Cell Reports, Volume *25*

Supplemental Information

Single-Cell Transcriptomics Characterizes Cell Types in the Subventricular Zone and Uncovers Molecular Defects Impairing Adult Neurogenesis

Vera Zywitza, Aristotelis Misios, Lena Bunatyan, Thomas E. Willnow, and Nikolaus Rajewsky

A

exp_ID	fixation	sex	# animals	age (weeks)	# cells	median UMIs/cell	median genes/cell
an002	MeOH	f	5	16.7	1788	1141	713.5
an003F	MeOH	f	5	13.7	855	1410	859
an003L	-	f	5	13.9	2362	1720	1024
an008	-	m	4	13.1	2994	924.5	625
an009	-	mix	2 f, 2 m	13.4	1805	938	624

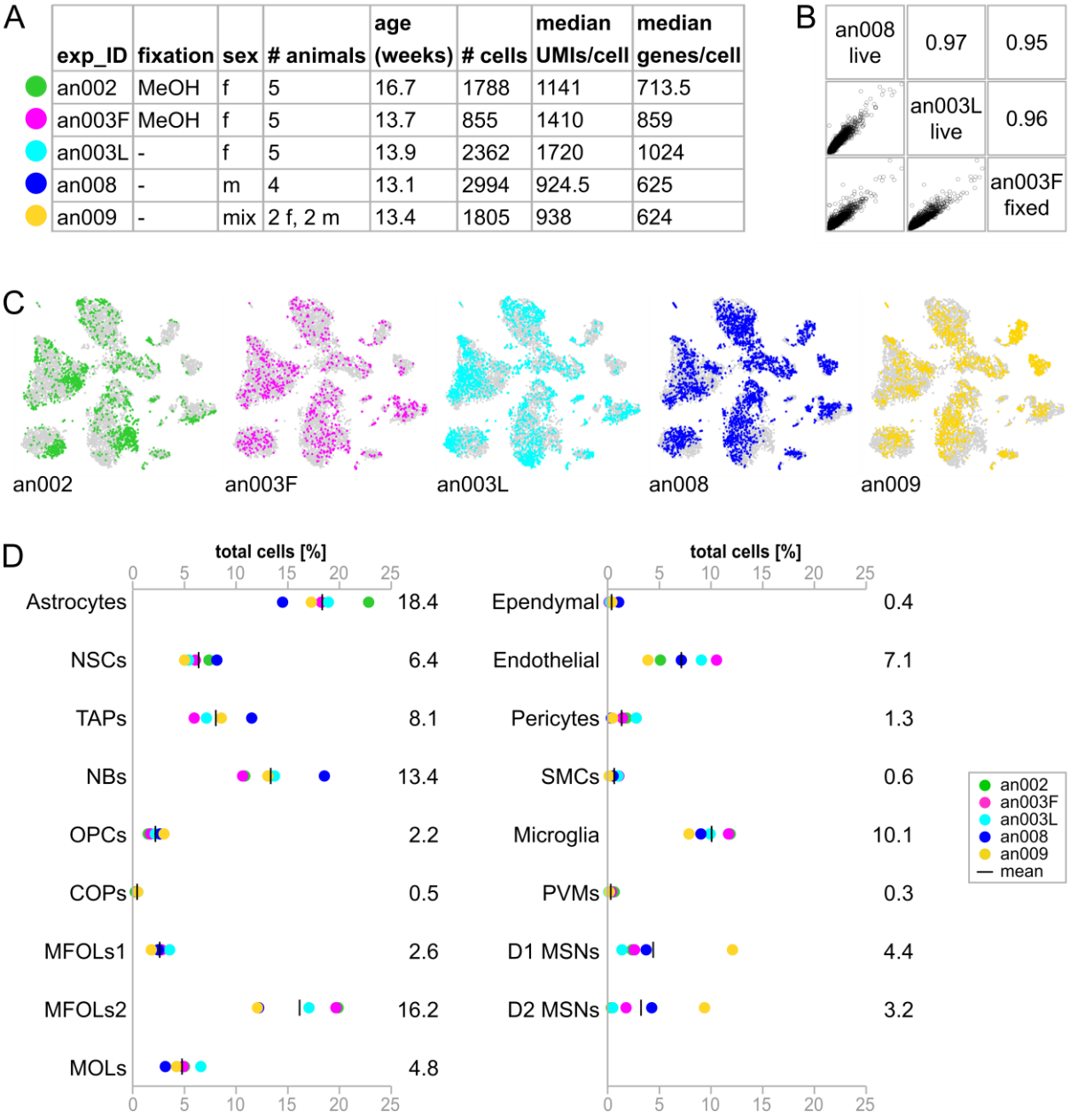


Figure S1. Cluster composition and distribution of replicates across cell types. Related to Figure 2.

(A) Overview of replicates.

(B) Gene expression levels from live and fixed cells correlate well. Upper right panels show Pearson correlation. Lower left panels show pairwise correlation of genes (averaged normalized UMI counts).

(C) Distribution of replicates in the all-cell tSNE plot. Cells from one replicate are colored and plotted on top of all other cells (in gray).

(D) Proportions of cell types. Numbers represent the mean.

Abbreviations: exp_ID: experiment identifier; MeOH: methanol; f: female; m: male; UMIs: unique molecular identifiers. Cell type abbreviations as in Figure 2.

Figure S2. Marker gene expression. Related to Figure 2.

tSNE plots of 9,804 cells colored by the expression of certain marker genes (yellow indicates low, red indicates high expression): *Gja1*: astrocytes and NSCs; *Egfr*: activated NSCs and neural progeny; *Pcna*: proliferating cells; *Foxj1, Ak7*: ependymal cells; *Dlx1, Dlx2*: committed neural progenitors; *Tubb3*: neuroblasts; *Rbfox3, Snap25*: mature neurons; *Tac1*: D1 MSNs; *Penk*: D2 MSNs; *Cx3cr1*: microglia; *C1qa*: immune cells (microglia and PVMs);

Lyz2: PVMs; *Acta2*: SMCs and ependymal cells; *Vtn*: pericytes; *Cspg4*: pericytes and OPCs; *Pdgfra, Olig2*: OPCs; *Olig1*: oligodendrocytes; *Fyn, Bmp4*: COPs; *Opalin, Mal*: MFOLs; *Klk6*: MOLs. *Bfsp2* is a putative new marker gene to discriminate MFOLs1 from MFOLs2. *Bfsp2* was highest expressed in MFOLs1, but also detected in COPs (green circle) indicating that MFOLs1 represent an earlier stage of myelin-forming oligodendrocytes than MFOLs2. Classification of oligodendrocyte clusters was based mainly on genes published in Marques et al. (2016). Squares are close ups of certain cell clusters. Red circles enclose putative ChAT neurons. At the bottom right: tSNE colored by cluster annotation.
Cell type abbreviations as in Figure 2.

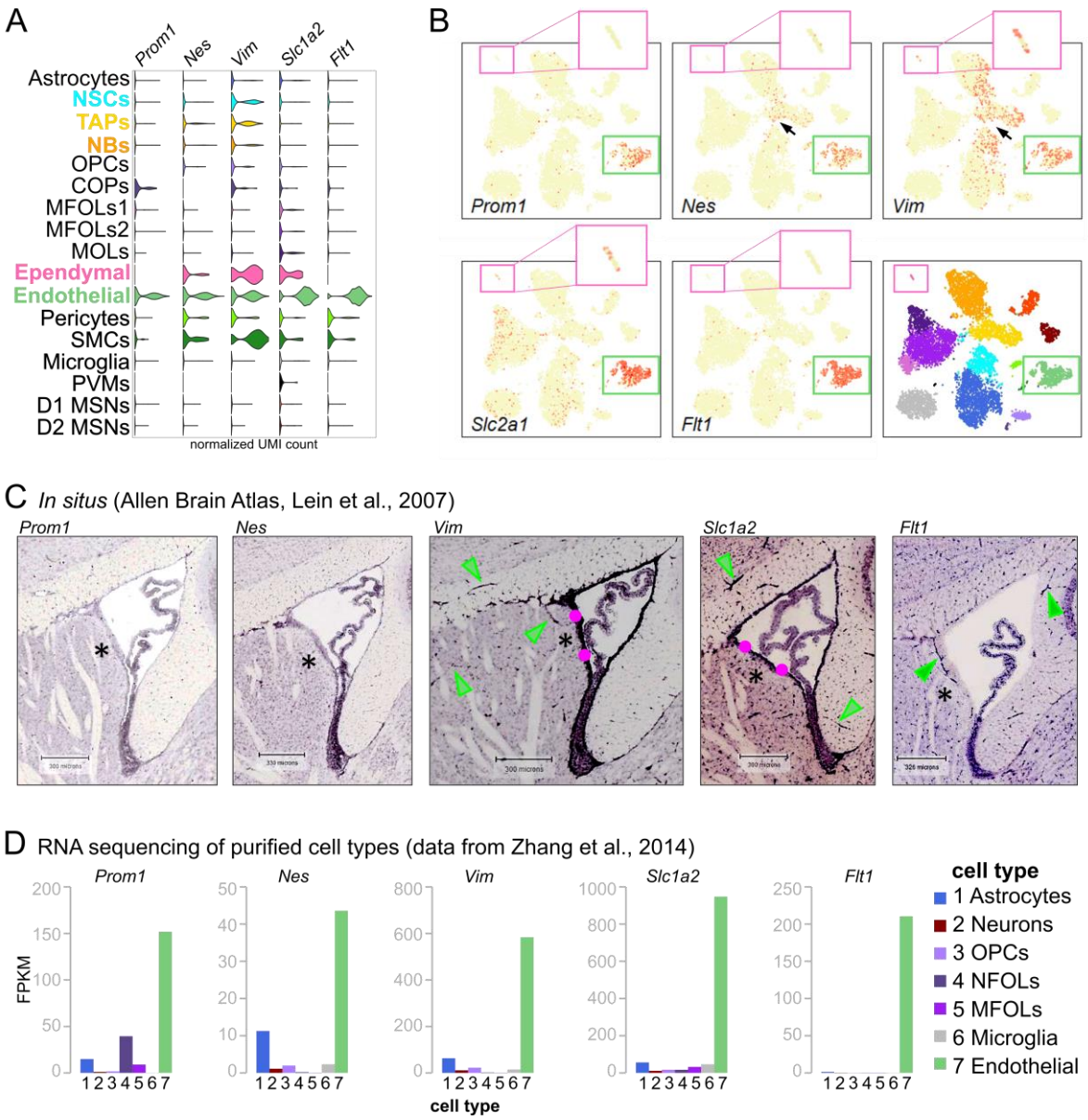


Figure S3. Characterization of endothelial cells. Related to Figure 2.

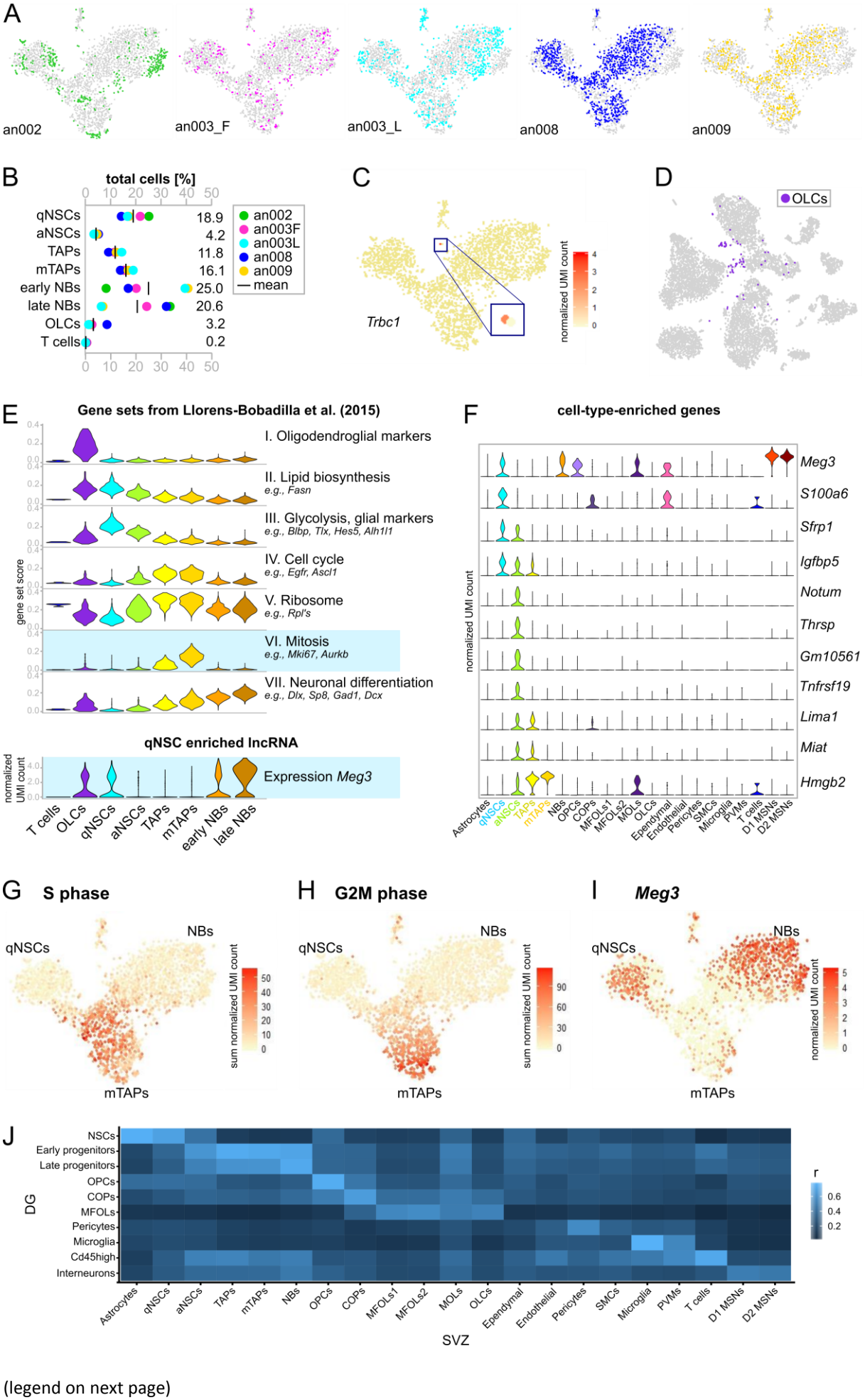
(A, B) Expression of genes associated with ependymal, endothelial, neural stem and progenitor cells.

(B) Pink boxes highlight and zoom into the ependymal cell cluster. Endothelial cells are highlighted by green boxes. Arrows point at neural progenitors. Yellow indicates low, red indicates high expression. At the bottom right: tSNE colored by cluster annotation.

(C) *In situ* images from the Allen Brain Atlas (Lein et al., 2007) depicting the expression of genes from A and B in the adult mouse brain. *Prom1* and *Nes* mRNA was barely detected. Green arrowheads point to blood vessels. Pink circles highlight staining of the ependymal cell layer. Asterisks label the lateral wall.

(D) Expression of genes from A-C in published bulk RNA sequencing data from sorted cell populations (data from Zhang et al., 2014).

Abbreviations as in Figure 2.



(legend on next page)

Figure S4. Subcluster characterization, NSC activation-state enriched genes, and comparison with DG. Related to Figures 2 and 3.

(A) Distribution of replicates in the tSNE plot of the subclustering. Cells from one replicate are colored and plotted on top of all other cells (in gray).

(B) Proportions of cell types after subclustering. Numbers represent the mean.

(C) tSNE plot of the subclustering colored by the expression of T cell receptor *Trbc1*. The expression of T cell receptors was specific for subcluster 1 (blue square represents a close up) disclosing these cells as contaminating T cells. Of note, the subcluster consists of four cells deriving from four different replicates.

(D) Localization of oligodendrocyte-like cells (OLCs, in purple) in the all-cell tSNE plot. In the subclustering analysis, subcluster 2 separated due to oligodendrocyte associated genes (gene set I, panel E). However, OLCs also expressed some glial and neuroblast genes (gene sets II, III and VII, panel E) and localized close to NSCs, TAPs, and NBs in the all-cell tSNE plot (panel D) suggesting that OLCs might represent cell doublets. As two previous single-cell RNA sequencing studies also observed OLCs within their FACS-sorted populations (Dulken et al., 2017; Llorens-Bobadilla et al., 2015) we doubt that these cells are doublets, but hypothesize that they might represent a so far uncharacterized cell type.

(E) Characterization of subclusters based on gene sets published in Llorens-Bobadilla et al. (2015). At the bottom: the expression of the long non-coding RNA (lncRNA) *Meg3* negatively correlates with the expression of mitotic genes (blue background). See also panels G-I.

(F) NSC activation-state enriched genes in the context of all identified cell types from Figures 2A and 3C.

(G, H) Expression sum of S (G) and G2M (H) phase associated genes (from Tirosh et al., 2016) in the tSNE plot of the subclustering.

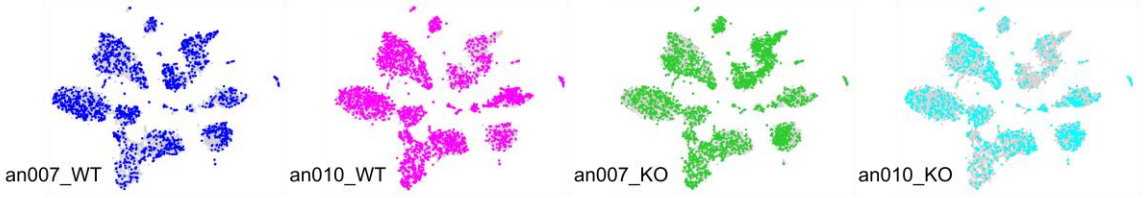
(I) Expression of the lncRNA *Meg3* in the tSNE plot of the subclustering.

(J) Gene expression profiles of SVZ and DG cell types correlate well. The heatmap depicts Pearson correlation (r, bright indicates high, dark indicates low correlation). Cell types on the y axis (DG) were characterized in Artegiani et al. (2017).

Cell type abbreviations as in Figures 2 and 3.

A

exp_ID	genotype	fixation	sex	# animals	age (weeks)	# cells	median UMIs/cell	median genes/cell
an007_WT	wildtype and heterozygous	-	m	4	11.3	2604	943	660
an010_WT	littermate controls (Ctr)	-	m	4	11.4	1713	1118	725
an007_KO	LRP2-deficient (KO)	-	m	3	11.3	2698	940	668
an010_KO		-	m	4	11.4	1041	1166	754



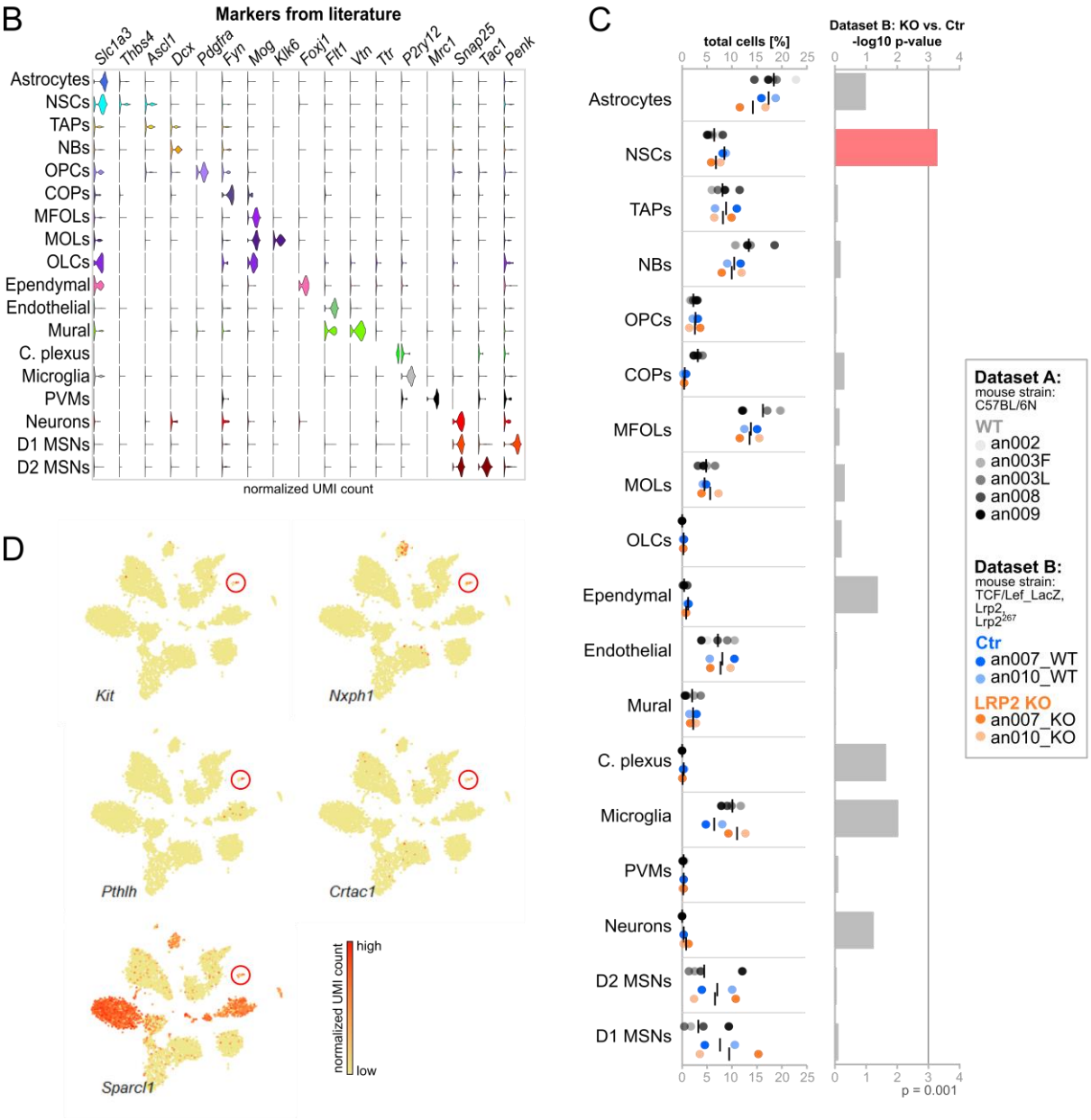


Figure S5. Cluster composition, cell type identification and distribution of replicates across cell types in the LRP2 KO/Ctr Drop-seq analysis (dataset B). Related to Figure 4. (legend on next page)

Figure S5. Cluster composition, cell type identification and distribution of replicates across cell types in the LRP2 KO/Ctr Drop-seq analysis (dataset B). Related to Figure 4.

(A) Overview of replicates and distribution of replicates in the all-cell tSNE plot. Cells from one replicate are colored and plotted on top of all other cells (in gray). an007_WT contained cells derived from one wild-type and three heterozygous mice. For an010_WT, two wild-type and two heterozygous mice were used.

(B) Identification of cell types based on known marker genes.

(C) Proportions of cell types in dataset A (gray, wild-type) and dataset B (blue, control; orange, LRP2-deficient mice). Right panel: P-values from the comparison of cell type proportions in dataset B: Ctr vs. KO. Statistical significance of data was determined using a generalized linear mixed model. The p-value for NSCs is highlighted in red.

Abbreviations: exp_ID: experiment identifier; m = male; KO: LRP2-deficient mice, Ctr: wild-type and heterozygous littermate controls; C. plexus: choroid plexus, OLCs: oligodendrocyte-like cells, other cell type abbreviations as in Figure 2.

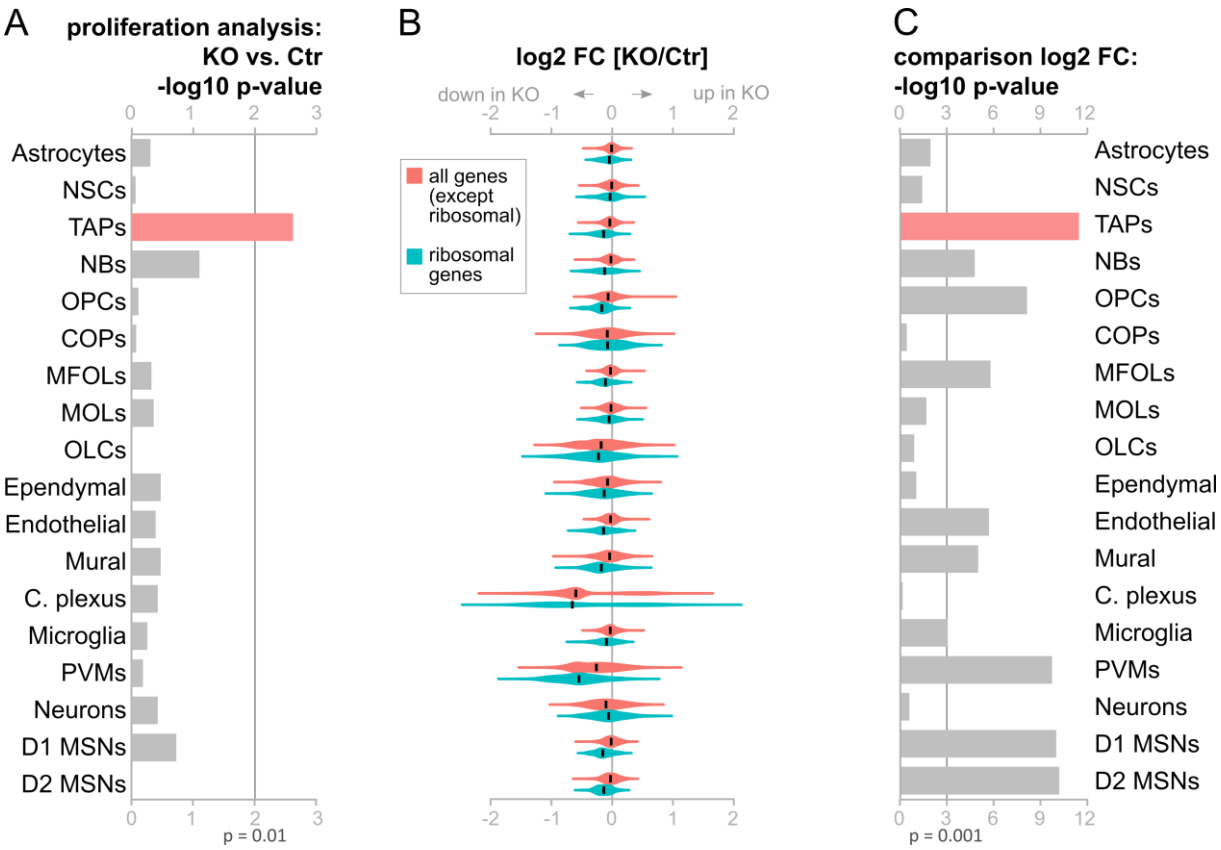


Figure S6. Proliferation is significantly reduced in TAPs, and ribosomal genes are decreased in the LRP2-deficient SVZ. Relate to Figures 4 and 5.

(A) P-values of the proliferation analysis for all cell types. Statistical significance of data was determined using Kolmogorov-Smirnov test between the cumulative fraction distributions of cells scored for cell cycle (LRP2 KO versus Ctr). The p-value for TAPs is highlighted in red.

(B) Distribution of log2 fold changes (LRP2 KO versus Ctr, cells averaged per cluster) of ribosomal (teal) and all remaining genes (red). Expression cut off ≥ 0.5 log2 (mean normalized UMI count) Ctr cells.

(C) P-values for panel (B). In 10 of 18 identified cell clusters, ribosomal genes were significantly lower expressed as all remaining genes in cells derived from LRP2 KO compared to Ctr SVZ (p-value < 0.001). TAPs (highlighted in red) had the smallest p-value. Statistical significance of data was determined using unpaired Student's t test.

Abbreviations: FC: fold change; all other abbreviations as in Figure S6.