervating nerve fibers of the *Ret*$^{CFP/CFP}$ stomach, nor in the intestines of either
328 genotype (data not shown), consistent with the previously described migratory defect and failure to
329 effectively colonize the ENS in the *Ret*$^{CFP/CFP}$ mice. We were able to detect CFP^{+}/GFRA2^{+} cells (putative
330 branch B neurons) in the esophagus of both genotypes, as well as CFP^{+}/GFRA2^{-} (non-neuronal
331 ENCDCs) and CFP^{-}/GFRA2^{+} (maturing branch B neurons) cells (Figure S6D), indicating that branch B
332 cells are spatially restricted to the esophagus in E12.5 *Ret*$^{CFP/+}$ samples and to the esophagus and
333 extrinsic nerve fibers along the stomach in E12.5 *Ret*$^{CFP/CFP}$. By E14.5, CFP^{+}/GFRA2^{+} branch B maturing
334 neurons are observed throughout the intestine in the *Ret*$^{CFP/+}$; however, CFP^{+}/GFRA2^{+} cells remain
335 restricted to the extrinsically innervating nerve fibers along the stomach in the *Ret*$^{CFP/CFP}$ mice (Figure
336 3H). Consistent with this difference, at E12.5 we see proportionally more *Ret*$^{CFP/CFP}$ cells in cluster 8
337 (branch B) than *Ret*$^{CFP/+}$ within our scRNA-Seq dataset (9/37 *Ret*$^{CFP/+}$, 28/37 *Ret*$^{CFP/CFP}$). By E14.5
338 however, *Ret*$^{CFP/+}$ are the predominant cells in cluster 8 (21/33 *Ret*$^{CFP/+}$, 12/33 *Ret*$^{CFP/CFP}$). This could be
339 related to an overall reduced number of E14.5 *Ret*$^{CFP/CFP}$ cells in the intestine, or reduced capture

340 efficiency as fluorescence decreases over biological time along this trajectory due to the downregulation
341 of *Ret (*and by proxy *CFP*) expression as cells progress along branch B (Figure 3D).

342 These data indicate that branch A and branch B represent two transcriptionally distinct
343 developmental trajectories of maturing ENS neuronal subtypes that colonize the length of the intestines
344 in the $Ret^{CFP/+}$ mice. Branch A neurons are never produced in the $Ret^{CFP/CFP}$ mice. Furthermore, Branch
345 B neurons are produced, at a potentially faster developmental rate, but remain restricted to the
346 extrinsically innervating nerve fibers throughout development in the $Ret^{CFP/CFP}$ mouse. As *Ret* is required
347 for cells to migrate into the developing ENS (Figure 2D) (Hirst et al. 2018) it follows that branch A, which
348 appears to require *Ret*, would be inaccessible while branch B, which is *Ret*-independent, remains
349 accessible in the absence of *Ret.*

350 ***HSCR-associated genes are expressed across different cellular states in the developing ENS***

351 While coding and enhancer LoF variants in *RET* contribute to ~50% of cases of Hirschsprung
352 disease, HSCR is also associated with rare coding pathogenic alleles (PAs) in 23 additional genes
353 (*GDNF, GFRA1, NRTN, SOX10, EDNRB, EDN3, ECE1, ZEB2, PHOX2B, KIFBP, TCF4, L1CAM,*
354 *IKBKAP, SEMA3C, SEMA3D, NRG1, ACSS2, FAM213A, ADAMTS17, ENO3, SH3PXD2A, SLC27A4*
355 and *UBR4*) (Tilghman et al. 2019; Chatterjee et al. 2016a; Kapoor et al. 2015; Chatterjee et al. 2021).
356 Many of these genes are a part of the *RET-EDNRB* gene regulatory network (GRN) (Chatterjee et al.
357 2016a). To assess the distribution of HSCR-associated genes, we investigated the expression pattern of
358 all 24 genes in our data to determine their co-expression patterns and their specific effects due to *Ret*
359 LoF (Figure 4A). Three HSCR-associated genes (*Gdnf, Edn3, and Nrtn*) are not expressed in our ENCDC
360 dataset as they are predominantly expressed from the surrounding gut mesenchyme. Of the remaining
361 21 expressed genes (Figure S7A-C), six are significantly differentially expressed (monocle likelihood ratio
362 test, $q<0.1$, Figure 4B-C, Table 6A) between $Ret^{CFP/+}$ and $Ret^{CFP/CFP}$ cells (*Gfra1, Nrg1, Phox2b, Ednrb,*
363 *Sema3c*, and *Sox10)*. To identify cell-type-specific differential effects from *Ret* LoF we further tested the
364 combinatorial effect of genotype and cell type. Three HSCR-associated genes (*Gfra1, Sema3d* and

365 *Ednrb*) exhibit significant cell-type-specific differential expression (monocle likelihood ratio test, q<0.05,
366 Figure 4B-C, Table 6D). Both *Gfra1 and Ednrb* reach statistical significance in both tests, suggesting a
367 cell autonomous effect of *Ret* LoF on the expression of these two genes. Alternatively, four HSCR-
368 associated genes (*Nrg1, Phox2b, Sema3c*, and *Sox10*) are globally differentially expressed but are not
369 differentially expressed with respect to genotype:cell type interactions (monocle likelihood ratio test,
370 q<0.05; Table 6A,D).

371 We have previously shown that the proportion of cells by genotype deviates significantly from the
372 naïve proportion in seven of the nine clusters in our data (Figure 1D), we also observe that genotype
373 proportions deviate significantly from expectation in E14.5 neurons and approach statistical significance
374 in E14.5 glia (chi-squared test for equal proportions; neurons/p=0.0158, glia/p=0.0702) (Figure 4D).
375 Several of the HSCR-associated genes (*Ednrb, Gfra1, Elp1, Sox10, Zeb2*) are expressed predominantly
376 in progenitor/glia cells (Figure S7A,C). *Phox2b, Tcf4,* and *Ubr4* are the three within this set that are
377 ubiquitously expressed across all clusters in our embedding (Figure S7A,C). *L1cam* is enriched in both
378 maturing neurons and maturing glia, but exhibits reduced expression within the cycling progenitor
379 population of cells (Figure S7A,C). *Ece1* is predominantly expressed in the population of maturing glial
380 cells. These results suggest that some of the observed differential expression of HSCR-associated genes
381 when observed at a global level, but not when assessed within discrete cell states, may be primarily due
382 to differences in cell type proportions across genotypes. Finally, similar to the observed persistence of
383 *Ret* expression along Branch A, a subset of HSCR-associated genes are differentially expressed
384 (monocle likelihood ratio test; q<0.05, Table 6C) between the branches, with higher expression in branch
385 A relative to branch B (*Ednrb, Gfra1, Nrg1*, and *Ret*), suggesting their potential role in the establishment
386 of Branch A identity in conjunction with *Ret*.

387 Despite significant genetic heterogeneity, ~67% of HSCR patients have mutations in the genes
388 of the *RET-EDNRB* GRN (Tilghman et al. 2019). This GRN includes bidirectional transcriptional feedback
389 between *RET and EDNRB,* and the transcription factors *SOX10, GATA2, RARB,* and *NKX2-5*, as well
390 as *GFRA1* and its physiological ligand, *GDNF*, and the ubiquitin ligase *CBL*, which breaks down

391 phosphorylated RET (Chatterjee and Chakravarti 2019; Chatterjee et al. 2016a, 2019). To ascertain
392 whether this feedback is cell autonomous or not we looked at the expression of these genes in our
393 dataset. *Nkx2-5* and *Gdnf* are not expressed in our ENCDC dataset. Of the remaining six genes, four
394 (*Gata2, Gfra1, Rarb, and Sox10*) are significantly differentially expressed between $Ret^{CFP/+}$ and $Ret^{CFP/CFP}$
395 ENCDCs (Table 6A), with *Gfra1* and *Rarb* more highly expressed in $Ret^{CFP/+}$ cells and *Sox10 and Gata2*
396 more highly expressed in $Ret^{CFP/CFP}$ cells, although these moderate differences in expression may also
397 reflect a disproportionate distribution of ENS cell types within our dataset with respect to genotype (Figure
398 4C-D, Figure S7C). Finally, with the exception of *Sox10*, which rapidly decreases expression during
399 neuronal commitment, and *Gata2*, which is specific to cluster 9, *Ednrb, Rarb, Gfra1*, and *Cbl* all exhibit
400 persistent expression along the branch A neuronal developmental trajectory (Figure 4B, Figure S7C),
401 consistent with a role for the *RET-EDNRB* GRN in establishing branch A identity. The expression of
402 *Sox10* exclusively within the progenitor population explains a previous observation that *Sox10*
403 haploinsufficiency only affects maintenance of progenitor cells and not neurons (Paratore et al. 2002).
404 These results highlight the diverse cellular states in which HSCR-associated genes are
405 expressed, suggesting that their roles in HSCR disease etiopathology may involve modulation of different
406 aspects of ENCDC proliferation and development. Collectively however, expression of *RET-EDNRB*
407 GRN genes is affected by the loss of Ret signalling and is persistent along the Branch A trajectory, which
408 is inaccessible in the $Ret^{CFP/CFP}$ mouse.

409 ***Ret LoF reduces the number of actively proliferating ENCDCs in the developing gut***

410 Reduction in proliferation capacity and precocious differentiation as a result of LoF of other genes
411 within the developing ENS have been proposed as significant drivers of colonic aganglionosis (Fujiwara
412 et al. 2018; Bergeron et al. 2016). Given the previously established role for *RET* as a proto-oncogene
413 (Grieco et al. 1990; Mulligan 2014) and its association with increased proliferation in other context (Park
414 and Bolton 2017; Perea et al. 2017; Gattelli et al. 2013), we asked whether *Ret* LoF was associated with
415 significant differences in the proliferation rate of ENCDC progenitors. Differential gene expression

416 analysis with respect to both genotype and cell type identified 331 genes (monocle likelihood ratio test,
417 q<0.05; Figure S8A, Table 6D), consisting of several genes associated with the mitotic cell cycle (C*dkn1c,*
418 *Cdk1, Top2a, Ccnd1, Cenpk, Cenpe, Cenpf, Ccna2, Ccnb1, Ccnb2*), DNA replication (*Pcna, Mcm3,*
419 *Mcm5, Mcm6, Mcm10, Gmnn*), and cell cycle regulation (*Rgcc, Mad2l1, Chek2, Rad51ap1*). Gene set
420 enrichment analysis of this differential gene list identified several cell cycle-associated gene sets
421 including mitotic cell cycle processes, cell division, DNA replication, regulation of cell cycle as top scoring
422 Gene Ontology Biological Processes (GO:BP) (hypergeometric test, q<0.05; Table 7D). To focus on cells
423 that are potentially engaging the cell cycle, we next performed a differential gene expression analysis on
424 just the subset of ENCDCs identified as progenitors or glia within the contiguous manifold (i.e. excluding
425 ENCDCs committed to a postmitotic neuronal fate and excluding cluster 9). We identified 364 genes with
426 significant differential expression with respect to genotype (monocle likelihood ratio test, q<0.05; Figure
427 5A, Table 6E), and again this list was enriched for gene sets associated with cell cycle regulation, in
428 addition to gliogenesis, cell fate commitment, and neural crest cell differentiation (hypergeometric test,
429 q<0.05; Figure 5B, Table 7E). Thus, both cell cycle occupancy and differentiation state, two processes
430 that are intrinsically linked during developmental fate specification, are disrupted in the *Ret*$^{CFP/CFP}$
431 progenitor population relative to *Ret*$^{CFP/+}$.

432 We next asked whether we could identify any significant differences in cell cycle state
433 assignments for ENCDCs in the absence of Ret. We used the tricycle package (Zheng et al. 2021) to
434 project the 1,003 single ENCDCs into the default reference cell cycle space provided (Figure 5C). This
435 analysis allows us to estimate the continuous position of each cell along the cell cycle (Θ) (Table 11),
436 comparisons of which by genotype and age (Figure 5D) suggests that at E12.5, there are minimal
437 differences in cell cycle occupancy between genotypes (Kolmogorov-Smirnov test; p-value = 0.660), but
438 by E14.5 there are consistently fewer *Ret*$^{CFP/CFP}$ cells engaging the cell cycle than *Ret*$^{CFP/+}$(Kolmogorov-
439 Smirnov test; p-value = 8.28×10^{-3}).

440 To further assess whether *Ret*$^{CFP/CFP}$ ENCDCs exhibit differential cell cycle occupancy relative to
441 *Ret*$^{CFP/+}$ we next sought to infer cell cycle stage annotations for cells in our dataset using a projection-

442 based approach with an independent, biologically relevant data source. Using the scCoGAPS patterns
443 previously learned from publicly available data, and on which we trained the random forest classification
444 model, we observed high usage of three patterns within a population of cycling glia. Gene ontology
445 analysis of pattern markers revealed these patterns correspond to different proliferative phases of the
446 cell cycle (hypergeometric test, $q<0.05$, Table 4A&B). Pattern 40 was associated with DNA replication,
447 corresponding to S phase; pattern 41 associated with cilium assembly, organization, and segregation,
448 and attachment of spindle microtubules to kinetochores, corresponding to prophase or metaphase (early
449 M phase); and pattern 42 associated with chromatid segregation and nuclear division, corresponding to
450 anaphase or telophase (late M phase). To predict the cell cycle stage for individual cells in our ENCDC
451 data, we scaled projected pattern weights linearly within each pattern to range between 0 and 1, and
452 used this as a measure of cell cycle stage occupancy probability (Figure S8C&D). We tested all six
453 possible comparisons of the four interaction conditions of genotype and age per pattern for differences
454 in probabilities using two-sample Kolmogorov-Smirnov tests on the empirical cumulative density functions
455 (Figure S8E), and used the Bonferroni method for multiple testing correction. Across the three patterns,
456 E14.5 cells have significantly lower proliferative cell cycle stage occupancy probabilities than E12.5 cells
457 of the same genotype (S phase: $Ret^{CFP/+}$ Bonferroni adjusted $p \leq 1.32 \times 10^{-15}$, $Ret^{CFP/CFP}$ $p \leq 2.00 \times 10^{-15}$; early
458 M: $Ret^{CFP/+}$ $p \leq 1.32 \times 10^{-15}$, $Ret^{CFP/CFP}$ $p \leq 3.84 \times 10^{-13}$; late M: $Ret^{CFP/+}$ $p \leq 1.38 \times 10^{-8}$, $Ret^{CFP/CFP}$ $p \leq 1.94 \times 10^{-}$
459 ^{9}), consistent with the expectation that as progenitor cells differentiate, the proportion of actively cycling
460 cells will decrease over biological time. Again, cell cycle stage occupancy probabilities for ENCDCs did
461 not differ significantly with respect to genotype at E12.5. By E14.5 however, $Ret^{CFP/CFP}$ cells have
462 significantly lower occupancy scores across all three patterns (S phase: Bonferroni adjusted $p \leq 9.66 \times 10^{-}$
463 ^{5}, early M: $p \leq 1.97 \times 10^{-4}$; late M: $p \leq 0.0193$) suggesting overall fewer cycling ENCDCs in the $Ret^{CFP/CFP}$.

464 This may be attributed to an increased differentiation rate, and consequently decreased proliferation rate,
465 in *Ret*$^{CFP/CFP}$ cells compared to *Ret*$^{CFP/+}$ cells between E12.5 and E14.5.

466 To validate this predicted decrease in cell cycle occupancy in *Ret*$^{CFP/CFP}$ ENCDCs, we next used
467 IHC to assess the number of CFP^{+} ENCDC co-expressing the mitotic marker phosphohistone H3 (pH3)
468 within the developing gut at E14.5 (Figure 5E). We confirmed a significant (Student's t-test, $p<1.0x10^{-4}$)
469 overall reduction in the number of CFP^{+} cells in the *Ret*$^{CFP/CFP}$ mouse model at E14.5 (Figure S8G)
470 consistent with the reduced number of CFP^{+} cells observed in our FACS-collected population (Figure 1A)
471 and previously published studies (Uesaka et al. 2008). We also identified a slight difference in the mean
472 number of pH3^{+} cells at E14.5 (Student's T-test, $p<0.05$); however, modest enough to suggest that the
473 overall number of mitotic cells within the developing gut remains unaffected (Figure S8H). This is
474 consistent with a cell-autonomous effect of *Ret* on cell cycle as ENCDCs make up a small portion of all
475 proliferating cells in the developing gut and would therefore not contribute significantly to the overall
476 proliferative capacity of the gut as a whole. The overall number of double positive pH3^{+}/CFP^{+} cells was
477 significantly reduced in *Ret*$^{CFP/CFP}$ animals (Figure 5F & S7I) (Student's T-test, $p=0.029$), and while not
478 statistically significant, we observed a difference in the mean percentage of pH3^{+} cells within the subset
479 of CFP^{+} ENCDCs in the *Ret*$^{CFP/CFP}$ ENS at E14.5 indicating that a smaller fraction of developing ENCDCs
480 are actively engaging the cell cycle in the absence of *Ret* (Figure 5G & Figure S8J) (Student's t-test,
481 $p=0.11$). In aggregate, these results are consistent with the previously established role of *Ret* in
482 enhancing proliferation and indicate that a loss of *Ret* in the developing ENS may contribute to an early
483 reduction in the progenitor pool, possibly through induction of precocious differentiation, leading to a
484 failure to colonize the more distal region of the developing gut.

485 **Discussion**

486 The enteric nervous system (ENS), initially derived entirely from the caudally migrating vagal and
487 cranially migrating sacral neural crest cells, contains more than 100 million neurons of a variety of
488 functionally distinct subtypes, highlighting the breadth of the functions it performs (Furness 2008; Rao

489 and Gershon 2016). In contrast to the rest of the peripheral nervous system, the ENS is endowed with
490 the ability to integrate multiple neuronal signals to control gastrointestinal (GI) behavior independently
491 from the brain and spinal cord (Furness 2008; Gershon 2010). One of the critical genes that enables this
492 transition of ENCDCs to diverse types of neurons is *RET*, which encodes a cell surface receptor tyrosine
493 kinase. *RET* LoF leads to colonic aganglionosis and is a known factor in the etiology of Hirschsprung
494 disease.

495 In spite of the critical role *Ret* plays in establishing the normal ENS, there is a limited understanding
496 of the diversity of *Ret*-expressing cells in the developing GI tract and the precise nature of the cell-type-
497 specific effects mediated by *Ret* LoF. By performing scRNA-seq on an enriched population of *Ret*-
498 expressing cells during ENS development we have generated a *Ret*-dependent fate transition map of
499 ENCDCs as they generate diverse classes of neurons and glia in the mouse GI tract to decipher the
500 specific cellular phenotypes mediated by *Ret* LoF.

501 ***Ret LoF affects the establishment of a specific developmental neuronal trajectory in the***
502 ***developing GI tract***

503 Our cellular reconstruction of ENCDC differentiation demonstrates that *Ret*-expressing cells
504 broadly acquire neuronal and glial fates early during normal development of the GI tract and helps explain
505 why *Ret* LoF leads to total aganglionosis caused by impaired migration, defects in proliferation, and loss
506 of immature ENCDCs (Schuchardt et al. 1994). Our data highlight that *Ret* is required for the induction of
507 an entire neuronal specification trajectory within the developing ENS which has previously been
508 associated with the specific development of at least two distinct subtypes of inhibitory motor neurons: an
509 interneuron subtype, and a population of intrinsic primary afferent neurons (IPANs) (Morarach et al. 2021).
510 Though previous studies have posited that branch A corresponds to a predominantly inhibitory motor
511 neuron (MN) trajectory and branch B corresponds to a predominantly excitatory MN trajectory (Morarach
512 et al. 2021), our data are derived from embryonic timepoints which are too early to directly support this
513 functional distinction between the branches. At E12.5, we observe putative branch B cells ($GFRA2^+/CFP^+$)

514 in the esophagus of the *Ret*$^{CFP/+}$, and in the esophagus and stomach of the *Ret*$^{CFP/CFP}$ mice (Figure 3F,
515 Figure S6C). We also observe putative branch A cells (CART^{+}/CFP^{+}) in the stomach of *Ret*$^{CFP/+}$ mice but
516 not the esophagus, and we do not observe putative branch A cells anywhere in the *Ret*$^{CFP/CFP}$ GI tract at
517 E12.5. By E14.5 we observe putative branch B cells in the stomach of the *Ret*$^{CFP/CFP}$, and throughout the
518 gut up to and including the wave front of migration in *Ret*$^{CFP/+}$ mice. The earlier establishment of branch B
519 cells within the stomach of the *Ret*$^{CFP/CFP}$ is consistent with a model of precocious differentiation we identify
520 in this study. The specific failure to produce cells along branch A in the *Ret*$^{CFP/CFP}$ suggests that the subset
521 of mature ENS neuronal subtypes derived from this developmental trajectory may be the most acutely
522 affected in disorders such as HSCR and should be the focus of future studies of other *RET*-associated
523 pathologies in the gut.

524 We provide evidence that aganglionosis observed in HSCR may arise from a combination of early
525 exit from the cell cycle for ENS progenitors, leading to precocious differentiation, and a failure to produce
526 the full repertoire of mature ENS cell types. We do not detect any transcriptional signatures of cell death,
527 suggesting that at this early time point apoptosis or necrosis may not be a significant contributing factor
528 in *Ret*-deficient aganglionosis. The presence of glial cells in our mouse models, independent of genotype,
529 is consistent with previous observations in HSCR patients that immature glial cells can be detected within
530 the aganglionic portion of the colon (Tani et al. 2017). Our data suggest that *Ret* expression is not required
531 for the establishment of a normal ENS glial identity; however, the observed effects on proliferation and
532 differentiation potential suggest that enteric glia may be a preferred fate in *Ret* LoF models.

533 ***Neuronal fates that use non-canonical RET co-receptors remain accessible in the absence of Ret***

534 A previous scRNA-seq study of the developing ENS at E15.5, when the migration of ENCDC
535 through the length of the gastrointestinal tract is near complete, reports a divergence in maturing neurons
536 that results in two branches of progenitors that produce different cell fates (Morarach et al. 2021). We
537 observe this divergence of the neuronal trajectories as early as E12.5, when ENCDCs are still migrating
538 and undergoing neuronal maturation. While we observe a near complete absence of *Ret*$^{CFP/CFP}$ cells in

539 branch A (cluster 7), it is interesting to note that branch B (cluster 8) contains maturing neurons from both
540 genotypes and is not grossly affected by the loss of *Ret*. As cells progress along branch B there is a
541 general reduction in the expression of *Ret*, *CFP*, and several *RET-EDNRB* GRN genes, including *Gfra1.*
542 Concomitant with this, cells in branch B subsequently upregulate *Gfra2*, which encodes a non-canonical
543 RET co-receptor. Though GFRA1 and GFRA2 are in the same family of receptors, GFRA1 preferentially
544 binds GDNF as its ligand, while GFRA2 preferentially binds the ligand NRTN. It is possible that GFRA2
545 is binding to residual RET after downregulation of transcription; however, it raises the possibility that
546 GFRA2 complexes with another co-receptor to activate an alternate signaling pathway responding to a
547 different ligand in this *Ret*-independent trajectory. Differential expression of *Gfra* family genes within the
548 aganglionic segment of HSCR patients has previously been observed (Lui et al. 2002), including a
549 specific lack of *GFRA1* expression in neuronal cells within the associated hypertrophied nerve fibers.
550 Additionally, *Ntrn* and *Gfra2* null mice exhibit a reduction in neuron fiber density and neuronal size in the
551 ENS, and in each case, result in the specific loss of cholinergic excitatory myenteric innervation
552 (presumably branch B-derived) within the mouse gut, leaving the nitrergic neuronal population (branch
553 A-derived) largely unaffected (Rossi et al. 1999; Heuckeroth et al. 1999), consistent with the segregation
554 of fates and corresponding gene expression differences described here. In aggregate, these results
555 suggest that the canonical signaling pathway of RET, GDNF and GFRA1 is responsible for the proper
556 establishment of ENS subtypes derived from the branch A developmental trajectory, and that the activity
557 of NRTN-GFRA2 signaling may play a greater role in the establishment of ENS neuron identities along
558 branch B, potentially independent from *Ret*.

559 ***Expression patterns of HSCR-associated genes highlight necessity of disease-relevant cellular***
560 ***context***

561 The specific expression patterns of many genes that harbor pathogenic alleles in HSCR patients
562 within the invading and developing ENCDCs of the ENS should lead us to reassess the nature of
563 aganglionosis in HSCR. Our data highlight the need to better understand the specific cellular states and

564 contexts in which these genes and GRNs operate, as well as the cell autonomous versus cell non-
565 autonomous effects in HSCR. Genes such as *Acss2, Eno2,* and *Adamts17* express very lowly in the ENS
566 population (Figure 4), yet are robustly expressed in both the mouse and human developing GI tract
567 (Tilghman et al. 2019; Chatterjee et al. 2019). Furthermore, while the nature of our study, using only
568 sorted *Ret*-expressing cells is by definition only able to capture cell-autonomous effects, additional non-
569 cell-autonomous effects have been observed in *Ret* LoF studies as well (Bogni et al. 2008). Hence
570 patients carrying pathogenic alleles in the genes expressed in other, non-ENCDC, may contribute to
571 other GI anomalies, and aganglionosis may then be a secondary effect rather than primary. Though
572 verification of this would require detailed patient phenotyping, the shift between neuronal trajectories we
573 have identified suggests developmental mechanisms that may guide future analysis of patient samples.

574 ***Ret LoF leads to proliferation defects in ENS progenitors***

575 During formation of the ENS, the gut itself is increasing dramatically in both size and length,
576 creating a huge proliferative burden on ENCDCs, the progenitors of the ENS. Our in silico estimates of
577 cell cycle stage occupancy in conjunction with an *in vivo* assay of proliferation suggest a decrease in
578 proliferation of ENCDCs in the absence of *Ret*. Even a small reduction in proliferative capacity could
579 produce a significant decrease in population size over the course of the several days it takes to colonize
580 the ENS, such as the absence of CFP+ cells we observe in the $Ret^{CFP/CFP}$ mouse intestines.

581 *Ret* deletion in the mouse leads to significant changes in fate acquisition, progenitor proliferation,
582 and ENCDC migration within the developing ENS. It is likely that *RET* deficiency alleles in humans lead
583 to similar phenotypes in distinct subsets of ENS cell types and states in HSCR patients, and that each of
584 these phenotypes may play some combinatorial role in establishing the aganglionosis that is the hallmark
585 of this disease. Furthermore, different *RET*-deficiency alleles and epistatic interactions may affect these
586 phenotypes in different ways, potentially explaining some of the observed variable expressivity and
587 incomplete penetrance of HSCR. Our study highlights the specific roles that *Ret* plays within different
588 cellular contexts of the developing enteric nervous system, and demonstrates that appropriate cellular

589 context is critical for the interpretation of multiple disease-related cellular phenotypes that may emerge

590 from a single causal insult.

591

592

593 ## Methods

594 ***Mouse strains:***

595 Mice heterozygous for the conditional *Ret* allele (*Ret*$^{fl/+}$)*(Gould et al. 2008; Uesaka et al. 2008)*
596 were obtained from David Ginty (Harvard University, Boston, MA). Briefly, the floxed *Ret* allele (*Ret*^{fl}) was
597 generated by inserting a gene cassette, comprising the floxed human *RET9* cDNA with SV40 intron polyA
598 sequence followed by a cyan fluorescent protein (CFP) cDNA with polyA sequence, into exon 1 of the
599 mouse *Ret* gene. *Ret CFP* knock-in mice (*Ret*$^{CFP/+}$; MGI:3777556) were generated by crossing *Ret*$^{fl/+}$
600 mice to β-actin-Cre mice to remove the *RET9* cDNA This mouse was backcrossed >20 generation to
601 C57BL/6 to maintain the line and all experiments were conducted in this isogenic strain. All mouse
602 experiments were conducted in accordance with the National Institutes of Health Guidance for the Care
603 and Use of Laboratory Animals. All procedures were approved by the Johns Hopkins Animal Care and
604 Use Committee (Protocol number: MO21M56). All mice embryos used were either E12.5 or E14.5 and
605 both male and female mice were used.

606 ***Mouse breeding and genotyping:***

607 Homozygous *Ret* null mice are perinatal lethal (P0), therefore we intercrossed *Ret*$^{CFP/+}$ mice to
608 generate all possible genotypes. Embryos were genotyped from yolk sac genomic DNA to allow for
609 genotype- and sex-specific analysis. Routine mouse genotyping was performed by PCR with primer pairs
610 and amplicons as follows: wild-type *Ret* allele forward primer 5'-
611 CAGCGCAGGTCTCTCATCAGTACCGCA-3' and reverse primer 5'-
612 CAGCTAGCCGCAGCGACCCGGTTC-3' resulting in a 449 bp PCR product in *Ret*$^{+/+}$ and *Ret*$^{CFP/+}$
613 embryos (modified from (Jain and Varadarajan 2014)), CFP knock-in allele forward primer 5'
614 ATGGTGAGCAAGGGCGAGGAGCTGTT-3' and reverse primer 5'-
615 CTGGGTGCTCAGGTAGTGGTTGTC-3' resulting in a 615 bp PCR product from *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$
616 embryos. Embryo sex was determined by PCR using yolk sac genomic DNA with the forward primer 5'-
617 CTGAAGCTTTTGGCTTTGAG-3' and reverse primer 5'-CCGCTGCCAAATTCTTTGG-3,' mapping to

618 exons 9 and 10 of the *Kdm5c/d* genes resulting in two 331 bp X chromosome-specific amplicons in
619 females and an additional 301 bp Y chromosome-specific amplicon in males (Clapcote and Roder 2005).

620 ***Dissection, Dissociation, Flow Cytometry and Sorting.***

621 Whole gut tissue (stomach, fore and hind gut) from $Ret^{CFP/+}$ and $Ret^{CFP/CFP}$ male and female
622 embryos at E12.5 and E14.5 were dissociated into single-cell suspensions using Accumax (Sigma, USA).
623 The cells were filtered serially through a 100 µm and 40 µm cell strainer, and centrifuged at 2000 rpm for
624 5 m. The cell pellet was resuspended in 5% FBS, 4 mM EDTA in Leibovitz L-15 medium for cell sorting
625 using MoFlo XDP (BD Biosciences) with a blue-violet excitation laser for CFP. Single cells were sorted
626 directly into 96 well plates containing 1.9 µL Triton X-100 (Sigma-Aldrich, 0.2% v/v, #T9284), 1 µL dNTP
627 mix (Amresco, 10 mM each, #N557), 1 µL 5' biotinylated oligo-dT_{30}VN primer (IDT, 10 µM), and 0.1 µL
628 RNasin Plus (Promega, 40 U/µL, #N2615) per well. As a no-template negative control, one well per plate
629 did not receive a cell. Immediately after collection, sorted plates were centrifuged (2500 rpm for ~5 s),
630 frozen on dry ice, and stored at -80°C until library preparation.

631 ***Single-cell RNA-Seq library Preparation***

632 Libraries were prepared from 2 embryos for each of the 8 conditions ($Ret^{CFP/+}$ and $Ret^{CFP/CFP}$,
633 female and male, E12.5 and E14.5) with two exceptions: libraries were prepared from 1 $Ret^{CFP/CFP}$ male
634 E14.5 embryo and 4 $Ret^{CFP/CFP}$ female E14.5 embryos owing to differential viability. Single-cell RNA-seq
635 libraries were prepared following a modified Smart-seq2 protocol (Picelli et al. 2014; Chevée et al. 2018).
636 Reverse transcription and preamplification (20 cycles for E12.5 cells, 22 cycles for E14.5 cells) followed
637 the original Smart-seq2 protocol, with the exception of using 5' biotinylated NGG template-switching
638 oligonucleotides. Sequencing libraries were generated following the Nextera XT protocol (Illumina, #FC-
639 131-1096) on a quartered scale, using approximately 200 pg input DNA per sample and 14 enrichment
640 PCR cycles. Nextera XT v2 Index Sets (Illumina, #FC-131-2004) were diluted 1:5 in water prior to use.
641 Libraries were sequenced (paired end, 50 bp read length) on an Illumina HiSeq 2500 to an average depth
642 of 9.91×10^5 (standard deviation 7.43×10^5) paired-end reads per cell.

643 ***Data preprocessing***

644 Single-cell RNA-seq libraries were aligned with hisat2 version 2.0.1-beta to a mouse mm10 hisat2
645 index with the inserted CFP cDNA sequence included in the index. Using samtools 1.2, sam files were
646 compressed to bam files then sorted and indexed. Using Cufflinks v2.2.1 indexed bam files were
647 quantified against the Gencode mouse vM10 assembly and normalized across all 1,369 samples (1,351
648 cells, 18 negative control wells). 51 samples, both in the bottom 5th percentile for total number of reads
649 and for which less than 50% of reads were mapped, were removed. 22 PCA outliers were removed
650 iteratively. The remaining 1,296 samples (1,280 cells and 16 negative control wells) were normalized
651 again, as before. 261 cells for which less than 1,000 genes were detected at non-zero fragments per
652 kilobase of transcript per million mapped reads (FPKM) were removed, and 4 more cells were
653 subsequently removed as PCA outliers. 28 cells were then removed due to a high ratio of mitochondrial
654 to nuclear gene expression, indicating poor library quality. 1,003 cells and 0 negative control wells passed
655 these quality control measures. The age, sex, and genotype of these 1,003 cells is as follows: 183
656 embryonic day (E)12.5 female *Ret*$^{CFP/+}$, 164 E12.5 male *Ret*$^{CFP/+}$, 173 E12.5 female *Ret*$^{CFP/CFP}$, 125 E12.5
657 male *Ret*$^{CFP/CFP}$, 161 E14.5 female *Ret*$^{CFP/+}$, 156 E14.5 male *Ret*$^{CFP/+}$, 37 E14.5 female *Ret*$^{CFP/CFP}$, and 4
658 E14.5 male *Ret*$^{CFP/CFP}$. FPKM values were transformed to copies per cell (CPC) using the Census
659 algorithm (Qiu et al. 2017). Raw and processed data have been deposited in the NCBI Gene Expression
660 Omnibus (GEO) under accession number GSE192676.

661 ***scRNA-Seq analysis***

662 To reduce the computational burden, analyses of gene expression were limited to expressed
663 genes defined as genes with non-zero CPC values in at least 20 cells, and greater than $1x10^{-20}$ standard
664 deviation. Differential gene expression tests were performed using the Monocle R package (Trapnell et
665 al. 2014). To reduce noise and batch effects in visualizations a set of 419 genes with high dispersion was
666 defined (Chevée et al. 2018). To determine these genes, we normalized expression by read depth across
667 all expressed genes, then subset cells into five sets according to the pool in which each cell was

668 sequenced. Estimated dispersions were calculated for normalized expression of expressed genes in
669 each set. Overdispersed genes were defined as those with greater than 1.4 times the dispersion fit across
670 all five sets. Uniform manifold approximation and projection (UMAP) was performed on these 419 high-
671 dispersion genes using the Umapr R package (McInnes et al. 2018; [CSL STYLE ERROR: reference with
672 no printed form.]). RNAVelocity estimates were obtained using the Velocyto R package(La Manno et al.
673 2018). Cluster assignment was performed using the mClust package (Scrucca et al. 2016). Specificity
674 scores for each gene were calculated using the cummeRbund package (Goff et al. 2013). Briefly, Jensen-
675 Shannon (JS) distance was calculated between a vector of mean expression level for each cluster to a
676 unit vector of exclusive expression in each cluster, and specificity is defined as one minus the JS distance.

677 ***Pattern discovery with scCoGAPS***

678 The Coordinated Gene Activity in Pattern Sets (CoGAPS) algorithm employs non-negative matrix
679 factorization to determine patterns of gene expression across cells and returns an amplitude (A) matrix
680 of gene weights by pattern and a pattern (P) matrix of cell weights by pattern. The optimal number of
681 patterns to learn was determined by running several iterations and selecting the maximum number of
682 patterns before observing redundancies and noise. Loom files for publicly available scRNA-seq data from
683 the enteric nervous system of adolescent mice were downloaded for analysis (mousebrain.org). 70
684 patterns were learned on log-transformed unique molecular identifier (UMI) counts from 11,624 libraries
685 prepared through the 10X Genomics platform using single-cell CoGAPS (scCoGAPS)(Stein-O'Brien et
686 al. 2019). We projected our data into these patterns using projectR (Sharma et al. 2019; Stein-O'Brien et
687 al. 2019), which returned projection weights for each pattern in each cell and each gene. For our data,
688 25 patterns were learned on a log-transformed, normalized expression matrix of CPC limited to
689 expressed genes, defined above. We projected the publicly available data into these patterns learned on
690 our data again using projectR.

691 ***Random Forest Model***

692 While gene expression patterns were learned on the complete public dataset, which consists of 10,519
693 glia and 1,105 neurons, we trained the random forest model on all neurons and a subset of 1,660 glia to
694 create a 60:40 glia:neurons training dataset to avoid biasing the model to classifying glial cells. The glia
695 were subsampled such that the distribution of glial cells across clusters in the subset matched the
696 distribution across clusters in the whole glial population. The model was trained on 90% of cells to predict
697 cluster assignment from 70 pattern weights with 10-fold cross-validation resulting in an accuracy of
698 89.8%. To keep model input consistent, and because pattern weights are constrained to [0,1], projected
699 pattern weights were normalized to [0,1] before being used as input to predict cluster assignment. The
700 parent cell type of each cluster was taken as the predicted cell type for our data.

701 ***Immunofluorescence assay on whole mouse gut***

702 Whole gut tissue (stomach, fore and hind gut) at E12.5 and E14.5 was dissected and washed in ice-cold
703 PBS followed by fixation in 4% PFA at 4°C for 1 h. The fixed tissues were permeabilized and blocked in
704 5% normal donkey serum + 0.3% Triton-X for 1 h at room temperature (RT). The tissues were incubated
705 with appropriate primary antibodies at 4°C overnight, washed with PBS, then incubated in appropriate
706 secondary antibodies at 4°C overnight (see Table 12 for antibodies and dilutions). Tissues were washed
707 in PBS and and incubated with 0.2 µg/mL of DAPI (4′,6-diamidino-2-phenylindole) for 10 m at RT,
708 followed by incubation for 20 m at RT with CUBIC-1 (Susaki et al. 2014) diluted 1:10 with PBS. Samples
709 were washed with PBS and mounted on Fisherbrand Superfrost Plus microscope slides with ProLong
710 Gold Antifade mounting media and imaged using a Nikon Eclipse Ti confocal microscope.

711 ***RNAscope***

712 Whole gut tissue (stomach, fore and hind gut) at E12.5 was dissected and washed in ice-cold PBS
713 followed by fixation in 1% PFA at room temperature (RT) for several hours. Tissue was dehydrated
714 through an ethanol gradient and cleared with xylenes before embedding in paraffin. Tissue blocks were
715 sectioned on a microtome at 10 µm thickness and mounted on Fisherbrand Superfrost Plus microscope

716 slides. Slides were baked for 1 h at 60°C, allowed to cool to RT, then deparaffinized with xylenes.
717 Fluorescence in situ hybridization was performed following the manufacturer's protocol for the RNAscope
718 Fluorescent Multiplex V2 Assay using the RNAscope Multiplex Fluorescent Reagent Kit v2 (ACDBio
719 #323100). Briefly, deparaffinized sections were incubated in hydrogen peroxide for 10 m at RT and
720 incubated in 1X Target Retrieval for 15 m at RT, then incubated in Protease Plus for 30 m at 40°C.
721 Sections were incubated in a 50:1:1 mixture of RNAscope Probes CFP (ACDBio #529381), Mm-Cartpt-
722 C2 (ACDBio #432001-C2), and Mm-Gal-C3 (ACDBio #400961-C3) for 2 h at 40°C. Sections were
723 incubated with Amp 1, 2, and 3 for 30 m, 30 m, and 15 m at 40°C, respectively. Next, for each of the
724 three channels sections were incubated in the appropriate HRP for 15 m at 40°C, the appropriate
725 fluorophore for 30 m at 40°C, then HRP blocker for 15 m at 40°C. The fluorophores used were 1:750
726 dilutions of TSA Plus fluorescein (PerkinElmer #NEL741E001KT), cyanine 3 (PerkinElmer
727 #NEL744E001KT), and cyanine 5 (PerkinElmer #NEL745E001KT). Sections were then incubated in 1X
728 Trueblack Lipofuscin Autofluorescence Quencher (VWR #10119-144) for 1 m at RT to reduce red blood
729 cell autofluorescence. Sections were then incubated in RNAscope DAPI for 30 s at RT and mounted with
730 ProLong Gold Antifade mounting media and imaged using a Nikon Eclipse Ti confocal microscope.

731 ***Image acquisition***

732 Tissues were visualized using a Nikon Eclipse Ti confocal microscope and processed using Fiji.
733 Confocal stacks were taken using a 10X air objective (Figure 2D), 20X air objective (Figure S6B-C), 40X
734 oil objective with Type A immersion oil (Figure 5E and proliferation quantification), or 60X oil objective
735 with Type A immersion oil (Figure 3E-G, Figure S6D). For image presentation, stacks were z-projected
736 to single images using the maximum value for each pixel. For Figure S6B, two fields of view were stitched
737 together (Preibisch et al. 2009).

738 ***Proliferation assay and quantification***

739 To measure cell cycle occupancy we stained replicate *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ whole GI tracts from both
740 male and female mice at E14.5 for phosphohistone H3 (Anti-H3S10p; Millipore Sigma #06-570), an

741 immunomarker specific to cells undergoing mitosis (Skaland et al. 2009), and CFP protein (Anti-GFP;
742 Aves labs #GFP-1010) using the immunofluorescence assay described above. Confocal stack images
743 were acquired on the Nikon Eclipse Ti confocal microscope with a 40X oil objective and Type A immersion
744 oil and quantified using a custom CellProfiler (Carpenter et al. 2006; McQuin et al. 2018) workflow. Briefly,
745 primary objects were identified from each input channel (pH3, CFP) using the Otsu Adaptive threshold
746 strategy with a smoothing scale of 1.34888 and a threshold correction factor upper bound limit of 0.99.
747 The adaptive window size was set at 40. The typical diameter size ranged between 30-120 pixels
748 depending on whether nuclei (pH3) or cell bodies were being identified (CFP). Objects were related and
749 filtered using standard parameters and cells were quantified as either CFP^+, $pH3^+$, or $CFP^+/pH3^+$ double
750 positive. The proportion of dividing ENS precursors for each genotype was determined by estimating the
751 mean % of pH3+ cells within the subset of CFP+ cells for each animal. To directly test the two genotypes,
752 we performed Student's T-Test on the percentage of $pH3^+$ CFP^+ cells between *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$
753 aggregated to each replicate.

754 **Acknowledgements**

755 We would like to thank Hao Zhang at the Bloomberg School of Public Health, Johns Hopkins University
756 Cell Sorting facility for help with single cell sorting. E.V. is supported in part by the Johns Hopkins
757 Predoctoral Training Grant in Human Genetics (T32-GM007814-39). L.A.G. and E.V. are supported by
758 an award from the National Institute of Aging (R01AG066768). L.A.G. is additionally supported by the
759 National Institute of Aging (R01AG072305), the National Science Foundation (IOS-1665692), and a
760 Johns Hopkins Catalyst award. SC is supported by a NYU Grossman School of Medicine Startup Grant
761 (59-D-37802-30609-NYUPRJ) and AC is supported by a National Institutes of Health (Eunice Kennedy
762 Shriver National Institute of Child Health and Human Development) R01 award HD028088.

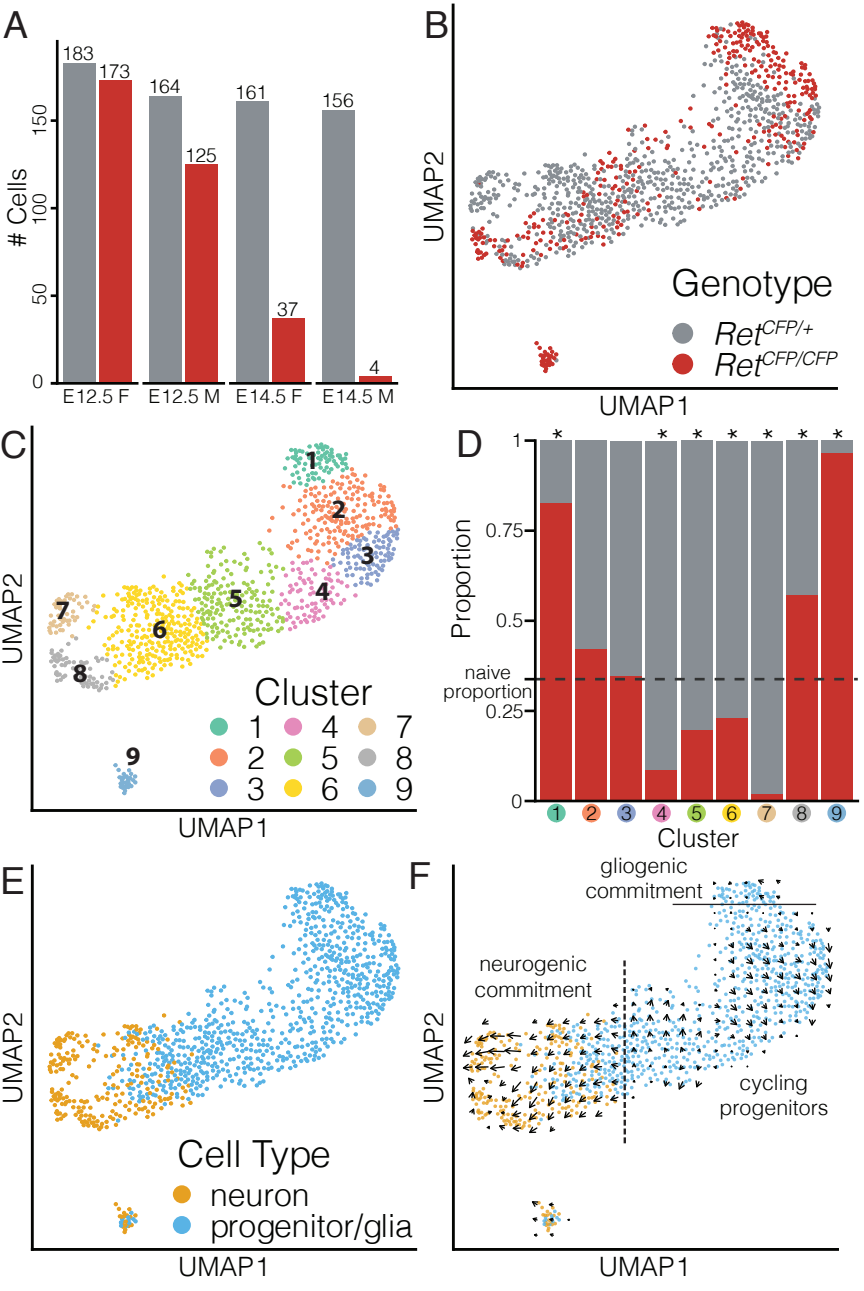
763 **Competing interests**

764 The authors declare no competing interests

Author Contributions

Elizabeth Vincent: Investigation, Data Analysis, Writing, Original draft preparation, Review and editing, Validation, Visualization; Sumantra Chatterjee: Conceptualization, Investigation, Writing, Original draft preparation, Review and editing, Supervision. Gabrielle Cannon: Investigation; Dallas Auer: Investigation; Holly Ross: Investigation; Aravinda Chakravarti: Conceptualization, Writing, Review and editing, Supervision, Funding acquisition; Loyal A. Goff: Conceptualization, Data Analysis, Writing, Original draft preparation, Review and editing, Supervision, Funding acquisition

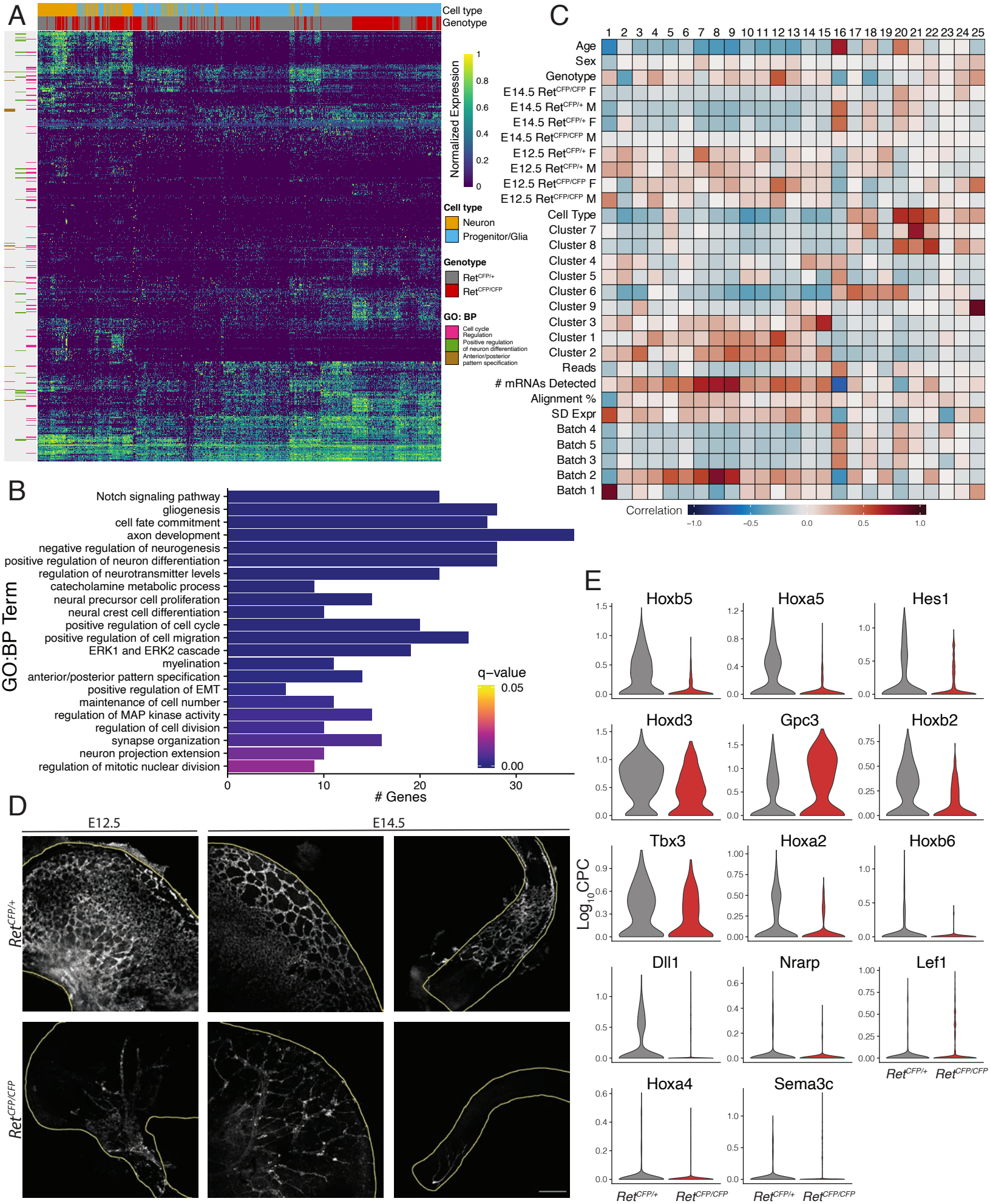
Figures


A
Cells
0
50
100
150
183
173
164
125
161
37
156
4
E12.5 F
E12.5 M
E14.5 F
E14.5 M
B
UMAP2
UMAP1
Genotype
Ret CFP/+
Ret CFP/CFP
C
UMAP2
UMAP1
1
2
3
4
5
6
7
8
9
Cluster
1 4 7
2 5 8
3 6 9
D
Proportion
1
0.75
0.5
naive proportion
0.25
0
Cluster
E
UMAP2
UMAP1
Cell Type
neuron
progenitor/glia
F
gliogenic commitment
neurogenic commitment
cycling progenitors
UMAP2
UMAP1

776 ***Figure 1: A cellular map of Ret expressing enteric nervous system cells.***

777 (A) Barplot of post quality control (QC) count of sorted *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ expressing cells from E12.5
778 & E14.5 embryonic gut in male and female mice. (B) Uniform Manifold Approximation and Projection
779 (UMAP) of 1003 cells colored by genotype. (C) UMAP embedding colored by 9 learned clusters (D)
780 Stacked barplot of the genotype proportions for each cluster. Asterisk denotes significant (Chi-squared
781 test for equal proportions, $p<0.05$ after multiple testing correction) deviation from expected naive
782 proportion. (E) UMAP embedding of *Ret*-expressing cells classified as glia/progenitor or neuron (F) RNA
783 velocity vector field along the UMAP embedding demonstrates direction of cellular differentiation. Dotted
784 line indicates neurogenic commitment, solid line indicates gliogenic commitment.

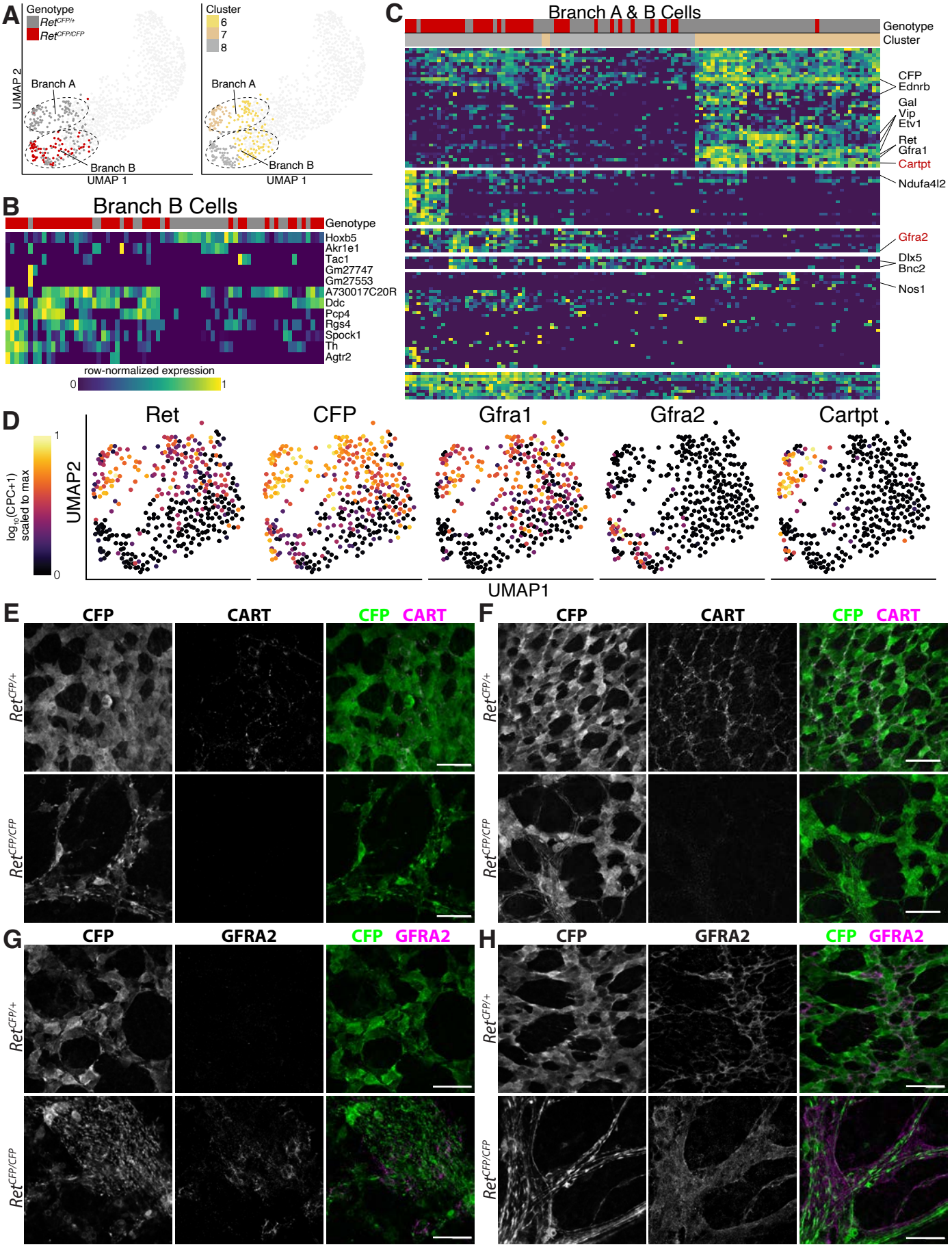
785


A
Cell type
Genotype
Normalized Expression
1
0.8
0.6
0.4
0.2
0
Cell type
Neuron
Progenitor/Glia
Genotype
Ret CFP/+
Ret CFP/CFP
GO: BP
Cell cycle Regulation
Positive regulation of neuron differentiation
Anterior/posterior pattern specification
B
GO:BP Term
Notch signaling pathway
gliogenesis
cell fate commitment
axon development
negative regulation of neurogenesis
positive regulation of neuron differentiation
regulation of neurotransmitter levels
catecholamine metabolic process
neural precursor cell proliferation
neural crest cell differentiation
positive regulation of cell cycle
positive regulation of cell migration
ERK1 and ERK2 cascade
myelination
anterior/posterior pattern specification
positive regulation of EMT
maintenance of cell number
regulation of MAP kinase activity
regulation of cell division
synapse organization
neuron projection extension
regulation of mitotic nuclear division
q−value
0.05
0.00
0
10
20
30
Genes
C
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Age
Sex
Genotype
E14.5 Ret CFP/CFP F
E14.5 Ret CFP/+ M
E14.5 Ret CFP/+ F
E14.5 Ret CFP/CFP M
E12.5 Ret CFP/+ F
E12.5 Ret CFP/+ M
E12.5 Ret CFP/CFP F
E12.5 Ret CFP/CFP M
Cell Type
Cluster 7
Cluster 8
Cluster 4
Cluster 5
Cluster 6
Cluster 9
Cluster 3
Cluster 1
Cluster 2
Reads
mRNAs Detected
Alignment %
SD Expr
Batch 4
Batch 5
Batch 3
Batch 2
Batch 1
Correlation
−1.0
−0.5
0.0
0.5
1.0
D
E12.5
E14.5
Ret CFP/+
Ret CFP/CFP
E
Hoxb5
Hoxa5
Hes1
Hoxd3
Gpc3
Hoxb2
Tbx3
Hoxa2
Hoxb6
Dll1
Nrarp
Lef1
Hoxa4
Sema3c
Log10CPC
Ret CFP/+
Ret CFP/CFP

786 ***Figure 2: Cellular and molecular changes driven by Ret loss of function***

787 (A) Heatmap of differentially expressed genes between *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ cells. (B) Significant
788 (hypergeometric test, $p<0.05$) gene ontology (GO) gene sets include neuron differentiation, gliogenesis
789 and cell fate commitment. C) Correlation plot of scCoGAPS patterns vs. annotated cellular features
790 including developmental age, sex and genotype. (D) Immunofluorescence staining of mouse GI tract at
791 E12.5 and E14.5 showing the organization of CFP-expressing cells. The first two columns show stomach
792 and the third column shows hindgut. 10X magnification, scale bar is 200 μm. (E) Violin plots of
793 significantly differentially expressed genes associated with the anterior-posterior patterning GO term.

794

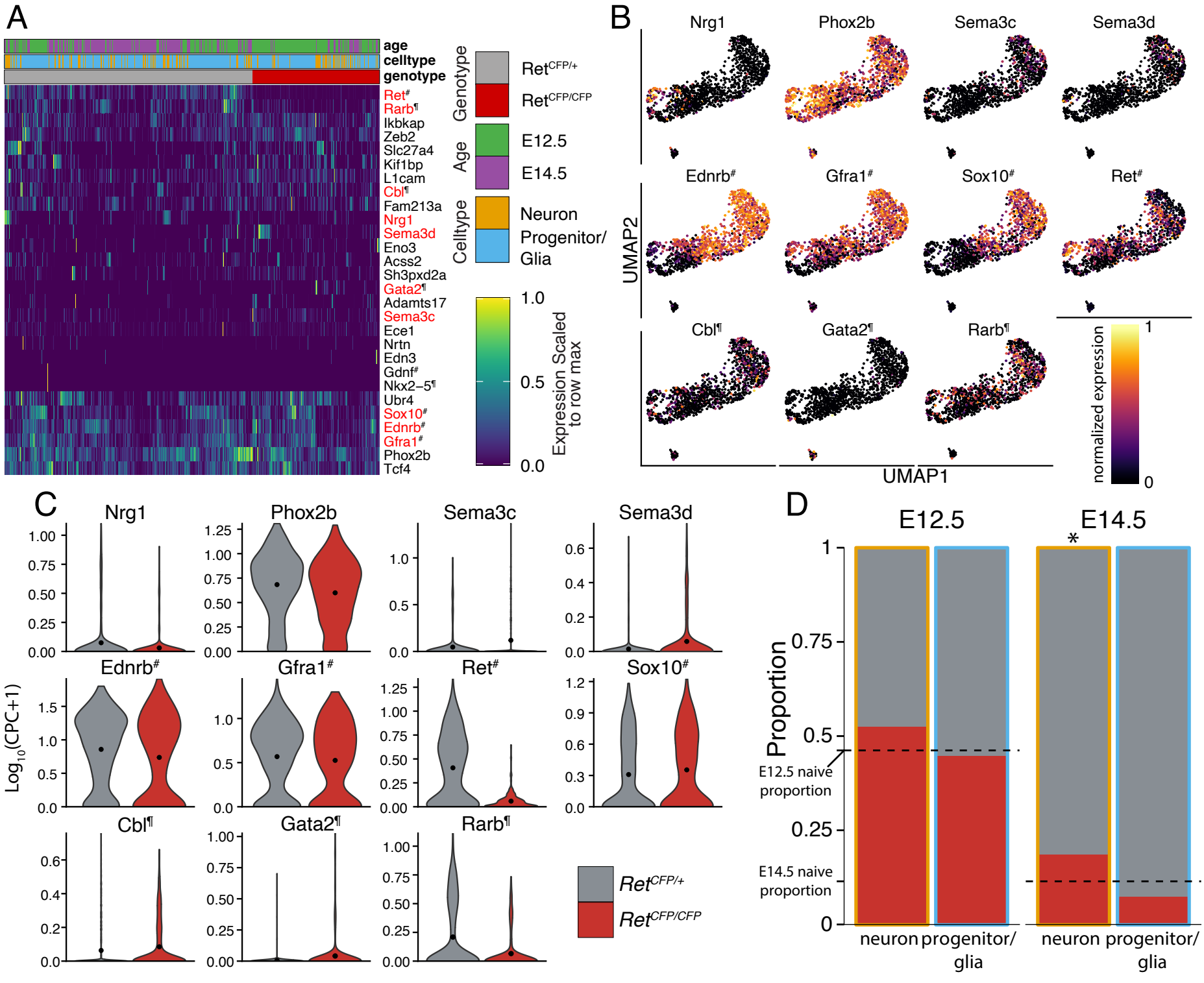

A
Genotype
Ret CFP/+
Ret CFP/CFP
Cluster
6
7
8
UMAP 2
Branch A
Branch B
UMAP 1
B
Branch B Cells
Genotype
Hoxb5
Akr1e1
Tac1
Gm27747
Gm27553
A730017C20R
Ddc
Pcp4
Rgs4
Spock1
Th
Agtr2
row-normalized expression
0
1
C
Branch A & B Cells
Genotype
Cluster
CFP
Ednrb
Gal
Vip
Etv1
Ret
Gfra1
Cartpt
Ndufa4l2
Gfra2
Dlx5
Bnc2
Nos1
D
Ret
CFP
Gfra1
Gfra2
Cartpt
log10(CPC+1) scaled to max
1
0
UMAP2
UMAP1
E
CFP
CART
CFP CART
Ret CFP/+
Ret CFP/CFP
F
CFP
CART
CFP CART
Ret CFP/+
Ret CFP/CFP
G
CFP
GFRA2
CFP GFRA2
Ret CFP/+
Ret CFP/CFP
H
CFP
GFRA2
CFP GFRA2
Ret CFP/+
Ret CFP/CFP

795 ***Figure 3: RET expression is required for access to the Branch A developmental trajectory***

796 (A) UMAP embedding of the neuronal cells comprising both developmental branches colored by
797 genotype (top) or by cluster assignment (bottom). Branch A (cluster 7) is almost entirely *Ret*$^{CFP/+}$ cells
798 while both genotypes are represented in branch B (cluster 8). (B) Heatmap of significant differentially
799 expressed genes with respect to genotype for cells in branch B. (C) Heatmap of significant differentially
800 expressed genes with respect to developmental lineage (branch) across all neurons from clusters 7 and
801 8. Row cluster segmentation represents the top six divisions of the hierarchical clustering of gene
802 expression values. (D) UMAP embeddings highlighting the persistent expression of *Ret, CFP,* and *Gfra1*
803 along the branch A developmental trajectory, the specific induction of the non-canonical Ret co-receptor
804 *Gfra2* along cells in branch B, and the specific expression of the branch A marker gene *Cartpt*. (E,F)
805 Immunofluorescence assay of branch A marker CART at E12.5 in *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ stomach (E)
806 and E14.5 in *Ret*$^{CFP/+}$ hindgut and *Ret*$^{CFP/CFP}$ stomach (F) demonstrates presence of CART positive cells
807 only in the *Ret*$^{CFP/+}$ mice. (G,H) Immunofluorescence assay of branch B marker GFRA2 at E12.5 in
808 *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ stomach (G) and E14.5 in *Ret*$^{CFP/+}$ hindgut and *Ret*$^{CFP/CFP}$ stomach (H). All IF assay
809 images are 60X magnification and scale bars are 40 µm.

810

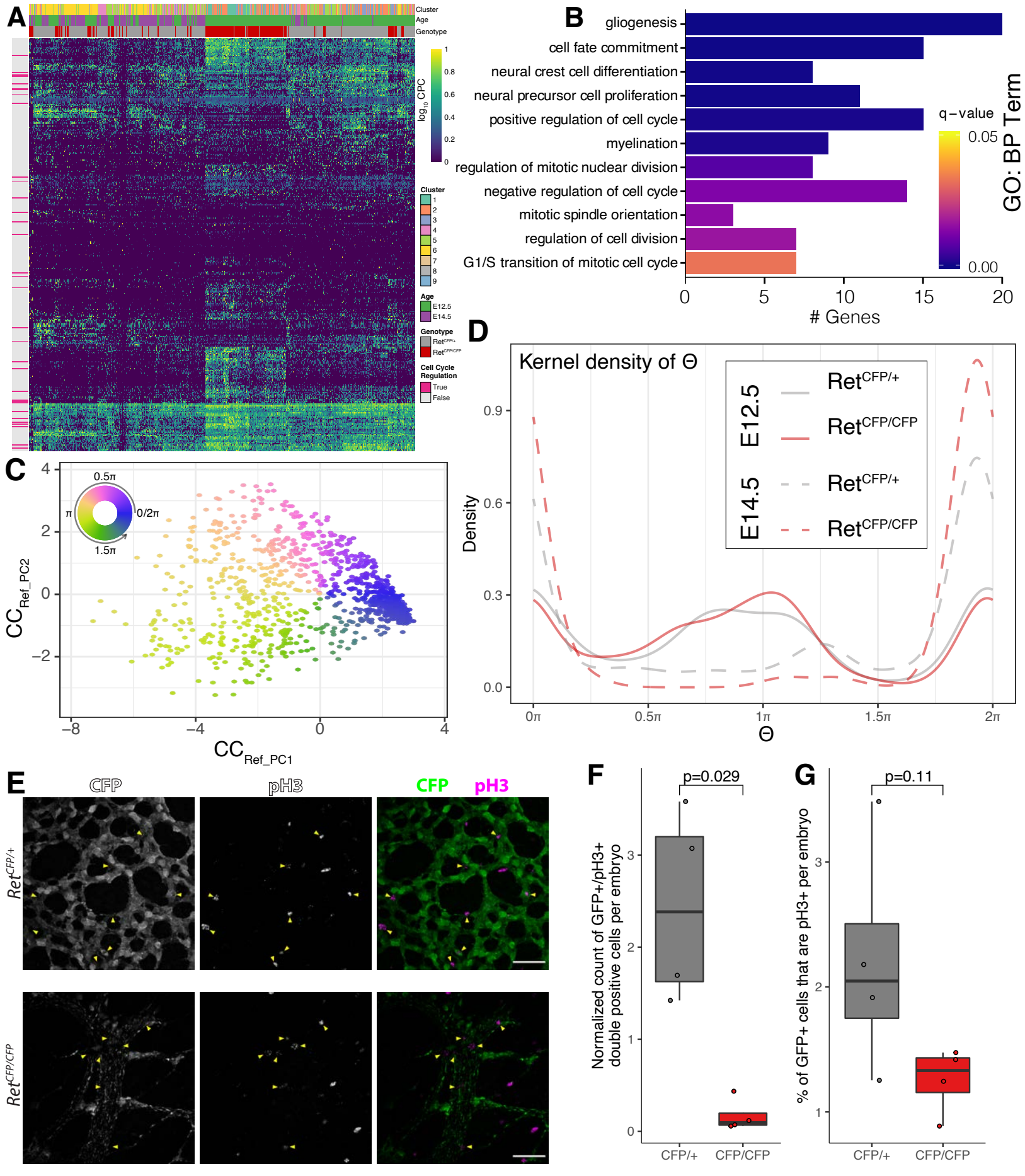
Vincent_Fig4



811 ***Figure 4: Cellular diversity of Hirschsprung disease (HSCR) genes***

812 (A) Heatmap of expression levels for 24 HSCR-associated genes across all isolated ENCDCs. Significant
813 differentially expressed genes are highlighted in red. (B) UMAP embeddings showing the expression of
814 11 genes that harbor pathogenic alleles in HSCR patients or are part of the *RET-EDNRB* GRN and are
815 differentially expressed (monocle likelihood ratio test $q<0.05$) with respect to genotype or with respect to
816 genotype-cell type interactions. Unannotated genes are HSCR-associated but not part of the *RET-*
817 *EDNRB* GRN, # indicates genes that are both associated with HSCR and are part of the *RET-EDNRB*
818 GRN, and ¶ indicates genes that are part of the *RET-EDNRB* GRN but have not been associated with
819 HSCR. (C) Violin plots of these 11 genes show expression with respect to genotype, with mean
820 expression value indicated by a dot. (D) Barplot showing the proportion of neurons and progenitors/glia
821 by genotype at E12.5 and E14.5. Dotted lines indicate the naive genotype proportion at E12.5 and E14.5.
822 Asterisk indicates a significant deviance from the naive proportion (Chi-squared test for equal proportions,
823 $p<0.05$).

824


A
Cluster
Age
Genotype
log10 CPC
Cluster
Age
E12.5
E14.5
Genotype
Ret CFP/+
Ret CFP/CFP
Cell Cycle Regulation
True
False
B
gliogenesis
cell fate commitment
neural crest cell differentiation
neural precursor cell proliferation
positive regulation of cell cycle
myelination
regulation of mitotic nuclear division
negative regulation of cell cycle
mitotic spindle orientation
regulation of cell division
G1/S transition of mitotic cell cycle
q−value
0.05
0.00
GO: BP Term
Genes
C
0.5π
0/2π
π
1.5π
CC Ref_PC2
CC Ref_PC1
D
Kernel density of Θ
E12.5
E14.5
Ret CFP/+
Ret CFP/CFP
Ret CFP/+
Ret CFP/CFP
Density
Θ
E
CFP
pH3
CFP pH3
Ret CFP/+
Ret CFP/CFP
F
p=0.029
Normalized count of GFP+/pH3+ double positive cells per embryo
CFP/+
CFP/CFP
G
p=0.11
% of GFP+ cells that are pH3+ per embryo
CFP/+
CFP/CFP

825 ***Figure 5:*** **Ret *LoF results in reduced cell cycle occupancy and significant reduction in the***

826 ***number of actively proliferating ENCDC progenitors.***

827 (A) Heatmap of significant differentially expressed genes between *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ within the non-

828 neuronal subset of isolated ENCDCs. (B) Barplots of significant biological process gene ontology terms

829 from gene ontology enrichment analysis of genes with significant differential expression between

830 genotypes in non-neuronal ENCDCs. (C) Tricycle embedding of ENCDCs to estimate cell cycle position.

831 (D) Kernel density estimates for cell cycle position (θ) for each genotype at each age. Proliferative phase

832 of the cell cycle (non-G0/G1) is between 0.25π and 1.75π. (E) Representative immunofluorescence

833 staining of the proliferation marker pH3 and CFP in the stomach of E12.5 *Ret*$^{CFP/+}$ and *Ret*$^{CFP/CFP}$ mice.

834 Yellow arrowheads indicate CFP^{+}/pH3^{+} cells. 40X magnification, scale bars are 20 μm. (F) Box &

835 whiskers plot showing the quantification of CFP^{+}/pH3^{+} double positive cells per embryo, normalized by

836 the total fields of view analyzed per embryo. (G) Box & whiskers plot showing the quantification of the %

837 of CFP^{+} cells with pH3^{+} staining in the GI tract of E14.5 *Ret*$^{CFP/+}$ or *Ret*$^{CFP/CFP}$ mice.

838

839

840 ## Supplemental Figures

841 ***Figure S1:*** Uniform Manifold Approximation and Projection (UMAP) embedding of 1,003 cells colored by
842 (A) age, (B) sex, and (C) sequencing batch.

843

844 ***Figure S2: Cluster-specific gene expression profiles***

845 (A) Scatter plot of mean gene expression per cluster vs. cluster specificity scores for each of the 9 learned
846 clusters of enriched ENCDCs. (B) UMAP embeddings highlighting the gene expression of top specific
847 cluster markers for each cluster. (C) UMAP embedding showing the expression levels for the ENS glial
848 marker *Ngfr* and the neuronal marker *Tubb3*.

849

850 ***Figure S3: Projection analysis of adult ENS cell types***

851 Each panel describes the usage of a single scCoGAPS pattern learned from (Zeisel et al. 2018). For
852 each of the 70 patterns, the left panel is a reduced dimensionality visualization of the glial cells and the
853 middle panel is for the neuronal cells from (Zeisel et al. 2018). Cells are colored based on their usage of
854 the learned pattern. The right panel is a UMAP embedding of the ENCDCs isolated in this study and cells
855 are colored by the corresponding projected pattern weight, highlighting the shared use of specific
856 biological features across both datasets.

857

858 ***Figure S4: Cell type annotation via transfer learning***

859 (A) Reduced dimensionality visualization of mature ENS cells from (Zeisel et al. 2018) colored by
860 previously annotated cluster designations. (B) UMAP embedding of sorted ENCDCs colored by the
861 random forest model predictions for corresponding clusters from Zeisel 2018. (C) UMAP embedding of
862 sorted ENCDCs colored by top-level ENS cell type designation. (D) Stacked barplot of cell type
863 proportions for each of the nine learned clusters of ENCDCs. (E) UMAP embedding of Zeisel 2018 mature
864 ENS cells by previously annotated cell type designations.

865

866 ***Figure S5: Learned gene expression patterns across all cells***

867 25 single-cell coordinated gene activity in pattern sets (scCoGAPS) patterns learned from sorted *Ret*-
868 expressing ENCDCs. Each UMAP embedding is colored by the usage of a specific pattern of gene co-
869 expression learned through non-negative matrix factorization.

870

871 ***Figure S6: RNAscope single molecule fluorescence in situ hybridization validation of branch A***
872 ***and branch B cells***

873 (A) UMAP embedding of ENCDC cells highlighting previously identified (Morarach et al. 2021) marker
874 gene expression for neuron developmental branches (branch A: *Nos1, Vip, Gal, Etv1*; branch B: *Bnc2,*
875 *Dlx5, Mgat4c, Ndufa4l2*). (B,C) RNAscope single molecule fluorescence *in situ* hybridization (smFISH)
876 validation of *Gal/Cartpt* in cross section of E12.5 stomach in both $Ret^{CFP/+}$ and $Ret^{CFP/CFP}$ mice. The yellow
877 box in (B) indicates the portion isolated in (C). 10X magnification, scale bars are 200 μm (B) and 50 μm
878 (C). (D) Whole mount immunofluorescence validation of branch B cells ($GFRA2^{+}/CFP^{+}$) in E12.5
879 esophagus of both $Ret^{CFP/+}$ and $Ret^{CFP/CFP}$ mice. 60X magnification, scale bars are 20 μm.

880

881 ***Figure S7: Expression profiles for Hirschsprung-associated and RET-EDNRB gene regulatory***
882 ***network genes***

883 (A) UMAP embeddings showing normalized expression of 21 expressed genes associated with
884 Hirschsprung disease (HSCR). (B) Violin plots showing expression of these same 21 genes by genotype.
885 (C) Violin plots showing expression of these same 21 genes by cluster. (D) UMAP embeddings showing
886 normalized expression of 3 genes in the *RET-EDNRB* GRN that have not been associated with
887 Hirschsprung disease. (E) Violin plots showing expression of these same 3 genes by genotype. (C) Violin
888 plots showing expression of these same 3 genes by cluster. Genes in red are significantly differentially
889 expressed (monocle likelihood ratio test $q<0.05$) either with respect to genotype or with respect to

890 genotype-cell type interactions. # indicates genes that are shared across the HSCR-associated and *RET-*
891 *EDNRB* GRN sets.
892
893 ***Figure S8: Cell cycle is modulated by Ret LoF***
894 (A) Heatmap of significant differentially expressed genes with respect to the combinatorial effect of both
895 genotype and celltype (i.e. genotype:celltype effect). (B) Barplot of significantly enriched gene ontology
896 terms for the list of genes identified as differentially expressed in A. Ontology coloring matches the row
897 color annotations for each gene in A. (C) UMAP embeddings showing the projected pattern weights for
898 the cell cycle stage-specific patterns (40-42) learned from (Zeisel et al. 2018). (D) UMAP embedding
899 colored by the maximum projection value across each projected cell cycle stage-specific pattern. (E)
900 Empirical cumulative density functions for projection values of cell cycle stage-specific patterns for each
901 of the interaction terms between age and genotype. Significant differences in distributions were
902 determined using the Kolmogorov-Smirnov test (* = $p<0.05$; ** = $p<0.01$ after multiple testing correction.
903 (F) Empirical cumulative density function (eCDF) curve across all projected (Zeisel et al. 2018)
904 scCoGAPS patterns. One curve for each of the interaction terms between age and genotype are
905 presented. (G-J) Box & whisker plots showing the distribution of quantifications for pH3 and CFP
906 immunofluorescence staining in E14.5 $Ret^{CFP/+}$ and $Ret^{CFP/CFP}$ mouse GI tracts. Each point represents a
907 separate sampled field of view from one of four replicate mice from each condition. Points are colored by
908 replicate embryos. G) Count of CFP^+ cells, H) Count of $pH3^+$ cells, I) Count of double positive $CFP^+/pH3^+$
909 cells, and J) Percent of CFP^+ cells that are $pH3^+$.
910

911 ## Tables

912 Table 1: Distribution of cells before and after QC
913 The age, sex, and *Ret* genotype of each embryo, as well as the number of libraries that were prepared,
914 sequenced, and passed all quality control filters.
915
916 Table 2: Gene specificity scores by cluster

Specificity scores for each gene (row) in each cluster (column). Specificity scores range from [0, 1], with a higher score indicating greater specificity.

Table 3: Gut atlas scCoGAPS A and P weights

Sheet “A-Pmatrix” contains the Pattern (P) matrix of pattern weight values for each cell (row) in each pattern (column) across 70 patterns learned on gut scRNAseq atlas data from Zeisel et al, 2018. Sheet “B-Amatrix” contains the Amplitude (A) matrix of pattern weight values for each gene (row) in each pattern (column).

Table 4: Gut atlas pattern markers

Sheet “A-patternMarkers” lists genes that are pattern markers for the 70 patterns in Table 3. Sheet “B-GO_terms” lists the corresponding gene ontology:biological process (GO:BP) terms for which the pattern marker lists are enriched. Sheet “C-GO_term_genes” lists the specific genes associated with each enriched GO:BP term. An entry of "not run" on sheet B indicates the pattern marker gene list for that particular pattern was not long enough (at least 10 genes) after converting to Entrez IDs, therefore gene ontology enrichment was not run on that list. An empty column under a pattern name on sheet B indicates no GO:BP terms reached significance (q<.05) for enrichment in pattern markers of that pattern.

Table 5: ENCDC data projected into gut atlas

Projected pattern weights (sheet “projection”) and p-values (sheet “pval”) for the 70 patterns (row) learned on the gut atlas into each cell (column) from the Ret mutant dataset.

Table 6: Differentially expressed gene lists

Each sheet presents results from a single differential expression test. At the top of each sheet we report which cells were tested (“cells”), the parameter tested (“test”), the full model formula (“full_model”), and the reduced model formula (“reduced_model”). Gene lists are subset to significantly differentially expressed (q < 0.05) genes.

Table 7: Differentially expressed gene GO lists

Each sheet presents results from a single gene ontology (GO) enrichment test, including the number of differentially expressed genes associated with each term (“Count”) and the official gene symbol for those genes (“geneID”) . The sheets in Table 7 report the GO enrichment results for the corresponding differentially expressed gene lists in Table 6. GO enrichment lists are subset to significantly enriched (q < 0.05) terms.

Table 8: ENCDC scCoGAPS A and P weights

Sheet “A-Pmatrix” contains the Pattern (P) matrix of pattern weight values for each cell (row) in each pattern (column) across 25 patterns learned on the Ret mutant dataset. Sheet “B-Amatrix” contains the Amplitude (A) matrix of pattern weight values for each gene (row) in each pattern (column).

Table 9: ENCDC patterns correlated with metadata

Correlation (sheet “A-correlation”) and the p-value for correlation (sheet “B-correlation-pvalue”) of various cell metadata parameters (row) with pattern weights for 25 patterns learned on the Ret mutant data (column).

Table 10: ENCDC data pattern markers

A list of marker genes each of the 25 patterns (column) learned on the Ret mutant data.

Table 11: Tricycle theta estimates

Theta estimates from tricycle (“tricyclePosition”) as well as relevant metadata for each cell (row).

Table 12: Antibodies and Concentrations Used

A list of all antibodies used in this study.

971 **References**

972 Alves MM, Sribudiani Y, Brouwer RWW, Amiel J, Antiñolo G, Borrego S, Ceccherini I, Chakravarti A,
973 Fernández RM, Garcia-Barcelo M-M, et al. 2013. Contribution of rare and common variants
974 determine complex diseases-Hirschsprung disease as a model. *Dev Biol* **382**: 320–329.
975 http://dx.doi.org/10.1016/j.ydbio.2013.05.019.

976 Amiel J, Sproat-Emison E, Garcia-Barcelo M, Lantieri F, Burzynski G, Borrego S, Pelet A, Arnold S,
977 Miao X, Griseri P, et al. 2008. Hirschsprung disease, associated syndromes and genetics: a
978 review. *J Med Genet* **45**: 1–14. http://jmg.bmj.com/content/45/1/1.short.

979 Bergeron K-F, Nguyen CMA, Cardinal T, Charrier B, Silversides DW, Pilon N. 2016. Upregulation of the
980 Nr2f1-A830082K12Rik gene pair in murine neural crest cells results in a complex phenotype
981 reminiscent of Waardenburg syndrome type 4. *Dis Model Mech* **9**: 1283–1293.
982 http://dx.doi.org/10.1242/dmm.026773.

983 Bogni S, Trainor P, Natarajan D, Krumlauf R, Pachnis V. 2008. Non-cell-autonomous effects of Ret
984 deletion in early enteric neurogenesis. *Development* **135**: 3007–3011.
985 http://dx.doi.org/10.1242/dev.025163.

986 Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH,
987 Lindquist RA, Moffat J, et al. 2006. CellProfiler: image analysis software for identifying and
988 quantifying cell phenotypes. *Genome Biol* **7**: R100. http://dx.doi.org/10.1186/gb-2006-7-10-r100.

989 Carrasquillo MM, McCallion AS, Puffenberger EG, Kashuk CS, Nouri N, Chakravarti A. 2002. Genome-
990 wide association study and mouse model identify interaction between RET and EDNRB pathways
991 in Hirschsprung disease. *Nat Genet* **32**: 237–244. http://dx.doi.org/10.1038/ng998.

992 Chatterjee S, Chakravarti A. 2019. A gene regulatory network explains RET-EDNRB epistasis in
993 Hirschsprung disease. *Hum Mol Genet* **28**: 3137–3147. http://dx.doi.org/10.1093/hmg/ddz149.

994 Chatterjee S, Kapoor A, Akiyama JA, Auer DR, Lee D, Gabriel S, Berrios C, Pennacchio LA,
995 Chakravarti A. 2016a. Enhancer Variants Synergistically Drive Dysfunction of a Gene Regulatory
996 Network In Hirschsprung Disease. *Cell* **167**: 355–368.e10.
997 http://dx.doi.org/10.1016/j.cell.2016.09.005.

998 Chatterjee S, Karasaki KM, Fries LE, Kapoor A, Chakravarti A. 2021. A multi-enhancer RET regulatory
999 code is disrupted in Hirschsprung disease. *Genome Res* **31**: 2199–2208.
000 http://genome.cshlp.org/cgi/pmidlookup?view=long&pmid=34782358 (Accessed July 19, 2021).

001 Chatterjee S, Nandakumar P, Auer DR, Gabriel SB, Chakravarti A. 2019. Gene- and tissue-level
002 interactions in normal gastrointestinal development and Hirschsprung disease. *Proceedings of the*
003 *National Academy of Sciences* **116**: 26697–26708. http://dx.doi.org/10.1073/pnas.1908756116.

004 Chatterjee S, Stine ZE, Mccallion AS, Chakravarti A. 2016b. RET Mutation and Function in HSCR,
005 MEN2, and Other Cancers. In *Epstein's Inborn Errors of Development* (eds. R.P. Erickson and A.J.
006 Wynshaw-Boris), pp. 517–523, Oxford University Press
007 http://oxfordmedicine.com/view/10.1093/med/9780199934522.001.0001/med-9780199934522-
008 chapter-69 (Accessed August 24, 2021).

009 Chevée M, Robertson JDJ, Cannon GH, Brown SP, Goff LA. 2018. Variation in Activity State, Axonal

010 Projection, and Position Define the Transcriptional Identity of Individual Neocortical Projection
011 Neurons. *Cell Rep* **22**: 441–455. http://dx.doi.org/10.1016/j.celrep.2017.12.046.

012 Clapcote SJ, Roder JC. 2005. Simplex PCR assay for sex determination in mice. *Biotechniques* **38**:
013 702, 704, 706. http://dx.doi.org/10.2144/05385BM05.

014 Dang R, Torigoe D, Suzuki S, Kikkawa Y, Moritoh K, Sasaki N, Agui T. 2011. Genetic Background
015 Strongly Modifies the Severity of Symptoms of Hirschsprung Disease, but Not Hearing Loss in
016 Rats Carrying Ednrbsl Mutations. *PLoS ONE* **6**: e24086.
017 http://dx.doi.org/10.1371/journal.pone.0024086.

018 Durbec PL, Larsson-Blomberg LB, Schuchardt A, Costantini F, Pachnis V. 1996. Common origin and
019 developmental dependence on c-ret of subsets of enteric and sympathetic neuroblasts.
020 *Development* **122**: 349–358. http://dx.doi.org/10.1242/dev.122.1.349.

021 Elmentaite R, Kumasaka N, Roberts K, Fleming A, Dann E, King HW, Kleshchevnikov V, Dabrowska
022 M, Pritchard S, Bolt L, et al. 2021. Cells of the human intestinal tract mapped across space and
023 time. *Nature* **597**: 250–255. http://dx.doi.org/10.1038/s41586-021-03852-1.

024 Fujiwara N, Nakazawa-Tanaka N, Miyahara K, Arikawa-Hirasawa E, Akazawa C, Yamataka A. 2018.
025 Altered expression of laminin alpha1 in aganglionic colon of endothelin receptor-B null mouse
026 model of Hirschsprung's disease. *Pediatr Surg Int* **34**: 137–141. http://dx.doi.org/10.1007/s00383-
027 017-4180-6.

028 Furness JB. 2012. The enteric nervous system and neurogastroenterology. *Nat Rev Gastroenterol*
029 *Hepatol* **9**: 286–294. http://dx.doi.org/10.1038/nrgastro.2012.32.

030 Furness JB. 2008. The enteric nervous system: normal functions and enteric neuropathies.
031 *Neurogastroenterol Motil* **20 Suppl 1**: 32–38. http://dx.doi.org/10.1111/j.1365-2982.2008.01094.x.

032 Gattelli A, Nalvarte I, Boulay A, Roloff TC, Schreiber M, Carragher N, Macleod KK, Schlederer M,
033 Lienhard S, Kenner L, et al. 2013. Ret inhibition decreases growth and metastatic potential of
034 estrogen receptor positive breast cancer cells. *EMBO Mol Med* **5**: 1335–1350.
035 http://dx.doi.org/10.1002/emmm.201302625.

036 Gershon MD. 2010. Developmental determinants of the independence and complexity of the enteric
037 nervous system. *Trends Neurosci* **33**: 446–456. http://dx.doi.org/10.1016/j.tins.2010.06.002.

038 Goff L, Trapnell C, Kelley D. 2013. cummeRbund: Analysis, exploration, manipulation, and visualization
039 of Cufflinks high-throughput sequencing data. *R package version* **2**.

040 Goldstein AM, Hofstra RMW, Burns AJ. 2013. Building a brain in the gut: development of the enteric
041 nervous system. *Clin Genet* **83**: 307–316. http://dx.doi.org/10.1111/cge.12054.

042 Gould TW, Yonemura S, Oppenheim RW, Ohmori S, Enomoto H. 2008. The neurotrophic effects of
043 glial cell line-derived neurotrophic factor on spinal motoneurons are restricted to fusimotor
044 subtypes. *J Neurosci* **28**: 2131–2146. http://dx.doi.org/10.1523/JNEUROSCI.5185-07.2008.

045 Grieco M, Santoro M, Berlingieri MT, Melillo RM, Donghi R, Bongarzone I, Pierotti MA, Della Porta G,
046 Fusco A, Vecchio G. 1990. PTC is a novel rearranged form of the ret proto-oncogene and is
047 frequently detected in vivo in human thyroid papillary carcinomas. *Cell* **60**: 557–563.
048 http://dx.doi.org/10.1016/0092-8674(90)90659-3.

049 Hao MM, Young HM. 2009. Development of enteric neuron diversity. *J Cell Mol Med* **13**: 1193–1210.
050 http://dx.doi.org/10.1111/j.1582-4934.2009.00813.x.

051 Heanue TA, Pachnis V. 2006. Expression profiling the developing mammalian enteric nervous system
052 identifies marker and candidate Hirschsprung disease genes. *Proceedings of the National*
053 *Academy of Sciences* **103**: 6919–6924. http://dx.doi.org/10.1073/pnas.0602152103.

054 Heuckeroth RO. 2018. Hirschsprung disease - integrating basic science and clinical medicine to
055 improve outcomes. *Nat Rev Gastroenterol Hepatol* **15**: 152–167.
056 http://dx.doi.org/10.1038/nrgastro.2017.149.

057 Heuckeroth RO, Enomoto H, Grider JR, Golden JP, Hanke JA, Jackman A, Molliver DC, Bardgett ME,
058 Snider WD, Johnson EM Jr, et al. 1999. Gene targeting reveals a critical role for neurturin in the
059 development and maintenance of enteric, sensory, and parasympathetic neurons. *Neuron* **22**:
060 253–263. http://dx.doi.org/10.1016/s0896-6273(00)81087-9.

061 Hirst CS, Stamp LA, Bergner AJ, Hao MM, Tran MX, Morgan JM, Dutschmann M, Allen AM, Paxinos G,
062 Furlong TM, et al. 2018. Publisher Correction: Kif1bp loss in mice leads to defects in the peripheral
063 and central nervous system and perinatal death. *Sci Rep* **8**: 9085.
064 http://dx.doi.org/10.1038/s41598-018-26661-5.

065 Jain PC, Varadarajan R. 2014. A rapid, efficient, and economical inverse polymerase chain reaction-
066 based method for generating a site saturation mutant library. *Anal Biochem* **449**: 90–98.
067 http://dx.doi.org/10.1016/j.ab.2013.12.002.

068 Kang Y-N, Fung C, Vanden Berghe P. 2021. Gut innervation and enteric nervous system development:
069 a spatial, temporal and molecular tour de force. *Development* **148**.
070 http://dx.doi.org/10.1242/dev.182543.

071 Kapoor A, Auer DR, Lee D, Chatterjee S, Chakravarti A. 2017. Testing the Ret and Sema3d genetic
072 interaction in mouse enteric nervous system development. *Hum Mol Genet* **26**: 1811–1820.
073 http://dx.doi.org/10.1093/hmg/ddx084.

074 Kapoor A, Jiang Q, Chatterjee S, Chakraborty P, Sosa MX, Berrios C, Chakravarti A. 2015. Population
075 variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1
076 susceptibility polymorphisms. *Hum Mol Genet* **24**: 2997–3003.
077 http://dx.doi.org/10.1093/hmg/ddv051.

078 Kapur RP, Yost C, Palmiter RD. 1992. A transgenic model for studying development of the enteric
079 nervous system in normal and aganglionic mice. *Development* **116**: 167–175.
080 http://dx.doi.org/10.1242/dev.116.Supplement.167.

081 Kulkarni S, Ganz J, Bayrer J, Becker L, Bogunovic M, Rao M. 2018. Advances in Enteric Neurobiology:
082 The “Brain” in the Gut in Health and Disease. *J Neurosci* **38**: 9346–9354.
083 http://dx.doi.org/10.1523/JNEUROSCI.1663-18.2018.

084 Kulkarni S, Saha M, Becker L, Wang Z, Liu G, Leser J, Kumar M, Bakhshi S, Anderson M, Lewandoski
085 M, et al. 2020. Neural crest-derived neurons are replaced by a newly identified mesodermal
086 lineage in the post-natal and aging enteric nervous system. *bioRxiv* 2020.08.25.262832.
087 https://www.biorxiv.org/content/10.1101/2020.08.25.262832v1 (Accessed July 19, 2021).

088 La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME,

089 Lönnerberg P, Furlan A, et al. 2018. RNA velocity of single cells. *Nature* **560**: 494–498.
090 http://dx.doi.org/10.1038/s41586-018-0414-6.

091 Lang D, Chen F, Milewski R, Li J, Lu MM, Epstein JA. 2000. Pax3 is required for enteric ganglia
092 formation and functions with Sox10 to modulate expression of c-ret. *J Clin Invest* **106**: 963–971.
093 http://dx.doi.org/10.1172/JCI10828.

094 Lasrado R, Boesmans W, Kleinjung J, Pin C, Bell D, Bhaw L, McCallum S, Zong H, Luo L, Clevers H,
095 et al. 2017. Lineage-dependent spatial and functional organization of the mammalian enteric
096 nervous system. *Science* **356**: 722–726. http://dx.doi.org/10.1126/science.aam7511.

097 Le Douarin NM, Smith J. 1988. Development of the peripheral nervous system from the neural crest.
098 *Annu Rev Cell Biol* **4**: 375–404. http://dx.doi.org/10.1146/annurev.cb.04.110188.002111.

099 Lui VCH, Samy ET, Sham MH, Mulligan LM, Tam PKH. 2002. Glial Cell Line–Derived Neurotrophic
100 Factor Family Receptors Are Abnormally Expressed in Aganglionic Bowel of a Subpopulation of
101 Patients with Hirschsprung's Disease. *Laboratory Investigation* **82**: 703–712.
102 http://dx.doi.org/10.1097/01.lab.0000017364.13014.ae.

103 McCallion AS, Stames E, Conlon RA, Chakravarti A. 2003. Phenotype variation in two-locus mouse
104 models of Hirschsprung disease: tissue-specific interaction between Ret and Ednrb. *Proc Natl*
105 *Acad Sci U S A* **100**: 1826–1831. http://dx.doi.org/10.1073/pnas.0337540100.

106 McInnes L, Healy J, Saul N, Großberger L. 2018. UMAP: Uniform Manifold Approximation and
107 Projection. *J Open Source Softw* **3**: 861. http://joss.theoj.org/papers/10.21105/joss.00861.

108 McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, Doan M, Ding L,
109 Rafelski SM, Thirstrup D, et al. 2018. CellProfiler 3.0: Next-generation image processing for
110 biology. *PLoS Biol* **16**: e2005970. http://dx.doi.org/10.1371/journal.pbio.2005970.

111 Morarach K, Memic F, Zeisel A, Hochgerner H, Marklund U, Linnarsson S. 2017. Single-cell molecular
112 interrogation of enteric nervous system development. *Mech Dev* S67.
113 https://www.infona.pl/resource/bwmeta1.element.elsevier-65aec147-c938-33e9-beeb-
114 d4eb732d66ef.

115 Morarach K, Mikhailova A, Knoflach V, Memic F, Kumar R, Li W, Ernfors P, Marklund U. 2021.
116 Diversification of molecularly defined myenteric neuron classes revealed by single-cell RNA
117 sequencing. *Nat Neurosci* **24**: 34–46. http://dx.doi.org/10.1038/s41593-020-00736-x.

118 Mulligan LM. 2014. RET revisited: expanding the oncogenic portfolio. *Nat Rev Cancer* **14**: 173–186.
119 http://dx.doi.org/10.1038/nrc3680.

120 Nagy N, Goldstein AM. 2017. Enteric nervous system development: A crest cell's journey from neural
121 tube to colon. *Semin Cell Dev Biol* **66**: 94–106. http://dx.doi.org/10.1016/j.semcdb.2017.01.006.

122 Natarajan D, Marcos-Gutierrez C, Pachnis V, de Graaff E. 2002. Requirement of signalling by receptor
123 tyrosine kinase RET for the directed migration of enteric nervous system progenitor cells during
124 mammalian embryogenesis. *Development* **129**: 5151–5160.
125 https://www.ncbi.nlm.nih.gov/pubmed/12399307.

126 Obermayr F, Hotta R, Enomoto H, Young HM. 2013. Development and developmental disorders of the
127 enteric nervous system. *Nat Rev Gastroenterol Hepatol* **10**: 43–57.

128 http://dx.doi.org/10.1038/nrgastro.2012.234.

129 Paratore C, Eichenberger C, Suter U, Sommer L. 2002. Sox10 haploinsufficiency affects maintenance
130 of progenitor cells in a mouse model of Hirschsprung disease. *Hum Mol Genet* **11**: 3075–3085.
131 http://dx.doi.org/10.1093/hmg/11.24.3075.

132 Park H-J, Bolton EC. 2017. RET-mediated glial cell line-derived neurotrophic factor signaling inhibits
133 mouse prostate development. *Development* **144**: 2282–2293.
134 http://dx.doi.org/10.1242/dev.145086.

135 Pattyn A, Morin X, Cremer H, Goridis C, Brunet JF. 1999. The homeobox gene Phox2b is essential for
136 the development of autonomic neural crest derivatives. *Nature* **399**: 366–370.
137 http://dx.doi.org/10.1038/20700.

138 Perea D, Guiu J, Hudry B, Konstantinidou C, Milona A, Hadjieconomou D, Carroll T, Hoyer N,
139 Natarajan D, Kallijärvi J, et al. 2017. Ret receptor tyrosine kinase sustains proliferation and tissue
140 maturation in intestinal epithelia. *EMBO J* **36**: 3029–3045.
141 http://dx.doi.org/10.15252/embj.201696247.

142 Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. 2014. Full-length RNA-seq
143 from single cells using Smart-seq2. *Nat Protoc* **9**: 171–181.
144 http://dx.doi.org/10.1038/nprot.2014.006.

145 Preibisch S, Saalfeld S, Tomancak P. 2009. Globally optimal stitching of tiled 3D microscopic image
146 acquisitions. *Bioinformatics* **25**: 1463–1465. http://dx.doi.org/10.1093/bioinformatics/btp184.

147 Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C. 2017. Single-cell mRNA quantification and
148 differential analysis with Census. *Nat Methods* **14**: 309–315. http://dx.doi.org/10.1038/nmeth.4150.

149 Rao M, Gershon MD. 2016. The bowel and beyond: the enteric nervous system in neurological
150 disorders. *Nat Rev Gastroenterol Hepatol* **13**: 517–528.
151 http://dx.doi.org/10.1038/nrgastro.2016.107.

152 Rossi J, Luukko K, Poteryaev D, Laurikainen A, Sun YF, Laakso T, Eerikäinen S, Tuominen R, Lakso
153 M, Rauvala H, et al. 1999. Retarded Growth and Deficits in the Enteric and Parasympathetic
154 Nervous System in Mice Lacking GFRα2, a Functional Neurturin Receptor. *Neuron* **22**: 243–252.
155 https://www.sciencedirect.com/science/article/pii/S0896627300810867.

156 Schuchardt A, D'Agati V, Larsson-Blomberg L, Costantini F, Pachnis V. 1994. Defects in the kidney and
157 enteric nervous system of mice lacking the tyrosine kinase receptor Ret. *Nature* **367**: 380–383.
158 http://dx.doi.org/10.1038/367380a0.

159 Scrucca L, Fop M, Murphy TB, Raftery AE. 2016. *mclust 5: Clustering, Classification and density*
160 *estimation using Gaussian finite mixture models. R J. 8: 289--317. doi: 10.32614*. RJ-2016-021.

161 Sharma G, Colantuoni C, Goff LA, Fertig EJ, Stein-O'Brien G. 2019. projectR: An R/Bioconductor
162 package for transfer learning via PCA, NMF, correlation, and clustering. *bioRxiv* 726547.
163 https://www.biorxiv.org/content/10.1101/726547v1.abstract (Accessed August 26, 2019).

164 Skaland I, Janssen EAM, Gudlaugsson E, Hui Ru Guo L, Baak J. 2009. The prognostic value of the
165 proliferation marker phosphohistone H3 (PPH3) in luminal, basal-like and triple negative phenotype
166 invasive lymph node-negative breast cancer. *Anal Cell Pathol* **31**: 261–271.

167 https://content.iospress.com/articles/analytical-cellular-pathology/clo464.

168 Southard-Smith EM, Kos L, Pavan WJ. 1998. Sox10 mutation disrupts neural crest development in
169 Dom Hirschsprung mouse model. *Nat Genet* **18**: 60–64. http://dx.doi.org/10.1038/ng0198-60.

170 Stein-O'Brien GL, Carey JL, Lee WS, Considine M, Favorov AV, Flam E, Guo T, Li S, Marchionni L,
171 Sherman T, et al. 2017. PatternMarkers & GWCoGAPS for novel data-driven biomarkers via whole
172 transcriptome NMF. *Bioinformatics* **33**: 1892–1894. http://dx.doi.org/10.1093/bioinformatics/btx058.

173 Stein-O'Brien GL, Clark BS, Sherman T, Zibetti C, Hu Q, Sealfon R, Liu S, Qian J, Colantuoni C,
174 Blackshaw S, et al. 2019. Decomposing Cell Identity for Transfer Learning across Cellular
175 Measurements, Platforms, Tissues, and Species. *Cell Syst* **8**: 395–411.e8.
176 http://dx.doi.org/10.1016/j.cels.2019.04.004.

177 Susaki EA, Tainaka K, Perrin D, Kishino F, Tawara T, Watanabe TM, Yokoyama C, Onoe H, Eguchi M,
178 Yamaguchi S, et al. 2014. Whole-brain imaging with single-cell resolution using chemical cocktails
179 and computational analysis. *Cell* **157**: 726–739. http://dx.doi.org/10.1016/j.cell.2014.03.042.

180 Tani G, Tomuschat C, O'Donnell AM, Coyle D, Puri P. 2017. Increased population of immature enteric
181 glial cells in the resected proximal ganglionic bowel of Hirschsprung's disease patients. *J Surg Res*
182 **218**: 150–155. http://dx.doi.org/10.1016/j.jss.2017.05.062.

183 Tilghman JM, Ling AY, Turner TN, Sosa MX, Krumm N, Chatterjee S, Kapoor A, Coe BP, Nguyen K-
184 DH, Gupta N, et al. 2019. Molecular Genetic Anatomy and Risk Profile of Hirschsprung's Disease.
185 *N Engl J Med* **380**: 1421–1432. http://dx.doi.org/10.1056/NEJMoa1706594.

186 Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS,
187 Rinn JL. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal
188 ordering of single cells. *Nat Biotechnol* **32**: 381–386. http://dx.doi.org/10.1038/nbt.2859.

189 Uesaka T, Nagashimada M, Enomoto H. 2015. Neuronal Differentiation in Schwann Cell Lineage
190 Underlies Postnatal Neurogenesis in the Enteric Nervous System. *J Neurosci* **35**: 9879–9888.
191 http://dx.doi.org/10.1523/JNEUROSCI.1239-15.2015.

192 Uesaka T, Nagashimada M, Yonemura S, Enomoto H. 2008. Diminished Ret expression compromises
193 neuronal survival in the colon and causes intestinal aganglionosis in mice. *J Clin Invest* **118**: 1890–
194 1898. http://dx.doi.org/10.1172/JCI34425.

195 Young HM, Ciampoli D, Hsuan J, Canty AJ. 1999. Expression of Ret-, p75(NTR)-, Phox2a-, Phox2b-,
196 and tyrosine hydroxylase-immunoreactivity by undifferentiated neural crest-derived cells and
197 different classes of enteric neurons in the embryonic mouse gut. *Dev Dyn* **216**: 137–152.
198 https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0177(199910)216:2%3C137::AID-
199 DVDY5%3E3.0.CO;2-6.

200 Young HM, Hearn CJ, Ciampoli D, Southwell BR, Brunet JF, Newgreen DF. 1998. A single rostrocaudal
201 colonization of the rodent intestine by enteric neuron precursors is revealed by the expression of
202 Phox2b, Ret, and p75 and by explants grown under the kidney capsule or in organ culture. *Dev*
203 *Biol* **202**: 67–84. http://dx.doi.org/10.1006/dbio.1998.8987.

204 Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Zwan J, Häring M, Braun E,
205 Borm LE, La Manno G, et al. 2018. Molecular Architecture of the Mouse Nervous System. *Cell*
206 **174**: 999–1014.e22. http://dx.doi.org/10.1016/j.cell.2018.06.021.

Zheng S, Stein-O'Brien G, Augustin J, Slosberg J, Carosso GA, Winer B, Shin G, Bjornsson HT, Goff LA, Hansen KD. 2021. Universal prediction of cell cycle position using transfer learning. *bioRxiv*. https://www.biorxiv.org/content/10.1101/2021.04.06.438463v2.abstract.

umapr: ARCHIVED UMAP dimensionality reduction in R. Github https://github.com/ropensci-archive/umapr (Accessed October 22, 2021).