Supplementary Information for

High resolution mouse subventricular zone stem cell niche transcriptome reveals features of lineage, anatomy, and aging

Xuanhua P. Xie,[1,2] Dan R. Laks,[1,2,8] Daochun Sun,[1,2] Asaf Poran,[3,9] Ashley M. Laughney,[3] Zilai Wang,[1,2] Jessica Sam,[4] German Belenguer,[5,6] Isabel Fariñas,[6] Olivier Elemento,[3] Xiuping Zhou,[7] Luis F. Parada[1,2,*]

[1]Cancer Biology and Genetics Program; [2]Brain Tumor Center Memorial Sloan Kettering Cancer Center, 1275 York Ave., New York, NY 10065, USA; [3]Institute for Computational Biomedicine, Department of Physiology and Biophysics; Weill Cornell Medicine, New York, NY 10065, USA; [4]Biochemistry, Cell & Molecular Biology Graduate Program, Weill Cornell Medicine, New York, NY 10065, USA; [5]Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid 28031, Spain; [6]Departamento de Biología Celular, Biología Funcional y Antropología Física, Universidad de Valencia, Valencia 46010, Spain; [7]Institute of Nervous System Diseases, Xuzhou Medical University, Jiangsu 221002, PR China; [8]Present address: Voyager Therapeutics, Cambridge, MA 02139, USA. [9]Present address: Neon Therapeutics, Cambridge, MA 02139, USA.

Corresponding author: Luis F. Parada
paradal@mskcc.org

646 888 3781

This PDF file includes:

- Supplementary text
- Figures S1 to S7
- Legend for Movies S1
- Legends for Dataset Tables S1 to S3
- SI References

Other supplementary materials for this manuscript include the following:

www.pnas.org/cgi/doi/10.1073/pnas.2014389117

Movies S1

Dataset Tables S1 to S3

Supplementary Information Text

Supplemental Materials and Methods

Single cell sequencing.

The Dropseq cell-bead collection, sample preparation, library preparation, and sequencing were performed as previously described (1). Beads were purchased from Chemgenes (#MACOSKO-2011-10) and the microfluidics chip was purchased from Nanoshift. All our reagents were purchased from the recommended sources as outlined in Macosko et al., 2015 and the McCarroll lab's online Dropseq protocol, v3.1, Dec. 2015 (http://mccarrolllab.com/download/905/). Our primers and oligonucleotides were purchased from IDT with the identical sequences outlined by Macosko et al., 2015. Sequencing was done on a NextSeq 500 at the Cornell Sequencing core, 64bp (R2). Alignment of reads and generation of cellular expression counts were performed with the Drop-seqAlignmentCookbookv1.2Jan2016 (http://mccarrolllab.com/wp-content/uploads/2016/03/Drop-seqAlignmentCookbookv1.2Jan2016.pdf) and with the Dropseq tools provided by the McCarroll lab (http://mccarrolllab.com/download/922/). We utilized the mm10 reference genome into which we incorporated the *CGD* transgene mRNA sequence. The transgene expression level is determined by capturing of the 3' end of the transcript that encodes the *DTR* gene. Mus_musculus.GRCm38.82 was used as our refflat GTF. The Digital Expression NUM_CORE_BARCODES output was set to 5000 as we knew we had less than or equal to 2000 cells per sample.

Single cell analysis by "Seurat"

All our analysis was performed in R version 3.3.0 with R Studio. The "Seurat" R package was used to filter, normalize, cluster and select differentially regulated genes for each cluster. Seurat package (version 1.4.0.16 and, when released, version 2.0) was utilized as described in the online clustering tutorials by the Satija lab (http://satijalab.org/seurat) (1). Cells were excluded if genes were detected in less than 3 cells, or had less than 200 unique genes. The resultant matrix was normalized to 10,000 transcripts, log transformed, and a regression to the number of unique molecular identifiers (UMIs) was performed before dimensional reduction. Principle components 1-19 were used to identify subpopulation clusters. We compared the lists of differential genes for each group in a cohort to select genes that were unique to a group, in other words, they were signature genes that were differentially regulated in a certain group but not in any other group. This produced two lists, differentially up-regulated genes and unique up-regulated genes. We used the differentially up-regulated gene list for gene ontology analysis by DAVID (2), (https://david-d.ncifcrf.gov/summary.jsp). We used the top 10 unique genes (DEGs when not enough, Table S2B & 2C) for each group to produce the heatmap with the Seurat package (Figure 3C). We performed batch correction in Seurat (version 2.0) to cluster and project both the Aged samples and the sorted two-month samples together (Fig. S7B-C). To do this we used the union of the top 2000 most variable genes for each data set and ran the canonical correlation analysis (CCA) in Seurat that identifies common sources of variation between the two datasets. For this analysis we used the first 19 dimensions and all the other settings and analyses were the same as above.

When Seurat version 3 was released, we re-did both the sorted wildtype and aged samples analyses with the latest version. As the samples were from the same tissue, preparation date, sequencing run, and there were no discernible batch effects in the sorted samples (GFP^{hi}, GFP^{lo}, GFP-, Unsorted) we did not use the integration method for this analysis but rather simply merged the samples into one matrix before the Seurat pipeline so as not to introduce any noise from unnecessary processing. In contrast, for the aged samples we used the anchor integration method (Stuart et al., 2019). In both instances, the original results were conserved and proved robust across different pipelines and versions of Seurat. One exception was the co-clustering of H1 with H2 into one cluster in the Seurat version 3 processing of the sorted two-month SVZ samples. As the results and clustering were largely conserved across different versions of Seurat, we present the original analysis with confidence that they are robust results irrespective of the Seurat versions and processing employed.

Pseudotime Analysis

Pseudotime analysis was performed using the Destiny package as detailed in (3) and in the vignette at https://www.helmholtz-muenchen.de/icb/research/groups/quantitative-single-cell-dynamics/software/destiny/index.html. We used H0-L2 groups from Fig. S3A for our Pseudotime analysis in Fig. 4F. In Destiny, we plotted the result with our established color codes. The animation was compiled in Sequimago and labels were added in Wondershare Filmora (Version 7.8.9). Pseudotime analysis was also performed in a similar manner on the H and L groups of the aged samples that were included with the aforementioned pseudotime matrix of H and L groups in Fig. 4F (Fig. 7B).

Allen Brain Atlas SVZ localization of unique genes

The gene list used for this experiment is the same as those used in the aforementioned Heatmap section (see above). We collected images from the Allen Brain Atlas (4), http://mouse.brain-map.org/) where the SVZ was at its largest in the sagittal sections. We authored a macro in ImageJ that got rid of the white space in our quantification. We then made 5 sections of the SVZ from Ventral to Dorsal and used ImageJ to quantify the in situ hybridization of RNA for each of the five sections. We then normalized the values to the most ventral section and plotted the values, performed linear regression, and in GraphPad (Prism7) we calculated whether the slopes were significantly different than 0. The p-values for this test are reported.

Venn Diagrams

Venn diagrams were created by InteractiVenn (http://www.interactivenn.net/) (5) and reproduced in Photoshop CS6 (Version 13.0.6) to improve the quality and size of the associated text.

Transcription factor analysis

Transcription factors related in a sign sensitive manner to our inputted lists of up- and down-regulated genes were identified by TFactS by target gene signatures curated from microarray gene expression data (http://www.tfacts.org) (5, 6).

String network

Gene networks were generated by String database (https://string-db.org) (7) with medium confidence (0.400) to assess known and predicted interactions (physical and functional) between genes in our inputted gene sets. The PPI enrichment score p-value is a measure of the significance of enrichment for network interactions within a gene set.

qRT-PCR with CD95+ sorted samples

We sorted 1000 GFP^{hi};CD95+, GFP^{hi};CD95-, bulk GFP^{hi}, GFP^{lo}, GFP-, or DAPI-viable cells into tubes containing 3000 Dropseq beads (Chemgenes) in 100ul lysis buffer with DTT (#15508013, Thermo Scientific) (1). The tubes were rotated at 4 degrees for 10 minutes. Then we washed the samples twice with 6xSSC (1), followed with another wash with reverse transcription buffer (Thermo Scientific). Reverse Transcription mix with Template Switch Oligo were added and incubated for 30 minutes with rotation at room temperature and an additional 90 minutes with shaking at 700 rpm at 42 degrees Celsius. We then processed the samples as Dropseq samples outlined in Macosko et al., 2015 with Exonuclease treatment, PCR for 30-35 cycles, and Agencourt Ampure XP (#A63881, Beckmam-Coulter) bead purification of the resultant DNA. The purified DNA content was measured on a Qubit with high sensitivity reagents (#Q32854, Thermo Scientific). qRT-PCR reactions were performed with SYBR Select Master Mix (#4472903, Applied Biosystems) in an Applied Biosystems QuantStudio Flex6, real time PCR system. The following primers were selected from PrimerBank (8),

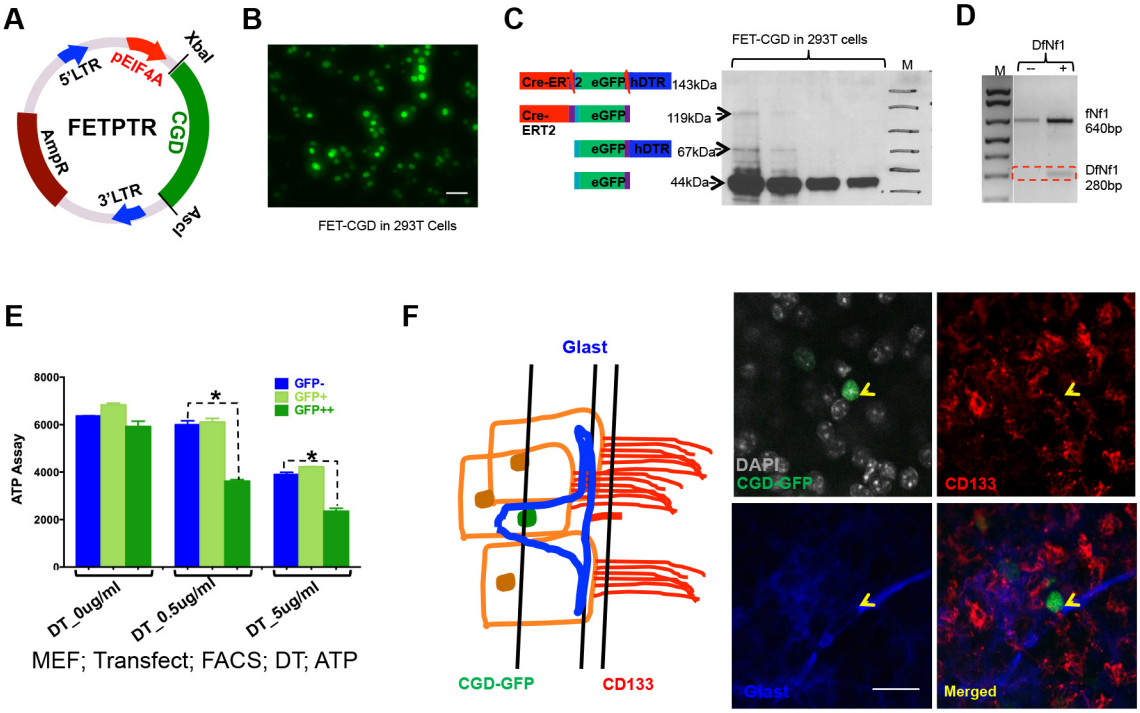
https://pga.mgh.harvard.edu/primerbank/index.html) except for Cd95 which was selected from Sabbagha et al. (9), and all primers were purchased from Eton Biosciences Inc.

qRT-PCR Primers for Mouse Genes

Gene (Mouse)	Fwd	Rev	Amplicon Size (bp)
1110017D15RIK	TGTCTCGGAACCGGACTCT	GGAGCAAGGGCTTTCATATCC	101
Fam183b	CGTGTGGGGCAGATGAAGAAT	GGTGAATGAGGTTCAGGAACTTG	148
Dynlrb2	GAATCCAGAGTCACAAAGGGG	GACCCGCATACTGAACCGTT	105
Rplp1	CTCGCTTGCATCTACTCCGC	AGAAAGGTTCGACGCTGACAC	109
Tmem212	GGTACACAGGATGGAGCGTTT	GCTTCCCACAAGTGTCTCTGG	121
Actb	GGCTGTATTCCCCTCCATCG	CCAGTTGGTAACAATGCCATGT	154
Hsp90ab1	TCAAACAAGGAGATTTTCCTCCG	GCTGTCCAACTTAGAAGGGTC	102
Cd95 (Fas)	TATCAAGGAGGCCCATTTTGC	TGTTTCCACTTCTAAACCATGCT	148

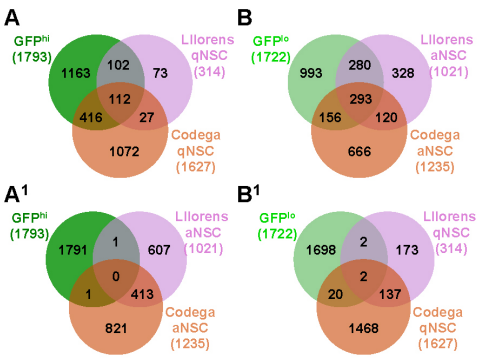
SCDE analysis

We employed the package SCDE (https://www.nature.com/articles/nmeth.2967) (10) to determine differentially expressed genes between 2 Months and 12 Months aged samples in both the combined H1-H2 groups (N=93 for 2 months, N=115 for 12 months) and in the combined H3-L0-L1 groups (N=84 for 2 months, N=63 for 12 months). We used non-log transformed, 'Seurat' normalized data matrices as the input. We chose genes with FDR P values <0.05 for further analysis including gene ontology (David, https://david-d.ncifcrf.gov/summary.jsp) (2).

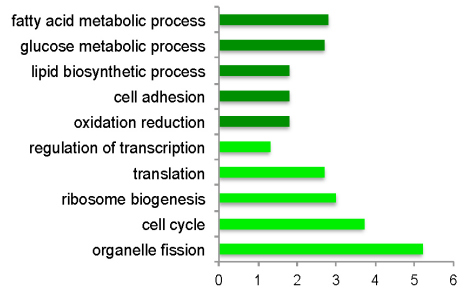


Adapted from Mirzadeh et al., 2008

Fig. S1. This figure accompanies Figure 1. *CGD* transgene construct characterization. (A) FETPTR expression vector. (B) eGFP reporter detection in FETPTR-*CGD* construct transfected 293T cells. Scale bar: 20μm. (C) Western blot with GFP antibody verifies efficient production of GFP monomers. 293T cells were transiently transfected with virus packaging a pEIF4A promoter driven *CGD* cassette and collected three days later in 2X laemmli for western blot. (D) CreER protein from the *CGD* transgene can remove the floxed region within an NF1 gene allele, detected by PCR. (E) Mouse embryo fibroblasts transfected with the FETPTR-*CGD* derived viruses are sensitive to diphtheria toxin in an ATP assay. (F) Left: Diagram of classic SVZ "pinwheel" structure in which a eGFP+/Glast+/CD133+ stem cell is surrounded by CD133+ ependymal cells. The three black lines indicate the relative positions for the microscope images on the right. Right: Corresponding IHC SVZ images with a *CGD*-GFP^{hi} nucleus and Glast+ cytoplasm surrounded by CD133+;GFP- ependymal cells. Scale bar: 20μm.



C



D

qNSC GFP^{hi} (1793)

Term	Fold Enrichment	FDR
hexose metabolic process	2.9	0.000
monosaccharide metabolic process	2.7	0.000
cell adhesion	1.9	0.000
biological adhesion	1.9	0.000
fatty acid metabolic process	2.7	0.000
oxidation reduction	1.8	0.000
glucose metabolic process	2.7	0.000
membrane organization	1.9	0.002
lipid catabolic process	2.4	0.002
endocytosis	2.1	0.002
membrane invagination	2.1	0.002
extracellular structure organization	2.3	0.004
sulfur metabolic process	2.7	0.005
organophosphate metabolic process	2.1	0.006
polysaccharide metabolic process	2.6	0.010
regulation of cell proliferation	1.6	0.012
cell-substrate adhesion	3.1	0.014
regulation of lipid metabolic process	3.1	0.018
ion transport	1.5	0.019
gland morphogenesis	2.7	0.019
cellular amino acid derivative metabolic process	2.2	0.020
lipid biosynthetic process	1.8	0.033
phospholipid metabolic process	2.1	0.035
regulation of cell adhesion	2.5	0.049

E

aNSC GFP^{lo} (1722)

Term	Fold Enrichment	FDR
cell cycle	3.6	0.000
mitotic cell cycle	4.9	0.000
cell cycle phase	4.2	0.000
cell cycle process	3.9	0.000
M phase	4.5	0.000
M phase of mitotic cell cycle	5.3	0.000
nuclear division	5.2	0.000
mitosis	5.2	0.000
organelle fission	5.0	0.000
cell division	4.2	0.000
DNA metabolic process	3.2	0.000
RNA processing	2.8	0.000
DNA replication	4.3	0.000
mRNA metabolic process	3.1	0.000
RNA splicing	3.7	0.000
mRNA processing	3.3	0.000
chromosome organization	2.6	0.000
DNA repair	3.3	0.000
translation	2.8	0.000
response to DNA damage stimulus	2.9	0.000
chromosome segregation	5.4	0.000
microtubule-based process	3.0	0.000
cellular response to stress	2.3	0.000
chromatin organization	2.4	0.000
DNA packaging	3.8	0.000
microtubule cytoskeleton organization	3.5	0.000
ribonucleoprotein complex biogenesis	3.2	0.000
chromatin assembly or disassembly	3.5	0.000
macromolecular complex subunit organization	2.2	0.000
cellular macromolecular complex subunit organization	2.5	0.000
ribosome biogenesis	3.2	0.000
cytoskeleton organization	2.2	0.000
establishment of RNA localization	3.8	0.000
nucleic acid transport	3.8	0.000
RNA transport	3.8	0.000
macromolecular complex assembly	2.1	0.000
spindle organization	5.8	0.000
RNA localization	3.8	0.000
nucleobase, nucleoside, nucleotide and nucleic acid transport	3.5	0.000
cellular macromolecular complex assembly	2.3	0.000
chromatin assembly	3.5	0.000
nucleosome organization	3.5	0.000
protein-DNA complex assembly	3.5	0.000
regulation of cell cycle	2.3	0.000
mRNA transport	3.7	0.000
chromatin modification	2.2	0.000
DNA recombination	3.4	0.000
regulation of transcription	1.3	0.000
transcription	1.4	0.000
DNA-dependent DNA replication	4.8	0.000
nucleosome assembly	3.3	0.001
DNA replication initiation	9.1	0.001
ncRNA metabolic process	2.2	0.001
rRNA processing	3.3	0.001
rRNA metabolic process	3.2	0.001
regulation of organelle organization	2.4	0.001
mitotic sister chromatid segregation	6.2	0.003
regulation of cytoskeleton organization	2.8	0.004
sister chromatid segregation	5.8	0.006
RNA catabolic process	3.6	0.015
microtubule-based movement	2.6	0.022
cytokinesis	4.5	0.045
cell cycle checkpoint	3.3	0.046
ncRNA processing	2.2	0.047
macromolecule catabolic process	1.5	0.048

Fig. S2. This figure accompanies Figure 2. Gene expression analysis of *CGD*-GFP^{hi} and GFP^{lo} cells point to a quiescent versus a proliferative state respectively. (A) Venn diagram analysis comparison of *CGD*-GFP^{hi} DEGs and published quiescent NSC signatures demonstrates significant overlap but not with activated NSCs (A^1). (B) GFP^{lo} cell signatures preferentially overlap with that of published activated NSCs but not with quiescent NSCs (B^1). (C) Gene Ontology analysis of *CGD*-GFP^{hi} and GFP^{lo} profiles is consistent with a stem versus progenitor state. (D) and (E) Known and unknown biological processes related to quiescent or proliferative NSCs are associated with GFP^{hi} (D) or GFP^{lo} cells (E).

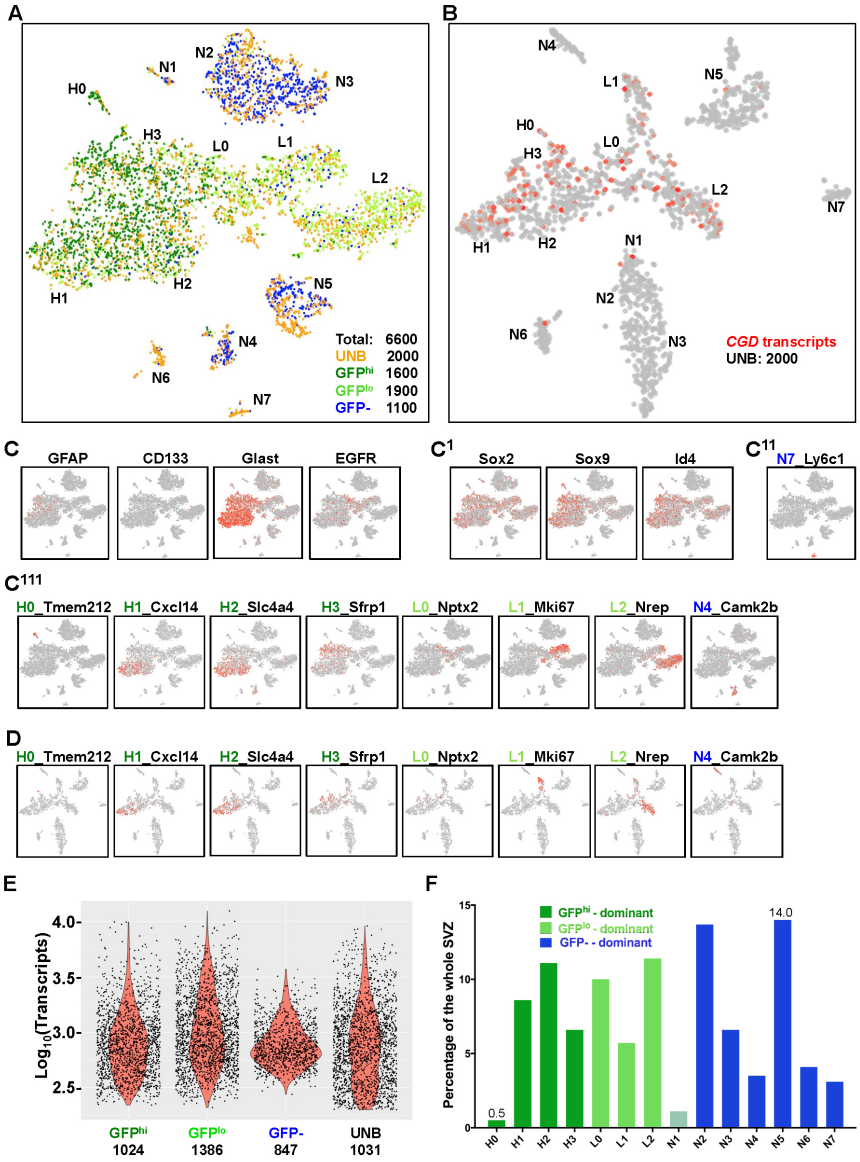
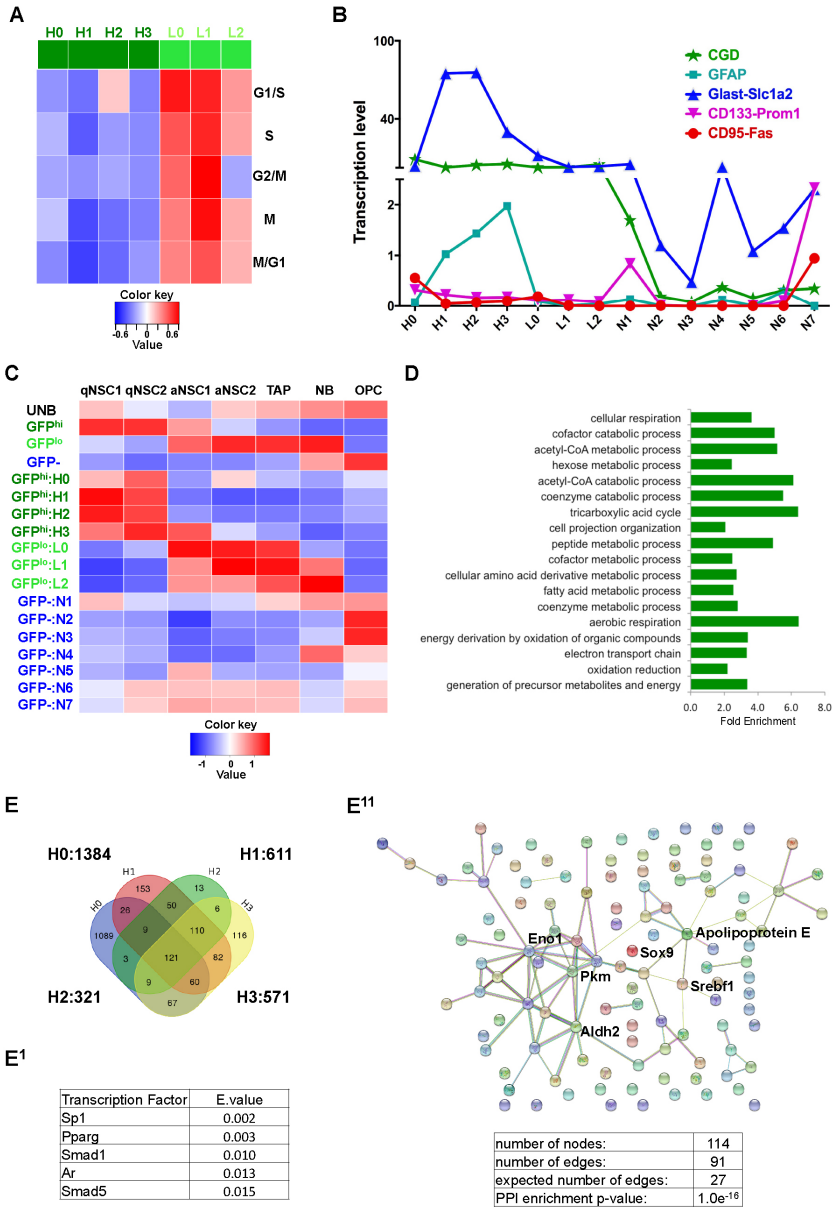


Fig. S3 This figure accompanies Figure 3 and Table 2. Drop-seq analysis of whole SVZ. (A) Combined 6600 cells yield fourteen groups that integrate the GFP^{hi}, GFP^{lo}, and GFP- subgroups. Note UNB cells exist in each of the fourteen groups. (B) tSNE projection of UNB cells shows similar groups and enriched *CGD* expression in H1-L2 lineage. (C) Distribution of sorting markers for stem/progenitor cells used in previous studies. (C^{1}) Representative distribution of NSC and progenitor specific transcription factors; (C^{11}-C^{111}) candidate markers identified in this study for GFP-:N7 and different SVZ stem/progenitor groups. (D) Representative markers for NSC/Progenitor/neuron lineage facilitate the identification of respective populations in UNB cells. (E) Transcript numbers for each of the four sorted and Dropseq analyzed samples present a “bell” shape distribution. Average transcripts per cell for each sample are indicated at the bottom. (F) SVZ UNB cells contain fourteen distinct subgroups ranging from 0.5% to 14% of the whole tissue.



E1

Transcription Factor	E.value
Sp1	0.002
Pparg	0.003
Smad1	0.010
Ar	0.013
Smad5	0.015

number of nodes:	114
number of edges:	91
expected number of edges:	27
PPI enrichment p-value:	$1.0e^{-16}$

Fig. S4. This figure accompanies Figure 4. Granular view provided by analysis of the fourteen groups in the adult SVZ. (A) Cell cycle signature analysis of the GFP^{hi} and GFP^{lo} (H&L) groups reveals heterogeneity. Subgroup L1 shows the highest mitotic index for all seven signatures (see Fig. 2E). (B) Transcription levels of the five NSC markers within the fourteen populations (all standard errors < 1.2, units are normalized values scaled to 10,000 counts/cell). (C) Published single cell gene signatures fail to distinguish between the four GFP^{hi} subgroups. (D) Known and unknown biological processes related to quiescent NSCs are revealed by a comprehensive list of genes derived from H0-H3 groups. (E-E^{11}) An overlap analysis of genes commonly expressed in H0-H3 groups results in 121 genes (E), putative transcription factors associated to this list (TfactS analysis) (E^{1}), and enrichment for a gene network is revealed (STRING analysis) (E^{11}).

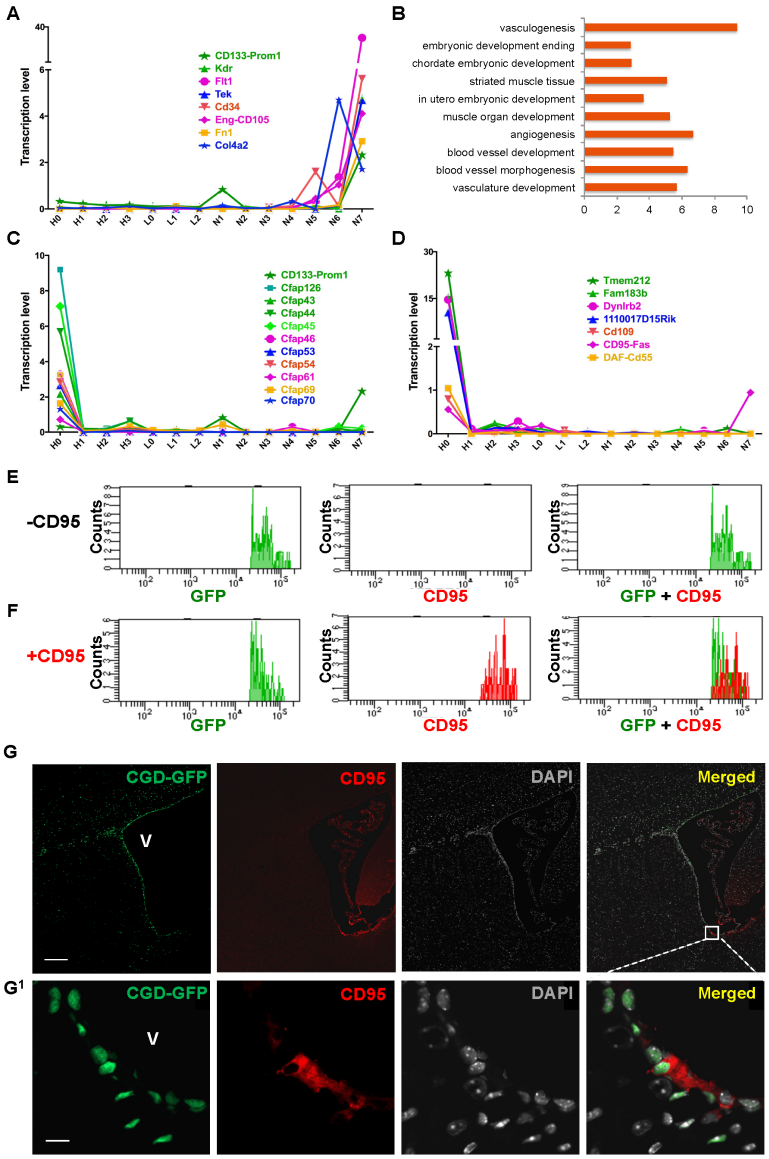


Fig. S5. This figure accompanies Figure 5. H0 subgroup gene analysis. (A) A gene network proposed as ependymal-like NSCs is enriched in the GFP-:N7 subgroup that by our analysis represents endothelial progenitor cells (all standard errors < 0.2, units are normalized values scaled to 10,000 counts/cell). (B) Gene ontology analysis of GFP-:N7 associate to vasculature and blood vessel development. (C) Cilia associated genes are uniquely expressed in GFP^{hi}:H0 subgroup (all standard errors < 0.2, units are normalized values scaled to 10,000 counts/cell). (D) Expression profiles of seven potential markers for H0 subgroup cells (all standard errors < 2.0, units are normalized values scaled to 10,000 counts/cell). (E) and (F) Among the GFP+ subgroups of cells, CD95 antibody preferentially labels cells with higher levels of GFP proteins. (G) IHC analysis reveals some CD95+ cells co-localized with *CGD*-GFP at ventral SVZ. Scale bars: 200μm (G) and 10μm (G^1). V: lateral ventricle in (G and G^1).

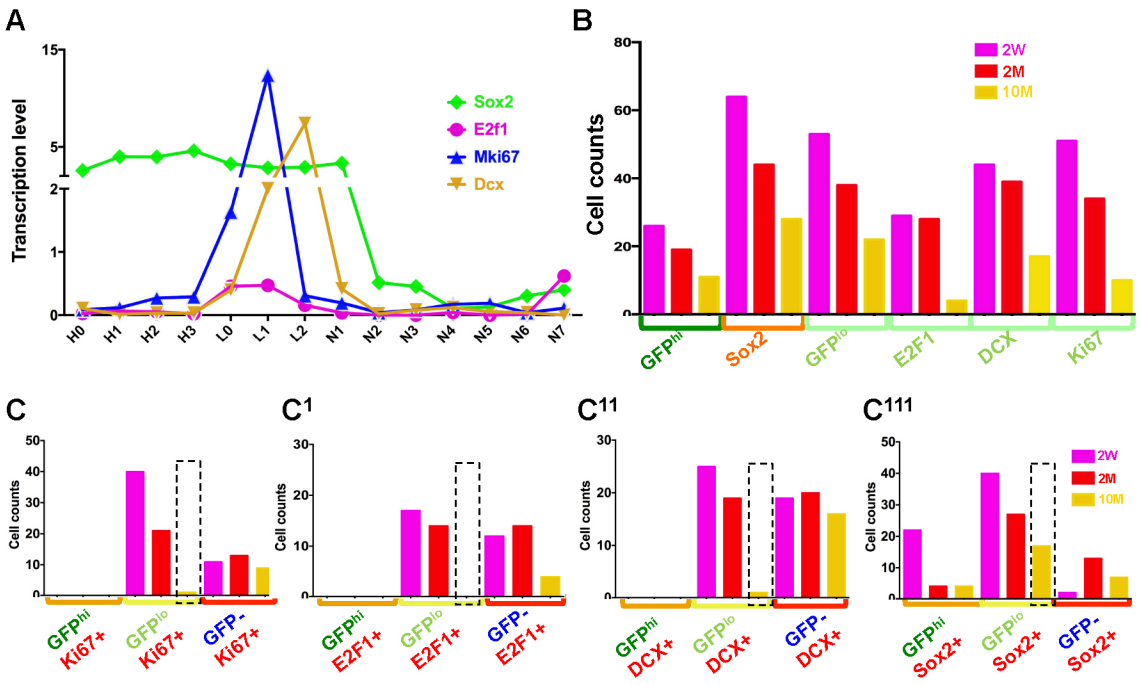


Fig. S6. This figure accompanies Figure 6. Age-related alteration of GFP^{hi} and GFP^{lo} lineage related genes. (A) Candidate genes for GFP^{hi} qNSC and GFP^{lo} progenitor cells illustrated in the adult SVZ analysis (all standard errors < 0.6, units are normalized values scaled to 10,000 counts/cell). (B) Concordant with age related GFP expression decrease, cells expressing four candidate genes decrease over age. (C-C^{111}) The *CGD*-GFP-associated expression of the three candidate genes for NSC activation (C-C^{11}, dashed line) drops more compared to the pan-NSC marker Sox2 (C^{111}, dashed line).

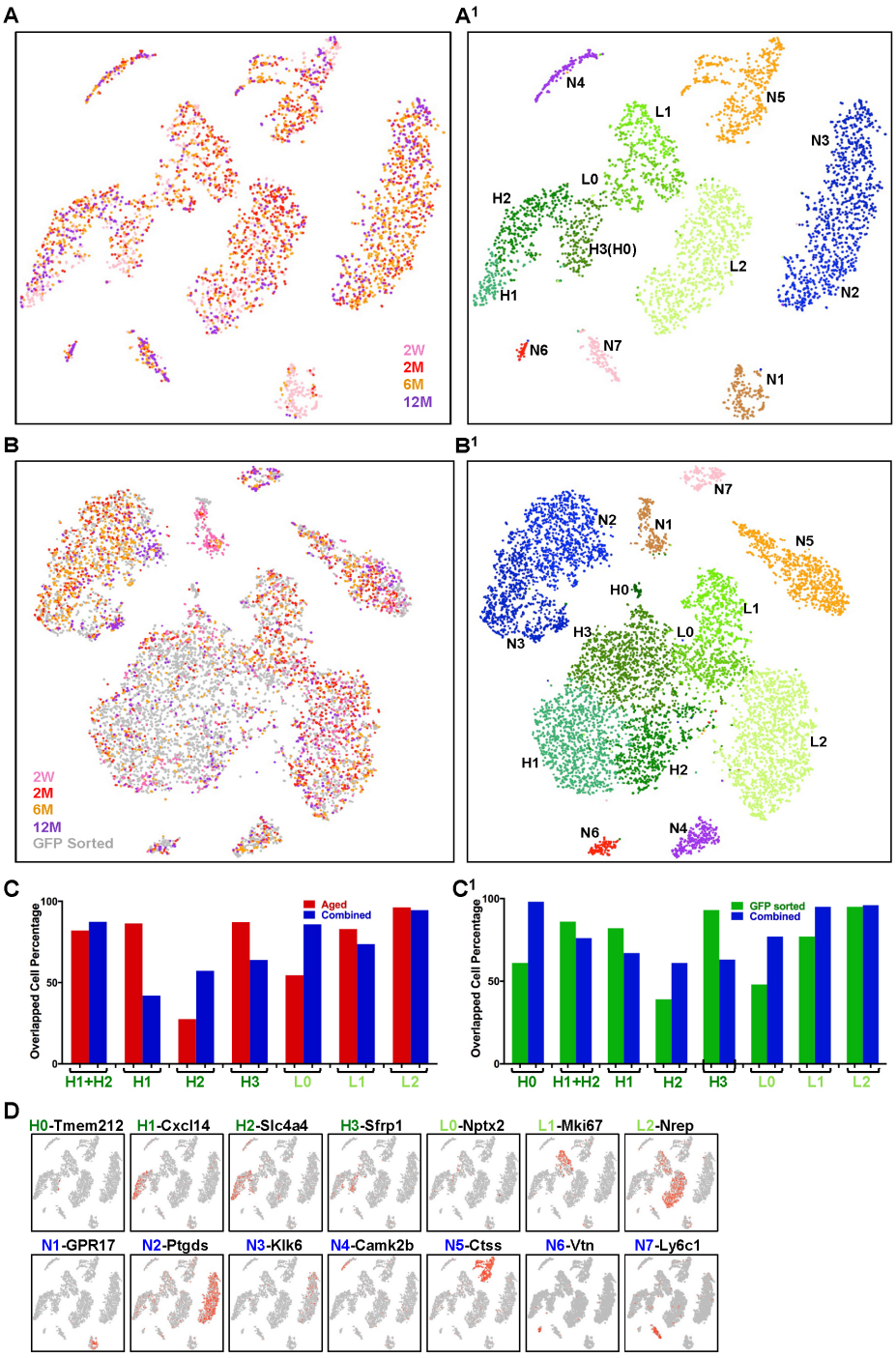


Fig. S7. This figure accompanies Figure 7. tSNE projection and marker gene analysis of the SVZ cells over age. (A-A[1]) tSNE projection of the 5,600 single cells from two-week, two-month, six-month, and twelve-month SVZs reveals sample distribution (A) and thirteen groups (A[1]). The groups were assigned by comparing the cell constituents of the GFP sorted cohort (Fig. 3A), the Aged sample cohort (Fig. S7A[1]), and the combined analysis (Fig. S7B[1]). They were further confirmed by comparing unique genes from this analysis (Table S3C) with the unique genes from the GFP sorted cohort analysis (Table

S2C). DEGs were used instead when there were not enough unique genes (H2). (B-B[1]) tSNE projection of the comprehensive populations of the combined GFP sorted cohort with the Aged sample cohort (12,200 cells) demonstrates sample distribution (B) and fourteen conserved populations (B[1]). (C-C[1]) Each stem/progenitor group identity in the Aged sample cohort is examined further by comparing the cell constituents in (C) the combined versus the Aged samples and (C[1]) the combined versus the GFP sorted cohort. More than half of the cells are overlapped between the two cohorts in most cases. A significant amount of H2 cells are assigned into the H1 group in the combined analysis, possibly due in part to batch effect correction. (D) Marker genes identified in the two-month SVZ samples guide the identification of Seurat based clusters of SVZ cells from different ages. H0 cells (9 cells in total) are too few to form a group in the Aged cohort by themselves.

Movies S1 Legend

To generate movie S1, the CGD-GFP^{hi} cells were sorted from freshly dissociated mouse SVZ, incubated in Cytation 5 (BioTek), and imaged every hour for GFP and cell morphology.

Dataset Table Legends

Dataset Table S1, related to Figure 2E-2F and S2. Transcriptomes (TS1A), differentiated expressed genes (TS1B), and GO analysis (TS1C&D) of the bulk RNAseq data derived from the *CGD*-GFP^{hi} vs GFP^{lo} samples. Complete lists of NSC genes derived from various studies including the *CGD*-GFP^{hi} vs GFP^{lo} samples are summarized in TS1E.

Dataset Table S2, related to Figure 3, S3, 4, S4, and S5A-D. Transcriptomes (TS2A), differentiated expressed genes (TS2B), unique genes (TS2C), and GO analysis (TS2D) of the fourteen cell groups derived from the two-month old murine SVZs. The 1914 genes derived from H0-H3 groups and associated GOs in Fig. S4D are summarized in (TS2E). The 121 overlapped G genes used in Fig. 4C&S4E, and related transcription factors are shown in (TS2F&G).

Dataset Table S3, related to Figure 7 and S7. Transcriptomes (TS3A), differentiated expressed genes (TS3B), and unique genes (TS3C) of the thirteen cell groups derived from the SVZs of two-week, two-month, six-month, and twelve-month mice. Transcriptomes of each cell group in the four aged SVZs are summarized in (TS3D). In lineage H1-H2, genes and GOs up-regulated in either two-month or twelve-month are shown in (TS3E&F). Transcription factors associated to the age-related change are indicated in (TS3G). In lineage H3-L1, genes and GOs up-regulated in either two-month

or twelve-month are shown in (TS3H&I). Potential transcription factors involved in the age-related blockage are summarized in (TS3J).

SI References

1. E. Z. Macosko *et al.*, Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202-1214 (2015).
2. W. Huang da, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).
3. P. Angerer *et al.*, destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241-1243 (2016).
4. E. S. Lein *et al.*, Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168-176 (2007).
5. H. Heberle, G. V. Meirelles, F. R. da Silva, G. P. Telles, R. Minghim, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* **16**, 169 (2015).
6. A. Essaghir *et al.*, Transcription factor regulation can be accurately predicted from the presence of target gene signatures in microarray gene expression data. *Nucleic Acids Res* **38**, e120 (2010).
7. D. Szklarczyk *et al.*, The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, D362-D368 (2017).
8. X. Wang, A. Spandidos, H. Wang, B. Seed, PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update. *Nucleic Acids Res* **40**, D1144-1149 (2012).
9. N. G. Sabbagha *et al.*, Alternative splicing in Acad8 resulting a mitochondrial defect and progressive hepatic steatosis in mice. *Pediatr Res* **70**, 31-36 (2011).
10. P. V. Kharchenko, L. Silberstein, D. T. Scadden, Bayesian approach to single-cell differential expression analysis. *Nat Methods* **11**, 740-742 (2014).