694 **Supplementary Methods**
695
696 **Mice**
697 All mouse strains were maintained at the University of California, San Francisco (UCSF)
698 specific pathogen-free animal facility and all animal protocols were approved by and in
699 accordance with the guidelines established by the Institutional Animal Care and Use
700 Committee and Laboratory Animal Resource Center. The day of birth was designated as
701 Postnatal day 0 (P0).
702 Unless otherwise started, wild-type Swiss Webster mice were used for histology
703 and spatial transcriptomics. *Aldh1l1-GFP* transgenic mice were generated by the
704 GENSAT project (37). *Emx1-cre* mice were obtained from The Jackson Laboratory (JAX
705 #005628) (38). Satb2-flox mice were a gift from Dr. Ralph Marcucio at UCSF (27).
706 Conditional knockouts were generated by breeding *Emx1-cre/+; Satb2-flox/+* with
707 *Satb2-flox/flox* mice to obtain mutants (*cre/+; flox/flox*) and littermate controls
708 (*cre/+;flox/+*). *Reeler* (*Reln/Reln*) mice were a gift from Dr. Eric Olson at SUNY Upstate
709 University (B6C3Fe a/a-Reln; JAX). Reln/+ mice were used as littermate controls. All
710 mice were maintained on a mixed background.
711
712 **Mouse tissue preparation**
713 Mice were transcardially perfused at P14 or P56 with ice-cold phosphate buffer saline
714 (PBS) and 4% paraformaldehyde (PFA) in 1X PBS. Brains were dissected and post-fixed
715 in 4% PFA for 24 hours at 4°C. Post-fixed brains were cryo-protected in 30% sucrose for
716 48 hr at 4°C and embedded in optimal cutting temperature compound (Tissue-Tek).
717 Cryosections (16 microns) were collected on superfrost slides (VWR) using a cryostat
718 (CM3050S, Leica) and stored at -80°C.
719
720 **Human tissue**
721 Human brain tissue was obtained with informed consent under protocol 16/LO/2168
722 approved by the NHS Health Research Authority at the Addenbrookes Hospital. Adult
723 brain tissue biopsies were taken from the site of neurosurgery resection for the original
724 clinical indication. For the purposes on this study samples were taken from peri-
725 contusional areas in traumatic brain injury (frontal cortex), lobectomy in epilepsy surgery
726 (temporal cortex) and peri-tumoural tissue (temporal cortex).
727 Tissue specimens were collected in Addenbrookes Hospital and transferred to a
728 CL2 facility where it was processed. Tissue was dissected and fixed in 4% PFA for 48-72
729 hours. Once fixed, samples were placed in 20% Sucrose for cryoprotection for 24-48
730 hours and mounted in OCT, stored at -80C.
731 Additional human brain tissue was collected in a de-identified manner with
732 previous patient consent in strict observance of the legal and institutional ethical
733 regulations of the University of California, San Francisco (UCSF) Committee on Human
734 Research. Protocols were approved by the Human Gamete, Embryo and Stem Cell
735 Research Committee (Institutional Review Board) at UCSF. For this study, one post-
736 mortem sample was taken from the superior frontal gyrus and processed as above.
737 Human brain blocks were cryosectioned to 16 microns.
738
739

740 **smFISH assay design and probes**

741 Mouse and human tissue smFISH was performed using the RNAScope LS Multiplex
742 Assay (Advanced Cell Diagnostics, ACD) (39). In this assay, the smFISH signal-to-noise
743 ratio (SNR) is amplified by branched DNA complexes formed on target transcripts and
744 tyramide signal amplification (TSA)-based labeling. Target RNAs are initially hybridized
745 to a series of single-stranded DNA "z-probes". Each z-probe is composed of (**1**) a 18-25
746 nucleotide region complementary to the target RNA, (**2**) a spacer sequence, and (**3**) a 14
747 nucleotide tail region. These probes are tagged by branched DNA-amplification trees:
748 pairs of z-probes are hybridized to oligo-preamplifiers, across their bridged tail
749 sequences, which are then tagged by 20 oligo-amplifiers (Fig 1 A). Each oligo-amplifier
750 is labeled with 20 Horse Radish Peroxidase (HRP) enzyme molecules. In general, a 1
751 kilobase region on the target transcript is hybridized by 20 z-probe pairs in tandem,
752 which can yield up to 8000 HRP labels per each target. The fluorescent smFISH signal is
753 consequently generated by the addition of tyramide-conjugated fluorophores. Tyramide is
754 enzymatically converted into a highly oxidized intermediate by HRP that covalently
755 binds to the proteins at or near the HRP label, depositing a large number of fluorophores
756 for probe detection (40). The combination of branched DNA and TSA-amplification
757 significantly boosts the sensitivity and SNR of the RNAScope assay, allowing fast
758 confocal imaging of large tissue areas with short exposure times (see imaging).
759 *Multiplexed detection*: To achieve 4-plex transcript detection with RNAScope,
760 target z-probes are assigned to one of four different channels (C1-C4) that contain
761 distinct tail-sequences. Tissue samples are hybridized to the mixture of C1-C4 probes,
762 followed by generation of channel-specific amplification trees. Finally, probes are
763 sequentially developed with TSA through incubation cycles of channel-specific HRP
764 labels, tyramide-conjugated fluorescent dyes and chemical enzymatic quenchers.
765 *Probe information*: All of the RNAscope probes used in this study and relevant
766 information including target sequences are listed in Supplementary Table 3. Further
767 information is readily available from the vendor (https://acdbio.com/catalog-probes). To
768 assess the background signal from the assay, target probes against the bacterial *DapB*
769 mRNA were used as negative controls. Target probes against mouse and human
770 housekeeping genes were used as positive controls. In the mouse cortical astrocyte
771 expression screen, the *Glast* probe was always assigned to C4 and multiplexed with other
772 probes in C1-C3 channels. With human tissue, the *Glast* probe was assigned to C3 and
773 multiplexed with C1-C2.
774
775 **Automated smFISH and IHC**

776 All histology on mouse and human brain cryosections was automated on a Leica BOND
777 RX robotic stainer after manual baking and dehydration. Tissue sections were first
778 processed for 3 or 4 gene smFISH using the RNAScope LS Multiplex Assay (ACD).
779 After smFISH, antibody staining was performed using TSA and slides were manually
780 coverslipped for imaging.
781
782 **i. Baking and dehydration**: Tissue cryosections were removed from -80°C and thawed
783 at RT for 15 min. Samples were then baked at 65°C for 45 min in vertical position on a
784 slide holder (Tissue-Tek) in an oven. After baking, samples were dehydrated in a series

785 of 50%, 70%, 100% and 100% ethanol (5 min each) in staining dishes (Tissue-Tek) and
786 air-dried for 10 min before automated RNAScope.
787
788 **ii. Automated histology design and setup:** For use on the Leica BOND RX, all
789 histology consumables were transferred to barcoded reagent containers. Staining
790 protocols that list the order and durations of reagent incubations and washes were created
791 on the BOND controller software. Slides were assigned unique barcode labels coupled to
792 staining protocols and placed onto temperature-controlled pads on the instrument. Flow-
793 through chambers were assembled across the whole slides using plastic coverplates.
794 During staining, a liquid volume of 150 µL was dispensed to each slide on every step
795 using automated liquid handling. Reagents were flushed at least once before main
796 incubations to ensure uniform coverage of the slide. Between reagent incubations,
797 multiple short washes were performed. All incubations were performed at room
798 temperature unless indicated otherwise. For the 4-plex RNAScope smFISH protocol, a
799 maximum number of 20 slides could be processed against 10 different probe mixtures in
800 a single run (e.g. 40 different genes screened across two sets of biological replicates). The
801 combined multiplexed RNAScope smFISH and IHC protocol for 20 slides ran overnight
802 on the Leica BOND RX lasting ~17 hours.
803
804 **iii. Automated RNAScope smFISH:** The RNAScope LS Multiplex assay (ACD) was
805 performed largely according to the instructions from the vendor and modifications are
806 noted below. Full details of the protocols 4-plex smFISH and the consumables used in
807 this study are provided in Supplementary Table 4.
808 *4-plex probe hybridization*: To perform 4-plex RNAScope on mouse brain
809 cryosections, samples were initially permeabilized with heat and protease treatment to
810 improve probe penetration and hybridization. For heat treatment, P14 and P56 samples
811 were incubated in BOND ER2 buffer (pH 9.0, Leica) at 95°C for 2 and 5 min,
812 respectively. For protease treatment, P14 and P56 samples were incubated in ACD
813 protease reagent at 42°C for 10 and 15 min, respectively. Prior to probe hybridization,
814 samples were incubated in hydrogen peroxide for 10 min to inactivate endogenous
815 peroxidases and ACD protease. Following pre-treatment, samples were incubated in
816 target z-probe mixtures (C1-C4) for 2 h at 42°C. The C2-C4 probes are provided at 50X
817 concentration by ACD and were diluted 1:50 in C1 probes. In exception, the C4 probe for
818 the high-expressing *Glast* mRNA was used at 1:100 in the astrocyte screen for reagent
819 conservation.
820 *smFISH signal amplification*: After probe hybridization, branched DNA
821 amplification trees were generated through sequential incubations in AMP1, AMP2 and
822 AMP3 reagents for 15-30 min each at 42°C with LS Rinse buffer high stringency washes
823 between incubations. Following amplification, probe channels were detected sequentially
824 via HRP-TSA labeling. To develop the C1-C3 probe signals, samples were incubated in
825 channel-specific HRP reagents for 15 min at 42°C, TSA fluorophores for 30 min and
826 HRP blocking reagent for 15 min at 42°C. The probes in C1, C2 and C3 channels were
827 labeled using Opal 520, 570 and 650 fluorophores (Perkin Elmer, diluted 1:2500)
828 respectively. Finally, to develop the C4 probe, the Atto425 fluorophore was used for 6-
829 color imaging on the Operetta system. The C4 probe complexes were first incubated with
830 TSA-biotin (Perkin Elmer, 1:500) for 30 min, followed by streptavidin-conjugated

831 Atto425 (Sigma, 1:400) for 30 min. Multiple short washes were performed between
832 incubations throughout the protocol using the BOND Wash buffer (Leica) and deionized
833 water (full protocol listed on Supplementary Table 4).
834 *3-plex smFISH*: To perform 3-plex RNAScope on human brain cryosections,
835 samples were heat-treated for 10 min and incubated in protease for 15 min. Probe
836 hybridization and branched DNA amplification were performed as described above. To
837 develop C1-C3 probes, Opal fluorophores (520, 570 and 650) were used at a lower
838 dilution (1:300) due to higher autofluorescence on postnatal human brain sections. To
839 distinguish RNA spots from lipofuscin autofluorescence, spots that appear identical
840 across Opal 520 and 570 channels were filtered out.
841
842 **iv. Automated immunohistochemistry:** RNAScope smFISH was directly followed by
843 antibody staining for the neuronal marker NEUN on the BOND RX system. To improve
844 antibody staining after IHC and perform 6-color imaging on the Operetta, the NEUN
845 signal was amplified using TSA-biotin and the Alexa 700 fluorophore. Samples were first
846 blocked in antibody blocking solution (Perkin Elmer) for 20 minutes. To block any
847 available TSA-biotin sites from the smFISH assay, samples were incubated in 0.2%
848 Avidin (Sigma) for 20 min and 0.05% Biotin (Sigma) for 30 min. After the avidin-biotin
849 block, samples were incubated in chicken anti-NEUN antibody (Milipore) diluted 1:500
850 in blocking solution for 1 hr. To develop the antibody signal, samples were incubated in
851 goat anti-chicken HRP (ThermoFisher, 1:500) for 1 hr, TSA-biotin (1:200) for 10 min
852 and streptavidin-conjugated Alexa 700 (Sigma, 1:200) for 30 min. Following antibody
853 staining, samples were incubated in DAPI (Sigma, 0.25 µg/ml) to mark cell nuclei and
854 washed multiple times in deionized water. After final washes, slides were briefly air-
855 dried and manually mounted using ~170 µL of Prolong Diamond Antifade (Fisher
856 Scientific) and standard coverslips (24x50 mm; Fisher Scientific). The full IHC protocol
857 is listed under Supplementary Table 4.
858
859 **Automated spinning disk confocal imaging**
860 Tissue sections were imaged on an Operetta CLS high-content screening microscope
861 (Perkin Elmer). To perform 6-color smFISH-IHC imaging, this system was equipped
862 with 8 LED light sources, 5X air and 40X water objectives, wide-field and spinning disk
863 confocal imaging modules and narrow band emission filters. The fluorophores, light
864 sources, exposure times and emission filters used for mouse and human tissue imaging
865 experiments are listed in Supplementary Table 5. Image acquisition and analysis were
866 controlled using the Harmony software (Perkin Elmer).
867
868 **i. Tissue identification**: To locate whole tissue sections or ROIs for high-resolution
869 imaging, entire slides were initially scanned under low magnification in wide-field mode.
870 Each slide was imaged for nuclear DAPI and NEUN staining if applicable using a 5X NA
871 0.16 objective (pixel size: 7.2 µm) under 5 minutes. To automatically locate the xy-
872 coordinates of tissue sections, a Harmony analysis script was used to detect DAPI+ areas.
873 Whole slide DAPI images were stitched, smoothed with Gaussian blurring and analyzed
874 with a global threshold. The detected DAPI+ areas were size filtered to remove staining
875 artifacts and slightly expanded to ensure complete tissue coverage. The resulting areas
876 were used to automatically set the xy-field positions of the subsequent 40X scan.

877 Alternatively, ROIs for 40X scans were manually selected on low magnification
878 previews. Selected 40X fields were imaged with a 7% overlap.
879
880 **ii. Confocal imaging**: The high-resolution smFISH images of tissue sections were
881 acquired on the spinning disk confocal mode using a sCMOS camera and a 40X NA 1.1
882 automated-water dispensing objective. The field-of-view was 320 x 320 µm and the pixel
883 size was 298 nm. A P14 mouse brain hemisection comprised 200 to 300 fields depending
884 on its anatomical position. Each field was imaged as a z-stack consisting of 20 to 30
885 planes with a 1 µm step size across each color channel. An IR laser was used to auto-
886 focus on the position of the coverslip and the relative z-heights of tissue sections were
887 manually identified by imaging DAPI on sample fields prior to tissue-wide scans. Each z-
888 plane was imaged across 4-6 channels depending on the experiment with exposure times
889 for mouse smFISH channels between 60 and 120 ms (Supplementary Table 5). The 40X
890 multi-channel settings and tissue heights were entered into an experimental layout on
891 Harmony and automatically executed after low magnification tissue identification scans.
892
893 **Image analysis**
894 To segment single neurons and astrocytes and quantify RNA spots from high-resolution
895 images, analysis scripts were created on Harmony software (Perkin Elmer) using
896 customizable pre-defined function blocks (*italicized below*). The complete single neuron
897 and astrocyte segmentation pipelines are provided in Supplementary Tables 6 and 7. Each
898 40X field was analyzed separately to optimize processing time of large datasets.
899
900 **i. Quantification of neuronal gene expression *in situ*:**
901
902 **a. Segmentation of neurons:** Maximum intensity z-projection images were calculated
903 across each channel to generate 2D images from z-stacks. NEUN+ neurons were
904 segmented in three steps. **1**) Supervised texture segmentation was performed at a coarse
905 scale to locate NEUN+ areas on images and filter staining artifacts on tissue sections
906 (*find texture regions*). Intensity and size thresholding then identified the NEUN+
907 neuronal soma (*find image region*). **2**) Neuronal nuclei were segmented within the
908 NEUN+ soma from Gaussian blurred DAPI images using intensity, size and contrast
909 thresholds (*find nuclei*). **3**) The neuronal cytoplasm was segmented around the nuclei
910 within the boundaries of the neuronal soma using NEUN intensity thresholding (*find*
911 *cytoplasm*).
912
913 **b. Filtering single neurons**: Neuron segmentation yielded single neurons as well as
914 doublets/triplets that overlap in z-projection images and neurons that are partially
915 contained in tissue sections (Supplementary Figure 1B). To automatically distinguish
916 single neurons, morphological (e.g. area, roundness) and intensity (e.g. pixel sum over
917 DAPI and NEUN) properties of segmented cells were measured and used to train a
918 supervised linear classifier (*select population*). For the training set, we manually selected
919 more one hundred single, doublet and partial cells across multiple tissue sections and
920 cortical areas. The resulting classification was validated across the cortex by manual
921 inspection of several fields.
922

923 **c. RNA spot calling in neurons**: RNA spots were identified by the detection of local
924 intensity maxima across each smFISH channel in the neuronal soma (*find spots*).
925 Individual spots were identified with an upper radius threshold of 750 nm. The number of
926 RNA spots per single neuron was calculated for each smFISH channel (*calculate*
927 *properties*). Last, all DAPI+ nuclei were identified across the given field for use in brain
928 region segmentation (see anatomical annotation below).
929
930 **ii. Quantification of astrocyte gene expression *in situ*:**
931
932 **a. Segmentation of astrocytes:** Maximum intensity z-projection images were generated
933 as above. Background illumination profiles of fluorescent channels were mapped to
934 correct uneven illumination (*flatfield correction*). *Glast*+ astrocytes were segmented in
935 three steps: **1**) Supervised texture segmentation was performed at a fine scale to identify
936 *Glast*+ astrocyte cytoplasm and main processes (*find texture regions*). To train texture
937 analysis, over a hundred points were selected inside versus outside *Glast*+ cortical
938 astrocytes across several tissue sections. Astrocyte cell areas were then filtered by size to
939 remove partial cells and holes across astrocyte nuclei, which are weakly labeled by *Glast*
940 smFISH, were filled (*select region*). **2**) Astrocyte nuclei were segmented within *Glast*+
941 cell areas from Gaussian blurred DAPI images using intensity, size and contrast
942 thresholds (*find nuclei*). To remove false positive non-astrocyte nuclei that overlap with
943 astrocyte processes in z-projections, additional morphology and Glast intensity filters
944 were used (*select population*). **3**) The astrocyte cytoplasm and processes were segmented
945 around the nuclei within the cell areas using *Glast* intensity thresholding (*find*
946 *cytoplasm*).
947
948 **b. Filtering single astrocytes**: To remove overlapping astrocyte doublets and partial
949 cells, cells were filtered based on morphological (e.g. area, roundness) and intensity (e.g.
950 pixel sum over DAPI and *Glast*) properties. As shown in Supplementary Figure 11, upper
951 layer astrocytes were slightly larger consistent with previous reports (*7*) and showed
952 higher *Glast* levels than those in deep layers. This analysis identified similar numbers of
953 astrocytes across cortical areas in technical replicates (Supplementary Figure 11). Manual
954 validation of this pipeline across ten tissue sections from two biological replicates
955 covering multiple cortical areas (somatosensory, visual, auditory) showed that >90% of
956 astrocytes were correctly identified using this pipeline.
957 Given their intimate cell-cell interactions, astrocyte processes occasionally
958 overlapped with neuronal soma and other nuclei in maximum-z projection images. To
959 remove overlapping non-astrocyte nuclei, the DAPI signal in the astrocyte cytoplasm was
960 identified with intensity thresholding and subtracted from the *Glast*+ cell area (cytoDAPI
961 filtered cells, Fig 2C). To remove overlapping neurons, NEUN signal was used to
962 segment neurons as above. Astrocytes that significantly overlap with neurons were
963 discarded (>50% overlap between the astrocyte nuclei and neuronal soma). NEUN+
964 neurons were then subtracted from the cytoDAPI-filtered Glast+ cell area (Fig 2C),
965 resulting in astrocytes filtered against overlapping neurons and nuclei.
966
967 **c. RNA spot calling**: RNA spots were quantified across single astrocytes as elaborated
968 for neuron previously. In addition to the filters used above, cells that are high-expression

outliers (above 99.5% of spot counts per gene across the brain) were also filtered out. RNA spots in neurons, cytoDAPI-filtered and double-filtered astrocytes were also quantified.

iii. Data processing:

The analysis was performed on a desktop workstation with two 6-core Intel i7-4930K 3.4 GHz CPUs and 64 GB of RAM. The neuronal dataset shown in Fig 1, consisting of 10 tissue sections and ~300,000 images, was analyzed under 6 hours. The data was exported from Harmony as a single cell matrix showing cell coordinates, morphological and intensity measurements, and RNA spot counts per cell. Brain region segmentation was performed in MATLAB (described below). Data organization and plotting were done in R.

Anatomical annotation of mouse cortical layers and areas

For mapping neuron and astrocyte subtypes across the cortex, one P14 mouse brain hemisphere was sectioned along the coronal plane to generate an 8-slide series containing 10 sections each. One slide was used to map the expression of cortical layer neuron markers (Fig 1) and the remaining slides were assayed with layer astrocyte markers (Fig 3 and 6). The cortical layers and areas were annotated using NEUN and DAPI staining as well as layer neuron marker expression patterns as anatomical landmarks (see Supplementary Figure 7 for areas). The Paxinos (*P6, plates 9 to 40*) and Allen Mouse Brain ISH Atlases (*P56, sections 38 to 88*) were used as anatomical references. Cortical areas were annotated broadly across the anterior-posterior and dorso-ventral axes, grouping functionally related areas (e.g. the anterior division of the somatosensory cortex contains the mouth and limb areas). The list of cortical area abbreviations and groupings are listed under Sup Table 8.

For screening layer astrocyte gene expression (Fig 2), two P14 mouse brain hemispheres (biological replicates) were sagitally sectioned to generate 18 slides. Each slide contained 4 sections through the somatosensory cortex from each replicate, corresponding to the areas used for RNAseq profiling of layer astrocytes.

For examining cortical layers in *Satb2* cKO and *Reeler* mice (Fig 5), three coronal sections through the somatosensory barrel cortex were collected from littermate control and mutant brains (n=3 biological replicates each). Each slide contained sections from one control and one mutant brain.

To normalize the layer depth of neurons across cortical areas, we automatically measured the normalized distance between individual neurons, cortical pia and white matter using the slideSegmenter application below.

Manual segmentation of brain regions

Brain areas were manually segmented on low magnification (5X) images of DAPI/NEUN stained brain sections. These segmentation masks were overlaid on xy-coordinates of high-magnification (40X) scans to annotate single cells. The offset between 5X and 40X objectives was manually corrected by aligning DAPI+ nuclei (identified across all cells in each field at the end of segmentation pipelines). To stitch images from Harmony, draw and name segmentation masks, align low-high magnification data and perform batch segmentations, the slideSegmenter application was created to work on the MATLAB

1015 environment and made publicly available
1016 (https://bitbucket.org/alexmatlab/slidesegmenter/).
1017
1018
1019 **Identification of neuronal subtypes from smFISH data**
1020 To identify cortical neuron subtypes in an unbiased manner from smFISH data, we
1021 adopted the following workflow:
1022
1023 **i. Filtering and normalization:** For downstream analysis, neurons were 1) selected from
1024 the 8 broad cortical areas that show the full complement of layers with respect to the 4
1025 genes profiled; and 2) filtered with a minimum cumulative spot-count threshold of 20.
1026 Spot counts then had a value of 1 added and were log transformed (log(spot counts +1)).
1027
1028 **ii. Clustering**: Clustering was performed for cells from each region individually with the
1029 4 genes profiled using graph-based clustering implemented by Seurat (FindClusters
1030 function, resolution 0.5) (41). Briefly, a K-nearest neighbor graph based on Euclidean
1031 distance is constructed from the expression values for each cell. Edges between cells
1032 were weighted based on shared overlap in neighborhoods determined by Jaccard distance.
1033 Cells were iteratively grouped together with the goal of optimizing the density of links
1034 inside communities as compared to links between communities.
1035
1036 **iii. tSNE:** For visualization, t-distributed stochastic neighbor embedding (tSNE)
1037 coordinates were calculated from the expression values for each cell (independent of the
1038 clustering) using perplexity 250 with Seurat (RunTSNE function). tSNE plots were then
1039 colored by the cluster assignments derived above, gene expression values, or other
1040 features of interest.
1041
1042 **iv. Hierarchical clustering:** The mean expression profiles of each of the Seurat clusters
1043 derived from each brain region were taken and hierarchically clustered together based on
1044 Euclidean distance using Ward.D2 clustering (hclust(dist, method = "ward.D2") R
1045 function). The resulting dendrogram was then cut at height 1.9 yielding 18 groups
1046 (cutree(hc, h = 1.9) R function). These groups were manually annotated to 10 major
1047 subtypes based on high expression differences (Supplementary Figure 5) and similarity
1048 among the spatial distributions of identified groups.
1049
1050 **Cortical layer and purified layer astrocyte RNA-Seq**
1051
1052 **i. Cortical layer dissection:** *Aldh1L1-GFP+* mice were transcardially perfused at P14
1053 with ice-cold Hanks Balanced Solution (HBSS) to wash away the blood. Brains were
1054 dissected and cortical hemispheres were cut sagitally on a vibratome in ice-cold HBSS to
1055 300 μm thick sections. Sections from 8 littermate pups were pooled for each experiment
1056 (n=3 biological replicates for astrocyte purification and n=2 biological replicates for
1057 whole layer RNA extractions) and microdissected to separate upper (L2-4) and deep (L5-
1058 6) cortical layers. The L4 of the somatosensory barrel cortex, which appears as dark
1059 barrels separated by light septa under bright-field illumination, was used as an anatomical
1060 landmark for layer microdissections (Supplementary Figure 8). To prevent contamination

1061 with white matter astrocytes, the most superficial layers that contain pial and L1
1062 astrocytes, and the deep subcortical white matter that contains fibrous astrocytes were
1063 discarded. For each experiment, the dissections were completed under 90 min and the
1064 tissue was kept in ice-cold HBSS.
1065
1066 **ii. Flow cytometry**: To purify cortical layer astrocytes, tissue dissociation was performed
1067 as described previously (17). Briefly, cortical layer tissue were minced with a forceps and
1068 enzymatically dissociated with papain (20 U/mL) in dissociation buffer (glucose 22.5
1069 mM, EDTA 0.5 mM, phenol red), L-cysteine (1 mM) and DNase (125 U/mL) for 80 min
1070 at 33°C. Tissue was then washed in inhibitor solution (dissociation buffer with
1071 ovomucoid (1.0 mg/mL)) and centrifuged for 5 min at 200 g. Supernatant was discarded,
1072 the tissue was resuspended in the inhibitor buffer and mechanically disrupted using a
1073 P1000 pipette. Dissociated cells were layered onto inhibitor buffer with concentrated
1074 ovomucoid (5 mg/mL) and centrifuged 5 min at 200 g. Finally, the cell pellet was
1075 resuspended in staining medium with DAPI. *Aldh1l1-GFP*$^{+}$ and *Aldh1l1-GFP*$^{-}$ cells were
1076 sorted as previously described (8) on a BD FACS Aria II and gated on forward/side
1077 scatter, live/dead by DAPI exclusion, and GFP, using GFP and DAPI controls to set gates
1078 for each experiment (Supplementary Figure 8).
1079
1080 **iii. RNA sequencing and analysis**: Total RNA from FACS-purified cortical layer
1081 astrocytes and whole cortical layers was extracted with Trizol LS (Invitrogen) and
1082 purified using the RNeasy Kit (Qiagen). cDNA was generated from full-length RNA
1083 using the NuGEN RNA-Seq V2 kit that employs the single primer isothermal
1084 amplification method to deplete ribosomal RNA, and sheared by Covaris to yield uniform
1085 size fragments. RNASeq libraries were generated using the NuGen Ultralow kit for
1086 adapters, barcoding, and amplification and purified using the Agencourt XP magnetic
1087 beads, quality controlled with an Agilent bioanalyzer, and quantified by qPCR.
1088 Five libraries were pooled per lane across three lanes for single end (SE50)
1089 sequencing on a HiSeq 4000. Read quality was assessed using FastQC (version 0.11.4)
1090 and 5 nucleotides at the 5' end were trimmed. 45 nucleotide long reads were aligned to
1091 the mouse reference genome (Ensembl GRCm38) using TopHat2 (version 2.0.11) (42).
1092 The multiple hit parameter was (-g) was set to 1 to exclude reads with multiple genomic
1093 alignments. On average, 68 million reads were uniquely mapped to each sample (range
1094 59-87M). Read counts per gene were calculated using SAMtoools (version 0.1.19) (43)
1095 and HTSeq (version 0.6.1p1) with default parameters (44). DESeq2 (45) was used to
1096 detect differentially expressed genes amongst upper and deep layer astrocytes (n=3
1097 biological replicates) and whole cortical layers (n=2 replicates). Purified deep layer
1098 astrocytes showed low levels of contaminating oligodendrocyte marker gene expression
1099 (e.g. MBP); these genes were excluded from analysis using a mild astrocyte-specific
1100 expression threshold (astrocyte vs whole layer expression > 0.1). To identify the top
1101 differentially expressed genes between upper and deep gray matter astrocytes, an
1102 expression threshold of 5 FPKM was used with a false-discovery rate (FDR) < 0.05. The
1103 resulting list of 159 differentially expressed layer astrocyte genes is provided in
1104 Supplementary Table 9.
1105
1106

1107 **Spatial reconstruction of astrocyte layer heterogeneity:**
1108
1109 **i. Pre-processing of astrocyte scRNA-seq data:** Single cell RNA-sequencing data of
1110 P56 cortical astrocytes was obtained from Batiuk and Martirosyan et al (23) which used a
1111 modified version of Smart-seq2 protocol. Number of unique molecules (UMI) was
1112 estimated from raw read counts with Census method from Monocle package (46). UMI
1113 version of the data was normalised with size factors according to standard scater
1114 workflow (47).
1115
1116 **ii. Spatial reconstruction:** We reconstructed spatial profiles of genes measured with
1117 scRNA-seq using a published method (22). Implementation of the method in Matlab was
1118 provided by Shalev Itzkovitz, see original publication for the description of
1119 implementation details. Here we summarize the method.
1120
1121 **a. Constructing the prior on the number of reads for markers genes in scRNA-seq**
1122 **cells**: Molecule counts of genes from the processing step of image analysis were fit by a
1123 Gamma distribution for each gene i and spatial bin j.
1124
1125 (1) $sc_{ij} = Gamma(os_{ij}, or_{ij})$
1126 Where sc is spot counts, os is observed shape of gamma distribution, and or is
1127 observed rate.
1128 To match this distribution to the statistical properties of scRNA-seq data and produce the
1129 prior distribution of scRNA-seq reads for each gene i and spatial bin j the following
1130 correction was applied for each scRNA-seq cell c:
1131
1132 (2) $f_c = (\frac{s_c}{sm})$
1133 Where f is the distribution rescaling factor, s is scRNA-seq sampling constant, sm is
1134 smFISH sampling constant
1135 (3) $p_{ijc} = \frac{f_c}{(\mathrm{or}_{ij} + f_c)}$
1136 where p is negative binomial probability
1137 (4) $prc_{ijc} = NegativeBinomial(r = os_{ij}, p_{ijc})$
1138 Where prc is the prior on expected single cell RNA-seq reads for each marker gene,
1139 bin and cell
1140
1141 smFISH sampling factor reflects the expected proportion of the total number of
1142 transcripts in a cell captured in smFISH images (we used 0.15). scRNA-seq sampling
1143 factor reflects the proportion of total number of transcripts (approximated as 180000)
1144 captured by scRNA-seq for each cell measured. The total number of transcripts was
1145 chosen based on Halpern et al that examined haptocytes: 180000 is roughly 1/4 of
1146 hepatocyte molecule number: 1/2 to account to tetraploidy and 1/2 to account for smaller
1147 cell size of astrocytes. In practice, sampling levels are computed for 8 bins of cells with
1148 similar sampling levels (e.g. 0-0.1%, 0.1%-0.3% …) to speed up the Monte Carlo

1149 sampling used to construct the prior. We used 5 spatial bins (j) for equally sized intervals
1150 of cortical depth.
1151
1152 **b. Using the prior to assign cells to spatial bins**. To find the probability of each
1153 scRNA-seq cell coming from spatial bin j given the expression of spatial markers we
1154 used Bayes' formula:
1155

1156 $$(5)\ P(Bin_j \mid gene_i\ reads) = \frac{P(gene_i\ reads \mid prc_{ijc}) * P(cells\ in\ Bin_j)}{\sum_{bin\ j=1}^{bin\ 5} P(gene_i\ reads \mid prc_{ijc}) * P(cells\ in\ Bin_j)}$$

1157
1158 where $P(cells\ in\ Bin_j)$ is the proportion of cells in spatial bin j, and the prc prior on
1159 reads comes from Equation (4).
1160
1161 The method assumes that expression of different genes is independent and the probability
1162 of assignment to spatial bins is computed as:
1163

1164 $$(6)\ P(Bin_j \mid gene_{i-n}\ reads) = \prod_{gene\ i=1}^{gene\ n} P(Bin_j \mid gene_i\ reads)$$

1165
1166 This result in posterior probability matrix of cells by bins. A probability weight matrix is
1167 obtained by normalizing the column sums to one
1168
1169 **c. Reconstructing expression of genes across spatial bins.** The probability weight
1170 matrix, P and the scRNA-seq expression matrix, E, are combined to obtain spatial
1171 profiles as a genes by bins matrix as a weighted average (Supplementary Tables 10 and
1172 11).
1173

1174 $$(7)\ S = E * P$$

1175
1176 **d. Identifying significantly zonated profiles.** To find which profiles are significantly
1177 zonated we assigned cells to the spatial bin with highest posterior probability. The
1178 Kruskal–Wallis test was used to find which genes are significantly different between cells
1179 assigned to different spatial bins.
1180
1181 **e. Leave one out validation.** To benchmark the model we used leave one out analysis
1182 (Supplementary Figure 18). We used 15 genes with smFISH profiles to reconstruct one
1183 left-out gene and compared the reconstructed profile to the average expression observed
1184 with smFISH. To measure the similarity of predicted and observed profiles we used
1185 Kullback–Leibler divergence base on the relative differences in expression between
1186 spatial bins rather than absolute number of genes per bin.
1187

1188 **Manual immunohistochemistry:**

1189
1190 Cryosections from *Aldh1L1-GFP* mice were manually stained for GFP, astrocyte and
1191 neuron marker antibodies. Samples were subjected to heat-induced antigen retrieval in
1192 10mM sodium citrate (pH 6) buffer for 5 min at 75°C, then permeabilized and blocked in
1193 10% goat serum in 1X PBS with 0.1% Triton X-100 (PBST) for 1 h. Primary antibodies

1194 were diluted in the blocking solution and incubated O/N at 4°C. After multiple PBST
1195 washes, samples were incubated in secondary antibodies and DAPI diluted in blocking
1196 solution for 1 h. After PBS and dH_2O washes, samples were mounted using
1197 Fluoromount-G (SouthernBiotech). The following primary antibodies were used: chicken
1198 GFP (GFP-1020, Aves, 1:2000), mouse NEUN (MAB377, Millipore, 1:500), rabbit GS
1199 (G2781, Sigma, 1:2000) and rabbit pSmad (13820S, Cell Signaling, 1:200). Goat
1200 secondary antibodies conjugated to Alexa fluorophores (Molecular Probes) were used for
1201 labeling. The *Aldh1L1-GFP* samples were imaged on a Leica TCS SPE laser confocal
1202 microscope with a 40x oil immersion objective.
1203
1204 **Statistics**:
1205
1206 The violin plots in Figure 1D, Supplementary Figures 2B-C, 11C and 21C and the box
1207 plots in Supplementary Figure 20A-B were plotted to show the data distribution as
1208 follows. The lower, middle and upper hinges correspond to 25, 50 (median) and 75
1209 percentiles. The whiskers show data points within 1.5 times the interquartile range
1210 (distance between the first and the third quartiles) of the 25 and 50 percentiles.
1211
1212 No statistical methods were used to pre-determine sample sizes but our sample sizes are
1213 similar to those reported previously (10-12). smFISH data (RNA spot counts per cell) and
1214 scRNAseq data distribution was assumed to be Gamma-Poisson but this was not formally
1215 tested.
1216
1217 For layer switch experiments (Figure 5), data collection and analysis were performed
1218 blind to the genotypes of the animals. Animals were randomly allocated to experimental
1219 groups. Two-tailed Student's t-tests were used to assess statistical significance in layer
1220 switch experiments (Figure 5).
1221
1222 **Data Availability:**
1223
1224 The raw bulk RNA-sequencing data are available at the Gene Expression Omnibus
1225 (GEO) under the accession code GSE140822. The single cell RNA-sequencing data will
1226 be made available under https://holt-sc.glialab.org/sc/. Other data are available as
1227 Supplementary Materials or from the corresponding author upon request.
1228
1229 **Code Availability:**
1230
1231 The code for spatial reconstruction of single cell astrocyte RNA-sequencing can be found
1232 at https://github.com/vitkl/cortical_astrocyte_mapping. The SlideSegmenter code is available
1233 at https://bitbucket.org/alexmatlab/slidesegmenter/src/master/. The Harmony image
1234 analysis scripts are provided as Supplementary Materials. Other code is available upon
1235 request.
1236
1237
1238
1239

Supplementary Methods References:

37. Gong S, Zheng C, Doughty ML, Losos K, Didkovsky N, Schambra UB, et al. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature. 2003 Oct 30;425(6961):917–25.

38. Gorski JA, Talley T, Qiu M, Puelles L, Rubenstein JLR, Jones KR. Cortical Excitatory Neurons and Glia, But Not GABAergic Neurons, Are Produced in the Emx1-Expressing Lineage. Journal of Neuroscience. Society for Neuroscience; 2002 Aug 1;22(15):6309–14.

39. Wang F, Flanagan J, Su N, Wang L-C, Bui S, Nielson A, et al. RNAscope. The Journal of Molecular Diagnostics. 2012 Jan;14(1):22–9.

40. Kerstens HM, Poddighe PJ, Hanselaar AG. A novel in situ hybridization signal amplification method based on the deposition of biotinylated tyramine. J Histochem Cytochem. 1995 Aug 4;43(4):347–52.

41. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. Nature Publishing Group; 2018 Jun;36(5):411–20.

42. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. BioMed Central; 2013 Apr 1;14(4):R36.

43. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009 Jul 15;25(14):1754–60. PMCID: PMC2705234

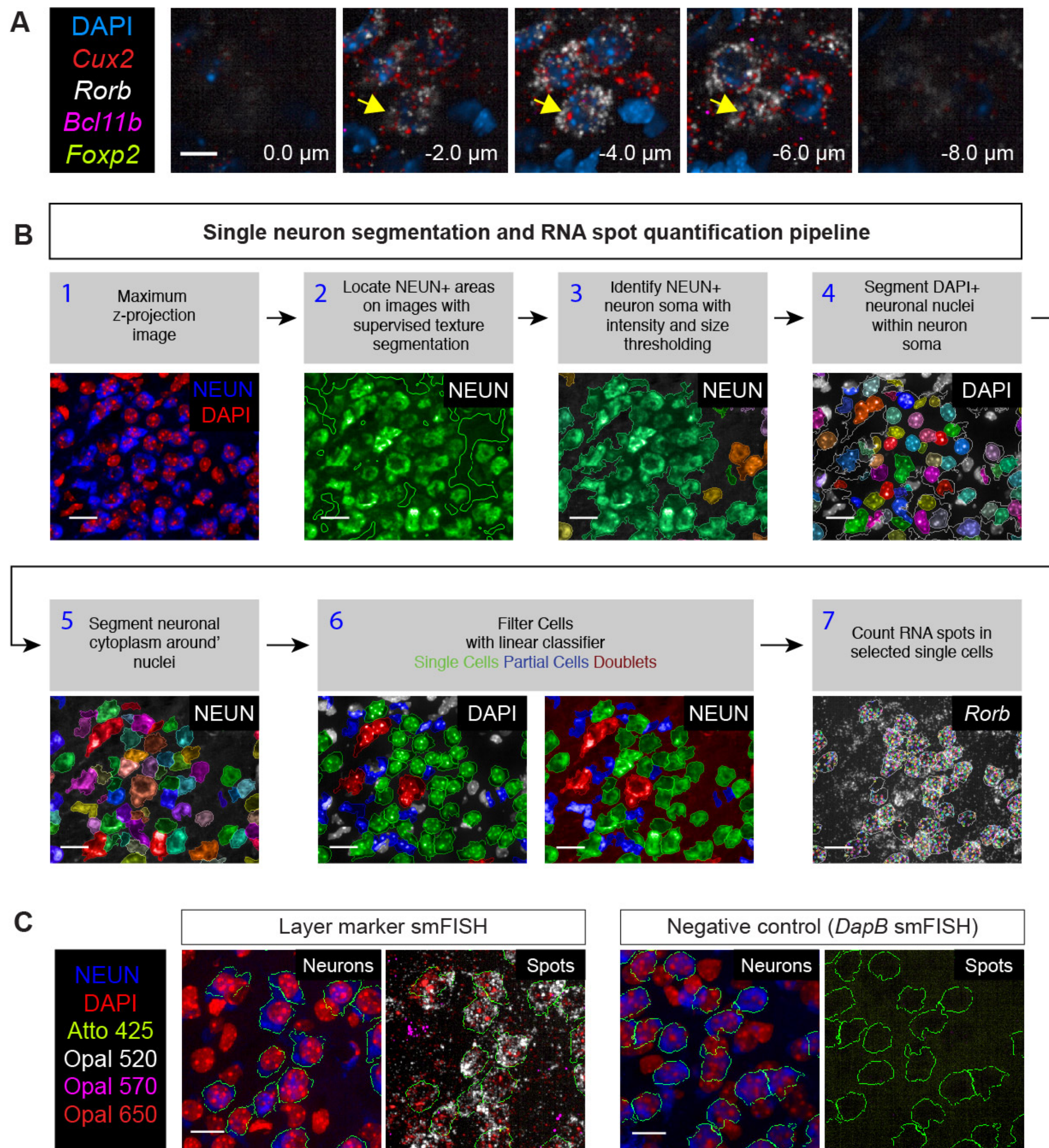
44. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. Oxford University Press; 2015 Jan 15;31(2):166–9.

45. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. BioMed Central; 2014 Dec 1;15(12):550.

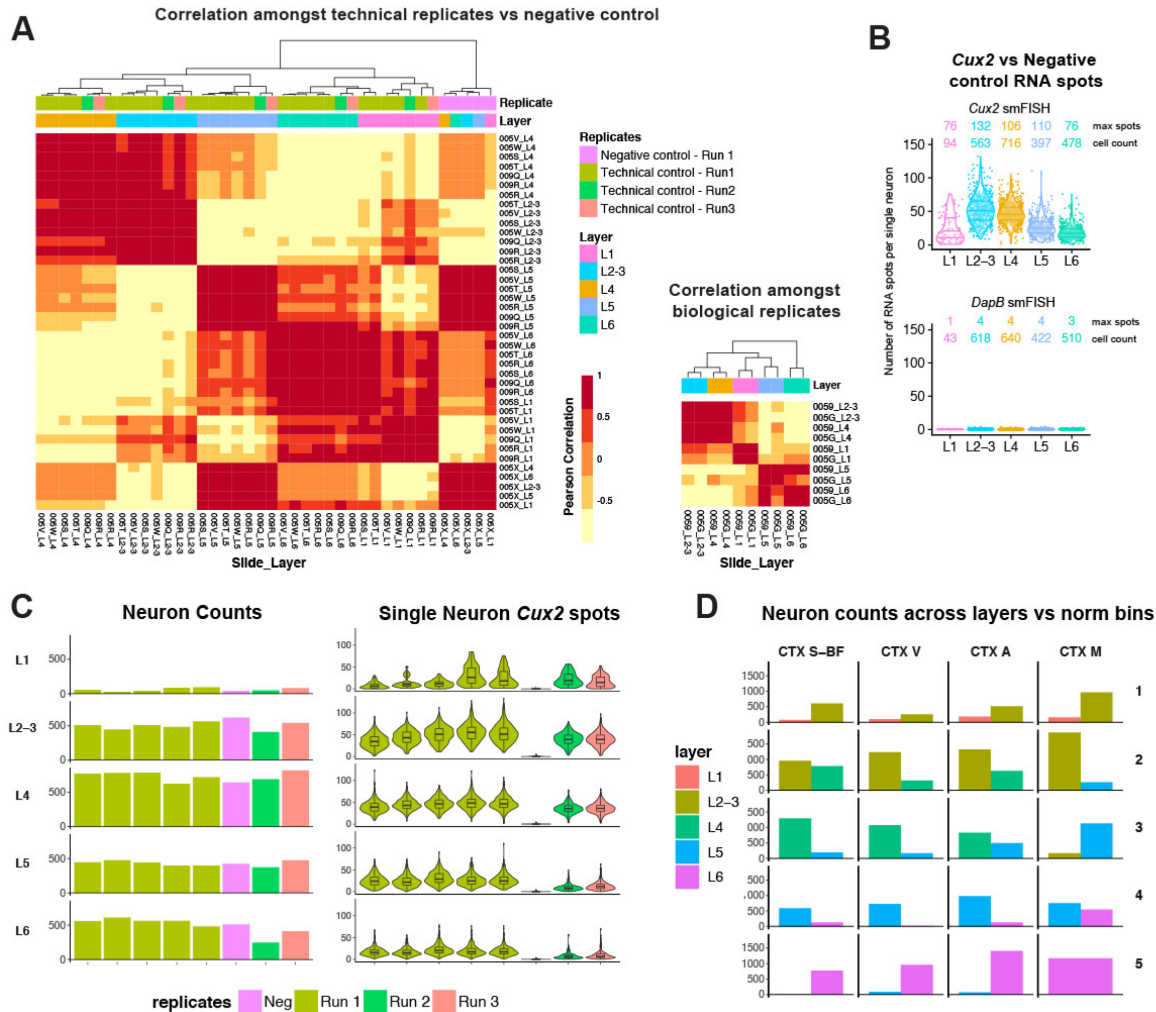
46. Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C. Single-cell mRNA quantification and differential analysis with Census. Nat Methods. 2017 Jan 23. PMCID: PMC5330805

47. Lun ATL, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Research. Faculty of 1000 Ltd; 2016;5:2122. PMCID: PMC5112579

1278

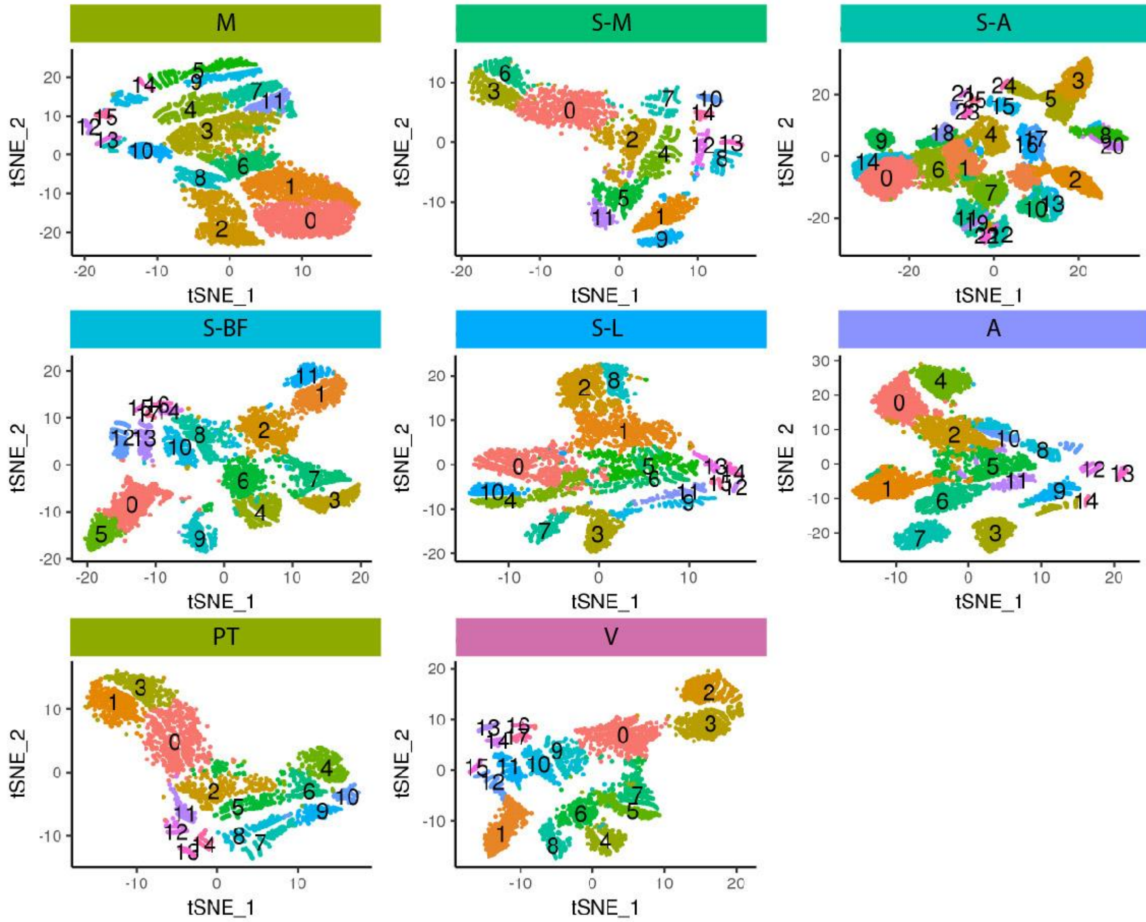


1279
1280
1281 **Supplementary Figure 1: Single neuron image analysis and gene expression quantification**
1282 **pipeline.**
1283 **A)** Individual 40X z-planes throughout *Rorb*$^+$ L4 neurons. Arrow indicates a single neuron across
1284 multiple z-positions. Nuclei are marked by DAPI. For neuronal segmentation, the z-stack is
1285 collapsed into a single plane via a maximum intensity projection.
1286 **B)** Segmentation of NEUN+ neurons and quantification of gene expression in single neurons.
1287 **C)** RNAScope smFISH assay shows high signal-to-noise ratio. The background signal is assessed
1288 by comparing smFISH against layer neuron markers to bacterial *DapB* transcript negative control
1289 (targeted with four different probes in different channels). *DapB* smFISH shows little to no signal
1290 on mouse tissue sections, as expected. Quantification shown in Supplementary Figure 2. *n*=12
1291 mice independently assayed, 3-10 tissue sections per replicate imaged.
1292 Scalebars: (A) 10 µm, (B,C) 20 µm
1293

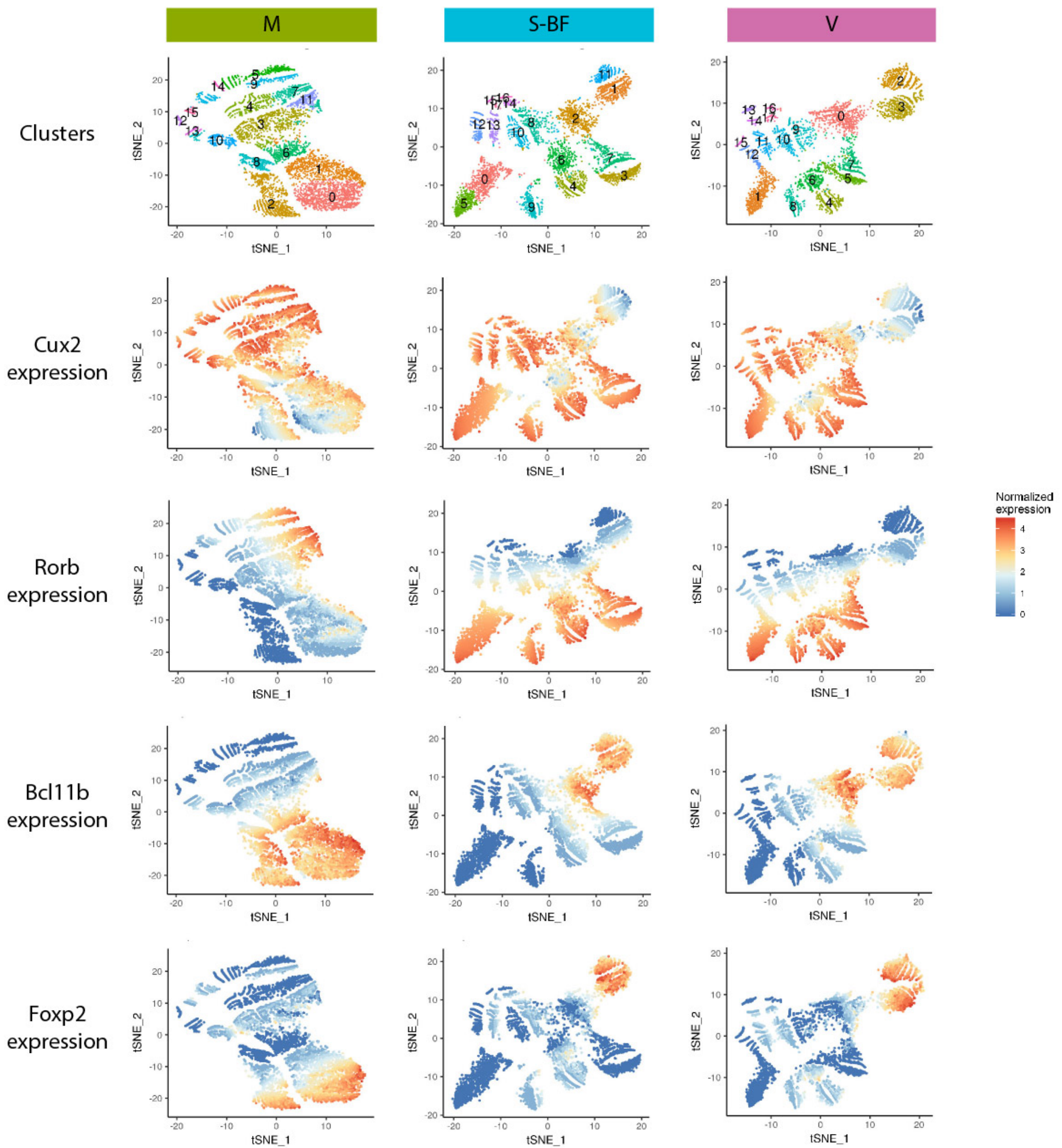


Supplementary Figure 2: Reproducibility of single neuron gene expression measurements. **A)** Neuronal gene expression measurements are highly consistent across technical and biological replicates. (Left) Heatmap showing the pearson correlation values across expression profiles of technical replicates and the negative control. Technical replicates are consecutive P14 brain sections on different slides assayed for smFISH against 4 layer markers. Negative control was assayed for smFISH against bacterial *DapB* transcripts. To assess technical variation within a staining run, multiple replicates were assayed simultaneously on the BOND RX (Run 1). To assess batch effects, replicates were assayed on different days using different consumable reagent kits (Runs 2 and 3). To calculate the expression profiles of replicates per cortical layer, single neuron RNA spot counts for 4 layer markers (*Cux2, Rorb, Bcl11b, Foxp2*) are averaged and log-transformed across each layer (L1, L2-3, L4, L5, L6) in the barrel cortex. The replicate-layer expression profiles were then hierarchically clustered. As expected, cortical layers clustered across technical replicates from the same staining run as well as different batches, indicating reproducibility, while negative control layers formed a distinct cluster. (Right) Heatmap showing the pearson correlation values across expression profiles of biological replicates. Biological replicates are barrel cortex sections from two littermate P14 animals. Average layer expression profiles were calculated as described above. As expected, cortical layers from biological replicates clustered together.

1315 **B)** Quantification of single neuron Cux2 vs negative control *DapB* expression across cortical
1316 layers. The background signal of RNAScope smFISH, assessed by the numbers of *DapB* spots
1317 per cell, is 0 to 2 spots per cell. Dot plot and the violin plot show single cell data. n=2 mice
1318 independently assayed, 3 tissue sections imaged.
1319 **C)** (Left) Quantification of neurons across cortical layers in the barrel cortex amongst replicates.
1320 Similar number of neurons are detected based on NEUN IHC and DAPI staining across technical
1321 and negative replicates. (Right) Quantification of single neuron *Cux2* expression across technical
1322 and negative replicate with violin and boxplots. Negative control shows the quantification of
1323 *DapB* expression in the same probe channel used for *Cux2* (Opal 650). The range of single cell
1324 *Cux2* expression across upper layers is highly consistent amongst technical replicates from the
1325 same staining run. Slightly lower expression is observed on batch replicate controls, yet the upper
1326 layer enrichment of Cux2 is highly similar. *n*=1 mouse, 3 tissue sections independently imaged.
1327 **D)** Distribution of layer neurons across normalized cortical depth bins. Cortical layers were
1328 manually annotated across four cortical areas, then neurons were sorted into five bins across
1329 normalized cortical depth between the pial surface and the white matter. The distributions of layer
1330 neurons to depth bins is largely similar across different cortical areas, with the exception of motor
1331 cortex that lacks a prominent layer 4.
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360

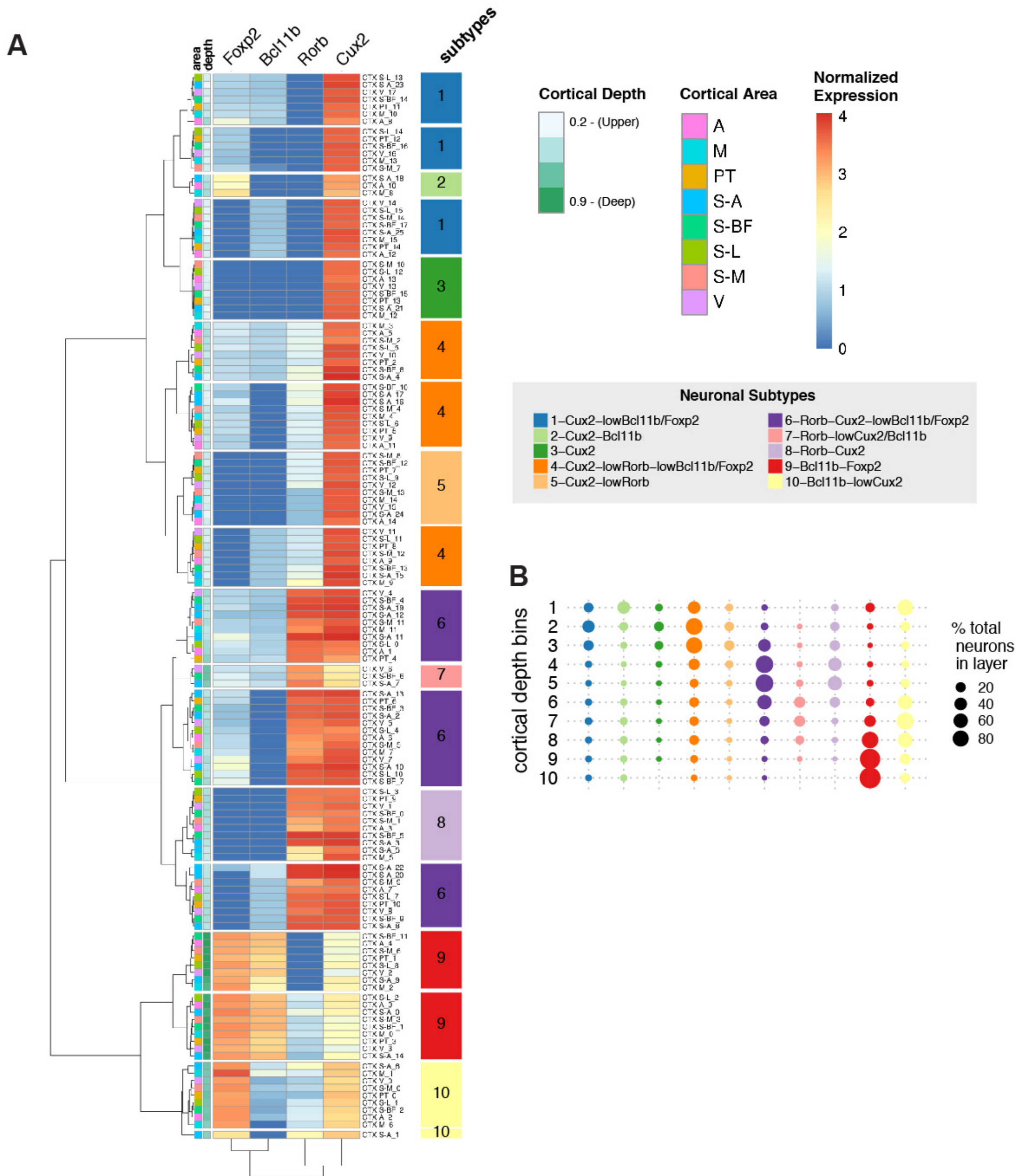


1361
1362
1363 **Supplementary Figure 3: Clustering of single neurons from different cortical areas.**
1364 **A)** To identify neuronal subtypes, clustering of single neuron gene expression was performed
1365 within each cortical area individually with 4 layer markers profiled. Single neuron tSNE
1366 coordinates were calculated from the expression profiles and tSNE plots were colored according
1367 to cluster assignments from above. See Supplementary Methods for details.
1368 Abbreviations: M, motor, S-A, anterior- somatosensory, S-M, medial-somatosensory, S-BF,
1369 somatosensory barrel, S-L, somatosensory-lateral, PT, parietal, A, auditory, V, visual.
1370
1371
1372
1373
1374
1375
1376

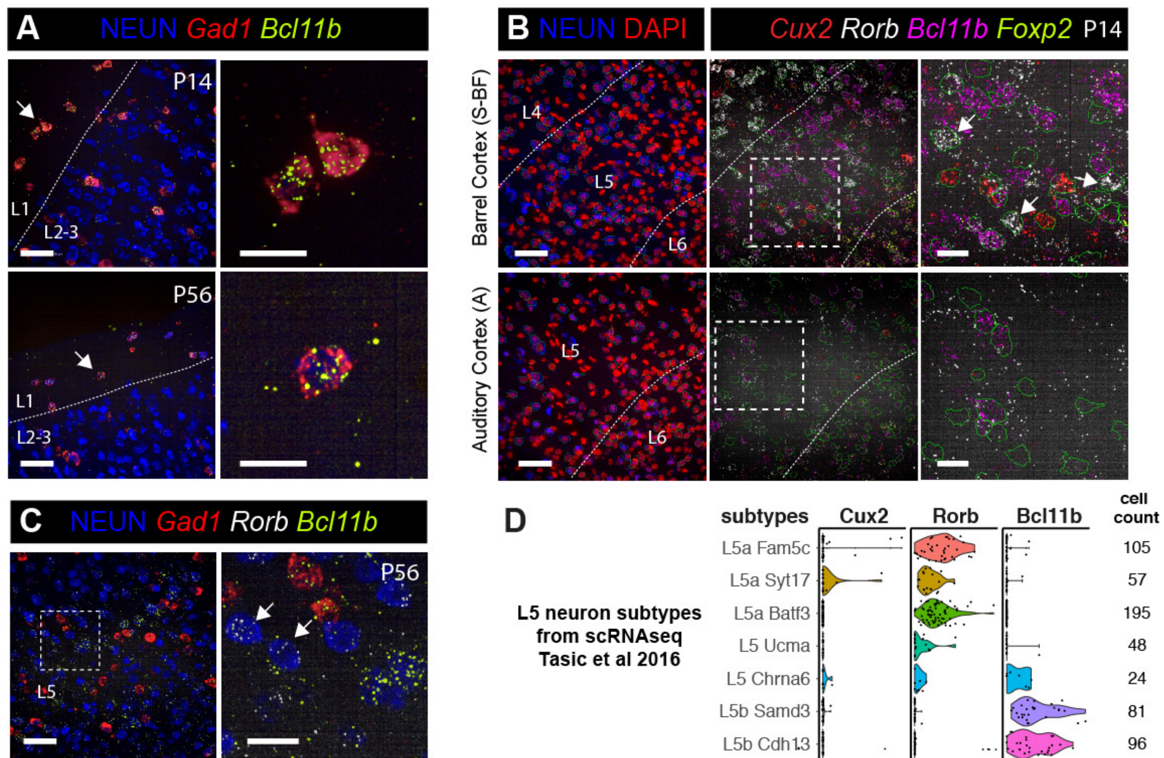


1377
1378
1379 **Supplementary Figure 4: Single neuron clusters are distinguished according to layer gene**
1380 **expression patterns.**
1381 **A)** tSNE plots across three different cortical areas (M, motor, S-BF, somatosensory barrel, V,
1382 visual) colored according to expression values.
1383
1384
1385
1386
1387

1388
1389

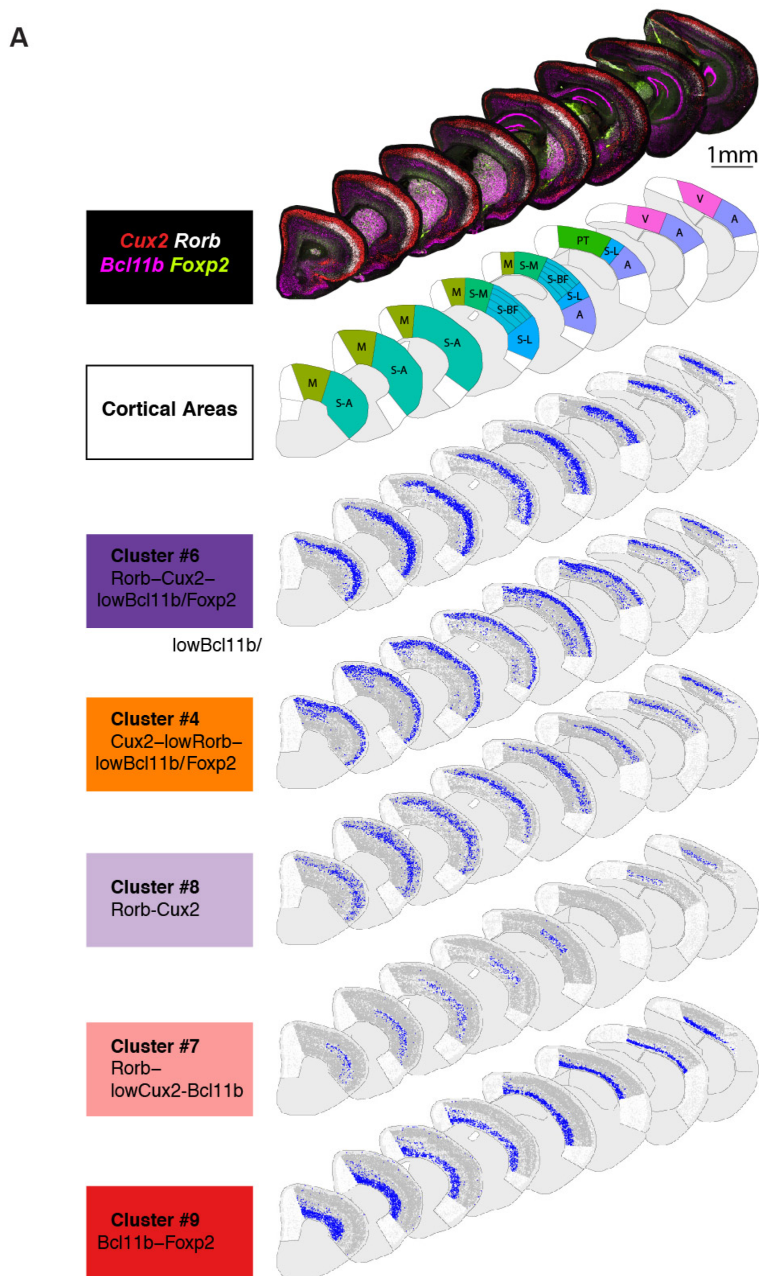


1390
1391
1392 **Supplementary Figure 5: Hierarchical clustering distinguishes neuronal subtypes.**
1393 **A)** Hierarchical clustering of clusters from 8 cortical areas (Supplementary Figure 3) according to
1394 mean expression profiles of each group. The clustering yielded 18 groups that were manually
1395 annotated to 10 major subtypes based on high expression differences and spatial distribution
1396 across the cortex. **B**) Neuronal subtypes plotted across 10 cortical depth bins.
1397
1398

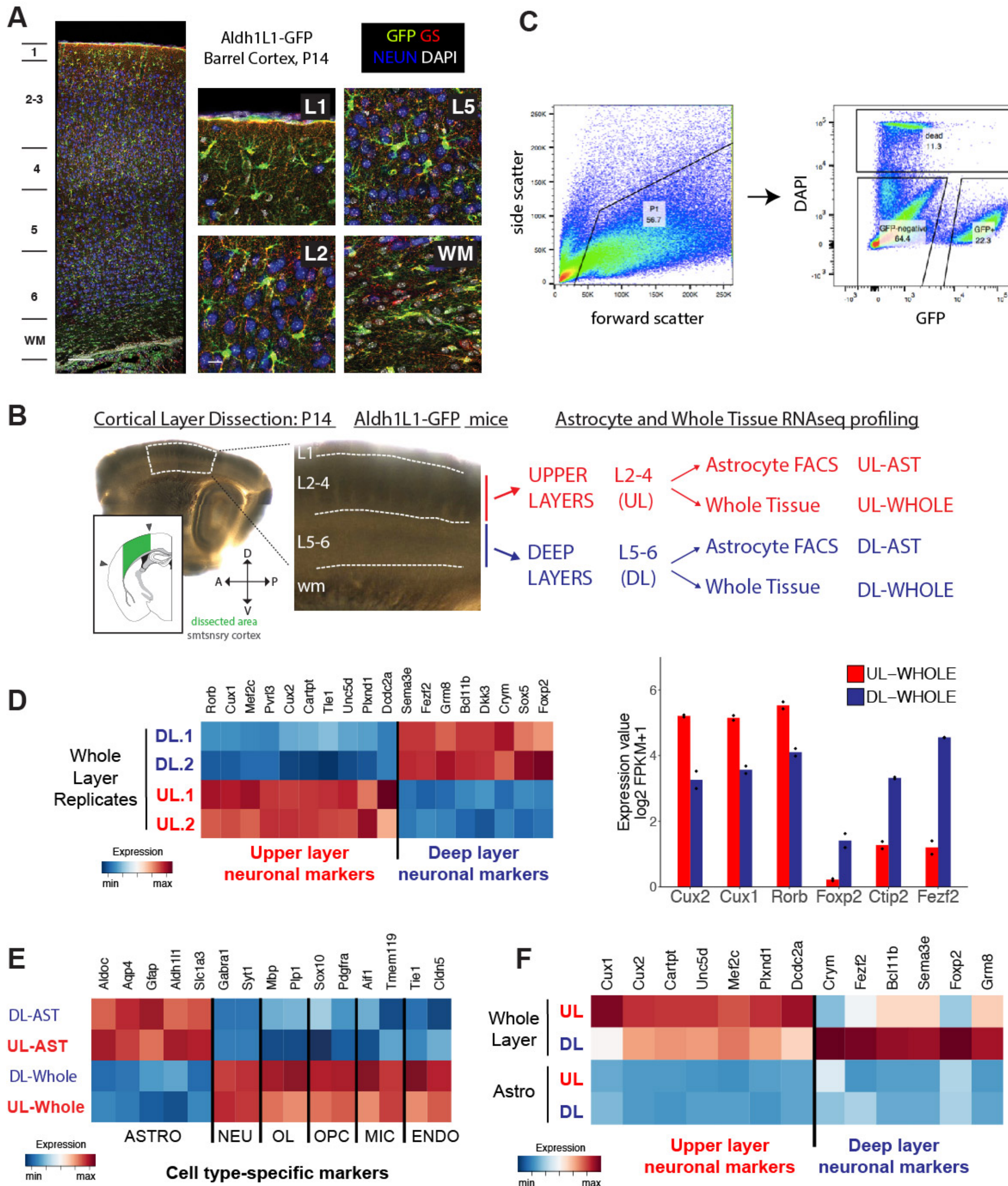


1399
1400
1401 **Supplementary Figure 6: smFISH images demonstrating *Cux2*^{mid}*Bcl11b*^{mid}-L1 and**
1402 ***Rorb*^{high}*Bcl11b*^{low}-L5 neuronal populations.**
1403 **A)** Neurons that co-express Cux2 and Bcl11b (cluster #2, arrows) are observed in L1. These
1404 interneurons, based on high *Gad1* expression, are present at P14 and are maintained into
1405 adulthood at P56. Right panels show higher magnification views of indicated neurons.
1406 **B)** Area enrichment of novel *Rorb*$^{+}$ L5 subpopulations. *Rorb*^{high}*Cux*^{mid}*Bcl11b*^{low} neurons (cluster
1407 #7, arrows) are observed in the L5 of the somatosensory barrel cortex, but are absent from the
1408 auditory cortex at P14. The higher magnification view of L5 areas outlined in dashed boxes
1409 shown on the right panels.
1410 **C)** *Rorb*^{high}*Bcl11b*^{low} neurons are maintained into adulthood at P56. *n*=2 mice independently
1411 assayed, 3 tissue sections imaged (A-C).
1412 **D)** Validation of *Rorb*^{high}*Bcl11b*^{low}-L5 subtypes in a published single neuron transcriptomics
1413 datasets. Dot plots of single neurons and violin plots showing the segregation of *Rorb* and *Bcl11b*
1414 expression amongst molecular subtypes of L5 neurons in the adult visual cortex identified by
1415 Tasic et al. Subtypes were named according to the nomenclature in the referenced study. Number
1416 of cells observed in each class is shown.
1417 Scalebars: (low magnification panels) 50 µm, (higher magnification panels) 20 µm.
1418
1419
1420
1421
1422
1423
1424
1425
1426

1427



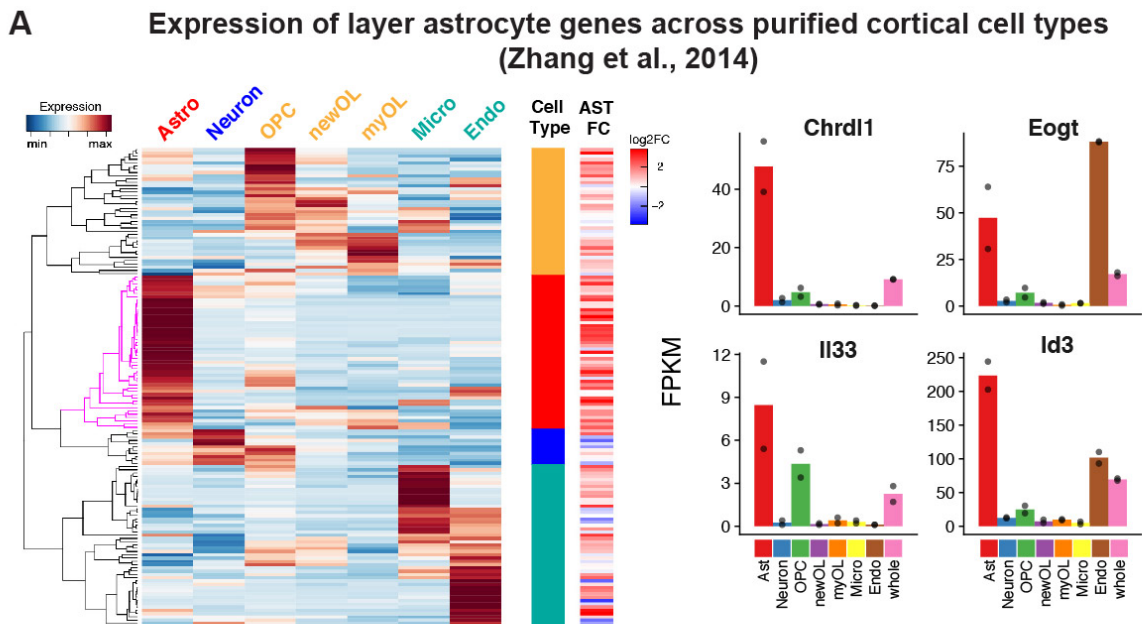
1428
1429 **Supplementary Figure 7: Maps showing the single cell level distribution of select neuronal**
1430 **subtypes.**
1431 **A)** (First row) Low magnification images of P14 hemisections from eight select anatomical levels
1432 assayed for neuronal layer marker smFISH. (Second row) Maps of broad cortical areas included
1433 in neuronal subtype analysis. (Bottom rows) Maps showing the spatial distribution of individual
1434 neuronal subtype clusters. *n*=1 mouse, 10 sections independently imaged.
1435 Scalebar: 1 mm
1436
1437
1438

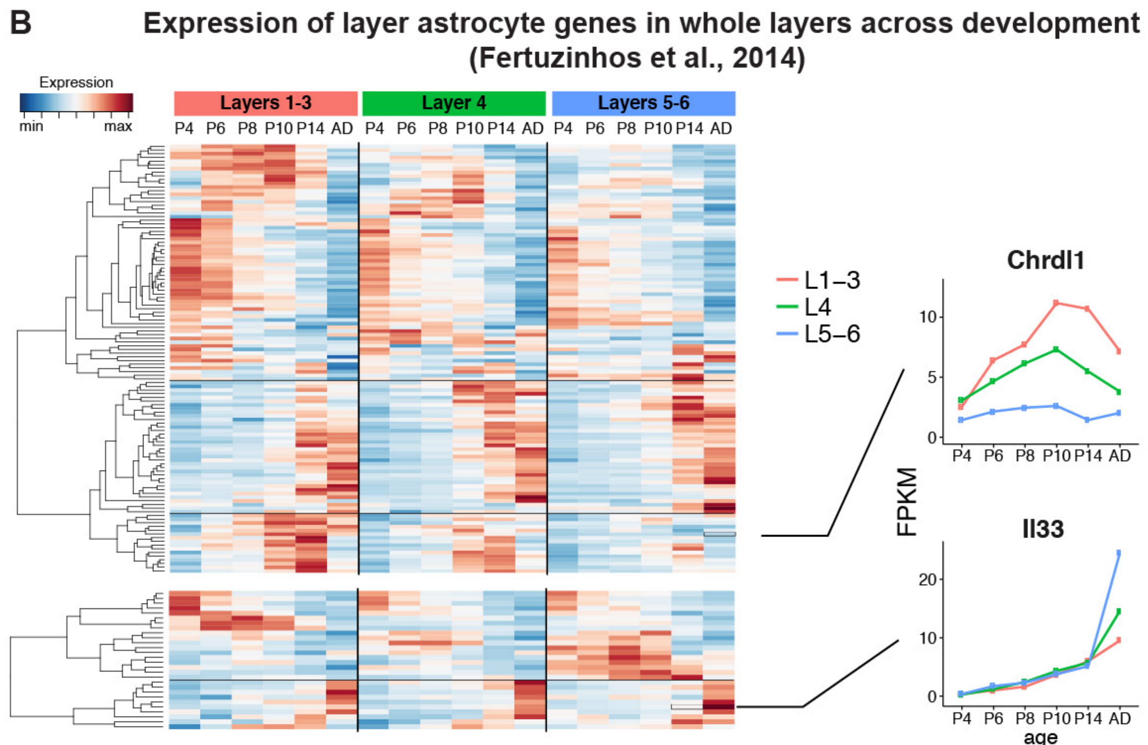


Supplementary Figure 8: Purification and RNAseq expression profiling of upper and deep layer astrocytes.
A) *Aldh11L1-GFP* labeling marks astrocytes across cortical layers and excludes neurons. Confocal images of antibody staining against GFP, NEUN (neuronal marker) and Glutamine Synthetase (GS, astrocyte marker) in the barrel cortex at P14. *Aldh1L1-GFP* labeling marks astrocytes throughout cortical gray matter, white matter and L1-subpia. *n*=2 mice independently assayed, 2 tissue sections imaged.
B) Schematic summarizing layer astrocyte purification and gene expression profiling. (Left) Bright-field images of a sagittal P14 mouse brain slice showing the outline of the layer microdissection in the somatosensory cortex (white dashed lines & also marked green in small

1451 diagram). L4 barrels were used as an anatomical landmark. (Right) FACS-purification and
1452 RNAseq profiling strategy.
1453 **C)** *Aldh1L1-GFP+* astrocyte isolation by FACS using scatter gates, doublet exclusion (not
1454 shown) and sorting for GFP+ cells with dead cell exclusion by DAPI staining.
1455 **D)** RNAseq expression pattern of known layer neuron markers across whole layer tissue, shown
1456 with an expression heatmap and bar-plots, validates the layer microdissection (n=2 biological
1457 replicates).
1458 **E)** Expression pattern of cell type-specific markers confirms the successful purification of
1459 astrocytes.
1460 **F)** The expression of known neuronal layer marker genes does not distinguish layer astrocytes.
1461 Scalebars: (A, large panel) 100 µm, (A, small panels) 25 µm.
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497

1498

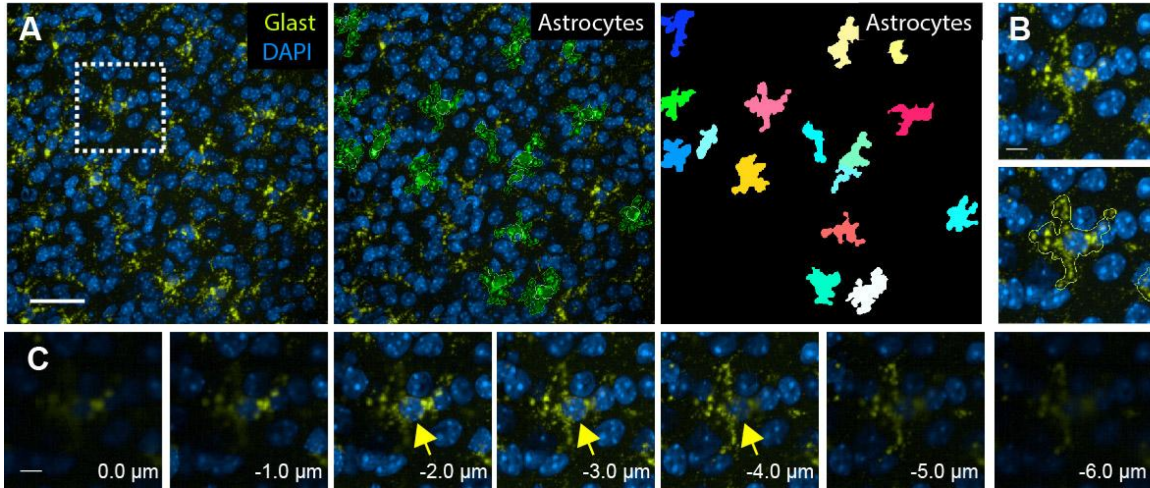




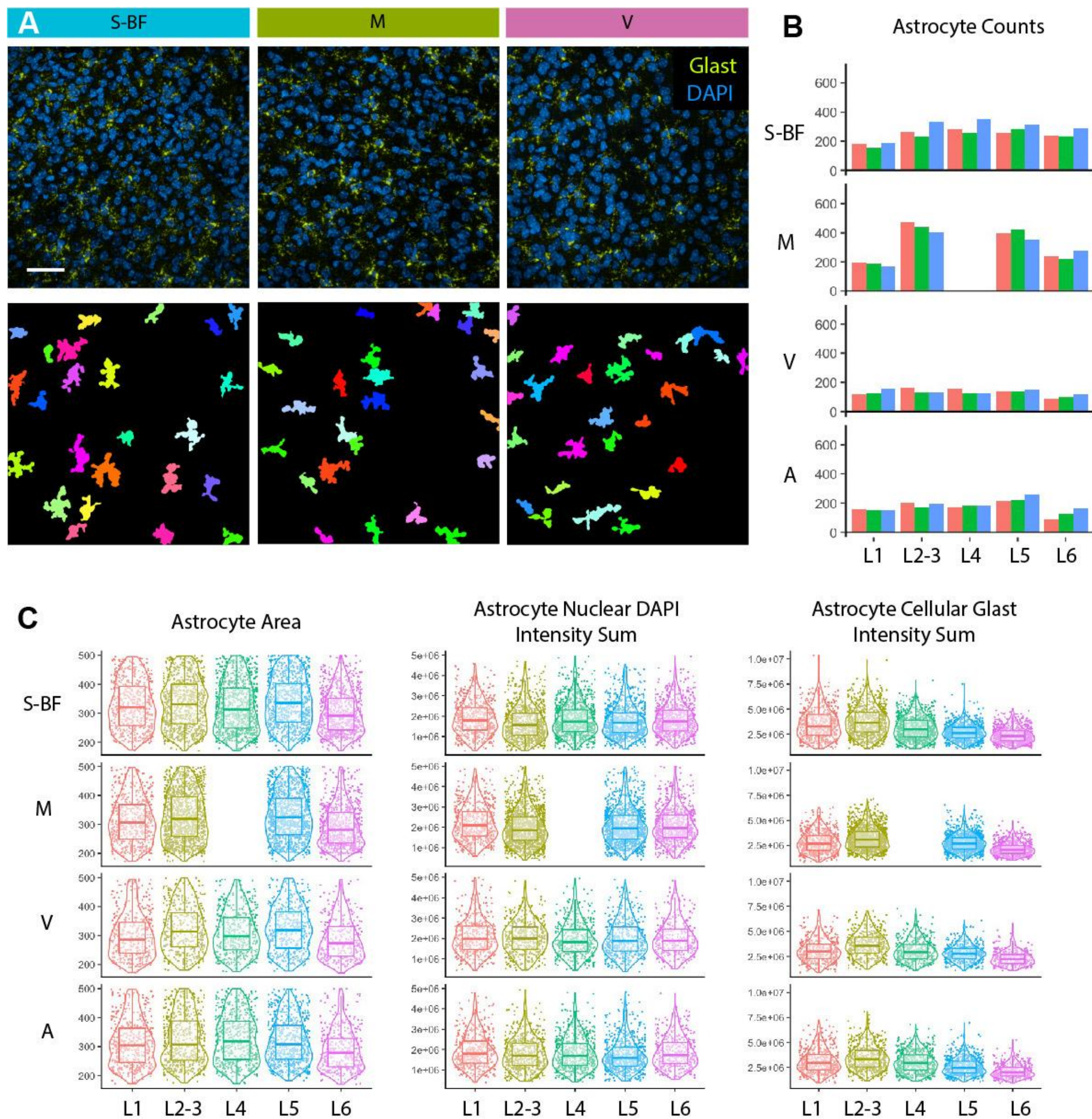
1499
1500
1501 **Supplementary Figure 9: Candidate layer astrocyte genes express in laminar and astrocyte-**
1502 **enriched manner across published cortical transcriptome dataset.**
1503 **A)** The expression pattern of 159 genes differentially expressed across upper and deep layer
1504 astrocytes across purified cortical cell types. Zhang et al performed RNA-seq analysis of purified
1505 mouse cortical astrocytes, neurons, oligodendrocyte precursor cells (OPCs), newly differentiated
1506 oligodendrocytes (newOL), myelinating oligodendrocytes (myOL), microglia, and endothelial

1507 cells (Zhang et al., J Neurosci. 2014, *n* =2 biological replciates per cell type). (Left) Heatmap
1508 shows that many candidate layer astrocyte genes show expression in astrocytes in Zhang et al's
1509 dataset. Many genes have enriched expression in astrocytes, others also express in additional cell
1510 types. (Right) Bar plots showing expression of select candidate layer astrocyte genes across cell
1511 types. *Chrdl1* expression is highly enriched in astrocytes while *Eogt* and *Id3* also show
1512 expression on endothelial cells (validated by smFISH, data not shown).
1513 **B)** The expression of 163 candidate layer astrocyte genes across whole cortical layer tissue
1514 throughout postnatal development and adulthood. Fertuzinhos et al performed RNA-seq analysis
1515 manually dissected upper (L1-3), mid (L4) and deep (L5-6) cortical layers at different timepoints
1516 during postnatal life and adulthood (Fertuzinhos et al., Cell Rep. 2014). (Left) Heatmap shows
1517 that many candidate layer astrocyte genes show laminar and developmentally regulated gene
1518 expression in Fertuzinhos et al's dataset. Top heatmap shows upper layer astrocyte enriched
1519 genes while the bottom heatmap shows deep layer astrocyte enriched genes. Many genes are
1520 upregulated during early postnatal life, consistent with the commencement and progression of
1521 cortical astrogenesis after birth (24) (clusters marked in magenta on the dendrogram). (Right)
1522 Most layer astrocyte candidate genes show temporally regulated expression throughput postnatal
1523 life. *Chrdl1* expression peaks during the second postnatal week yet persists into adulthood. *Il33*
1524 expression increases into adulthood.
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554

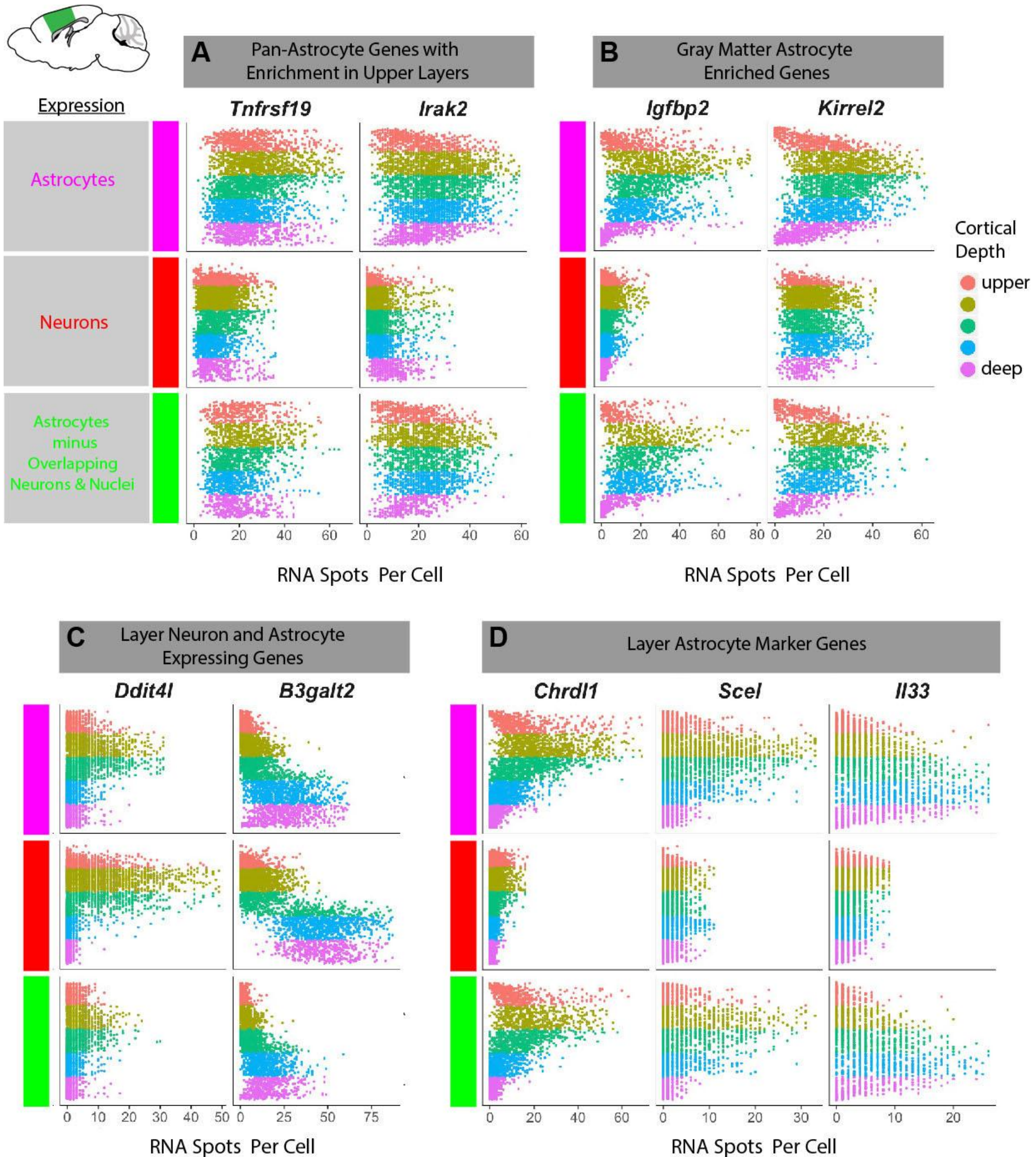
1555
1556
1557
1558
1559
1560
1561



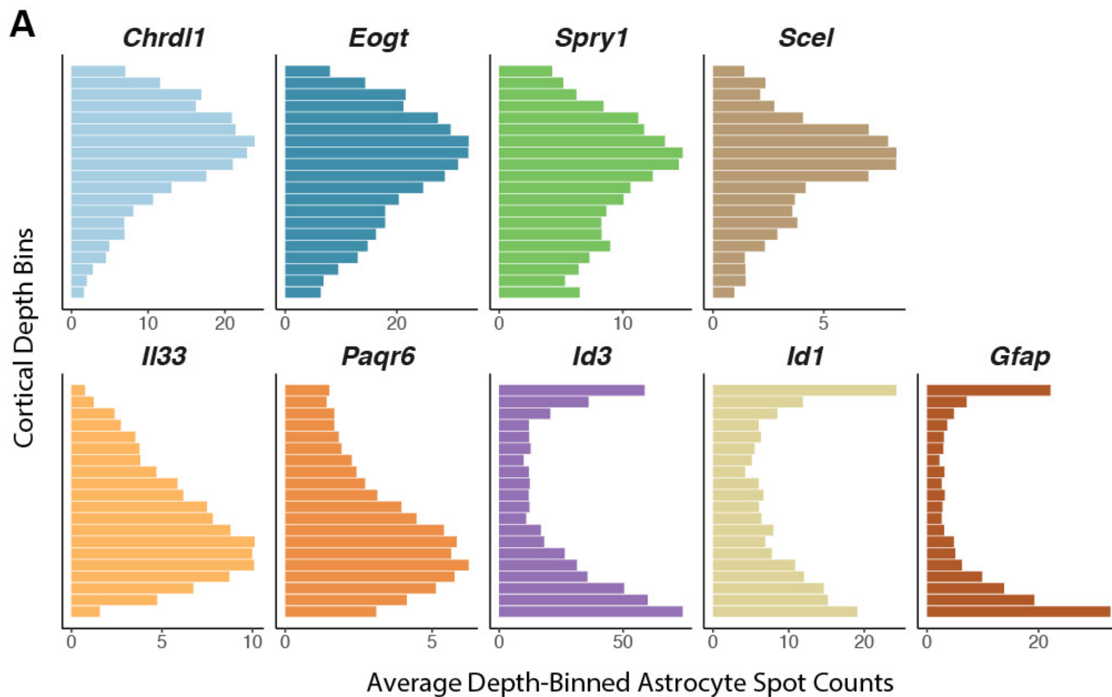
1562
1563
1564 **Supplementary Figure 10: Identification of cortical astrocytes with *Glast* smFISH.**
1565 **A)** (Left) Maximum z-projection image showing astrocytes in the P14 barrel cortex upper layers.
1566 (Middle) Segmentation of single astrocytes, outlined are astrocyte cell areas (green) and nuclei
1567 (white). (Right) Segmentation masks of individual astrocytes.
1568 **B)** Higher magnification image of an astrocyte indicated with dashed box in A. Bottom panel also
1569 shows the outline of the astrocyte cell area in dashed lines.
1570 **C)** Individual 40X z-planes throughout the same astrocyte. The arrow indicates the astrocyte
1571 nuclei marked by Glast and DAPI. For astrocyte segmentation, the z-stack is collapsed into a
1572 single plane via a maximum intensity projection. n=14 mice assayed across 4 independent
1573 batches, 5-10 tissue sections imaged.
1574 Scalebars: (A) 50 µm, (B,C) 10 µm
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589

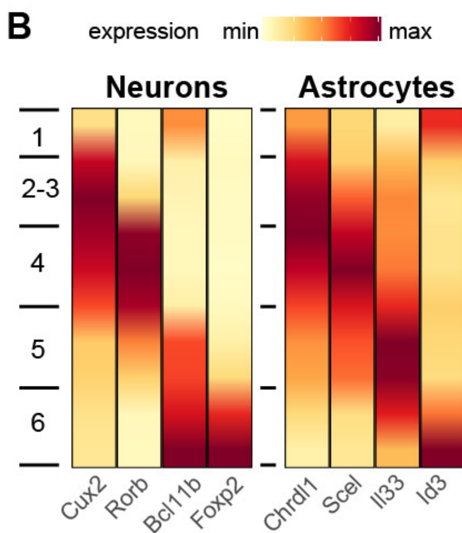


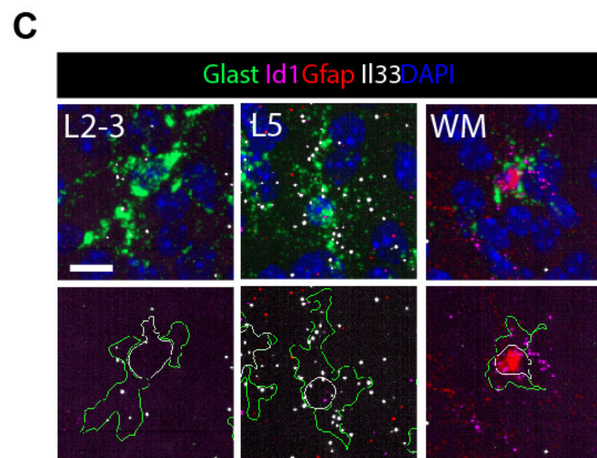
Supplementary Figure 11: Identification of astrocytes across different cortical layers.
A) Images (top) and segmentation masks (bottom) of astrocytes from barrel, motor and visual cortex. Midcortical layers (L4-5) are shown.
B) Astrocyte cell counts across cortical layers and areas are consistent across three technical replicates (different colors).
C) Astrocyte segmentation performs consistently across cortical areas. Violin, box and dot plots showing the cellular features of single astrocytes measured across four cortical areas. Deep layer astrocytes are slightly smaller and show lower expression of Glast than upper layer astrocytes. Number of astrocytes plotted across cortical astrocytes: A: 2628, M: 3777, S-BF: 3878 and V: 1968. *n*=1 mouse, 5-10 tissue sections independently imaged.
Scalebar: (A) 50 µm



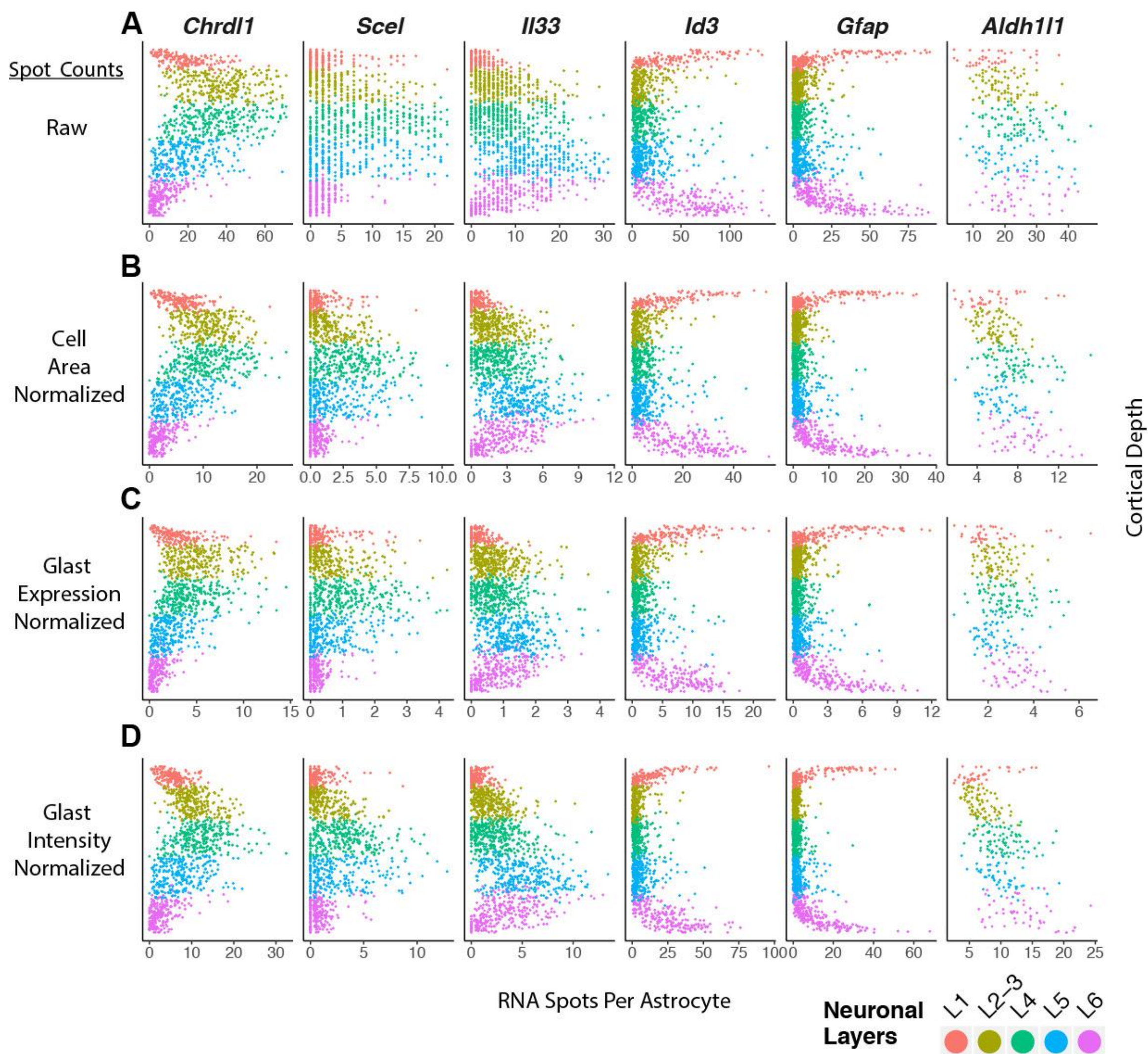
1607
1608
1609 **Supplementary Figure 12: Screening and selection of top layer astrocyte markers.**
1610 Quantification of single astrocyte, neuron and filtered astrocyte (i.e. removal of z-overlapping
1611 neurons and non-astrocyte nuclei) in situ expression of candidate layer astrocyte genes identifies
1612 several spatial and cell type-specific expression patterns. Screened genes show pan-astrocyte (A),
1613 gray matter astrocyte (B), astrocyte and neuron (C), and layer astrocyte enriched (D) expression
1614 patterns. *n*=2 mice assayed independently, 3 tissue sections imaged per replicate.
1615
1616
1617



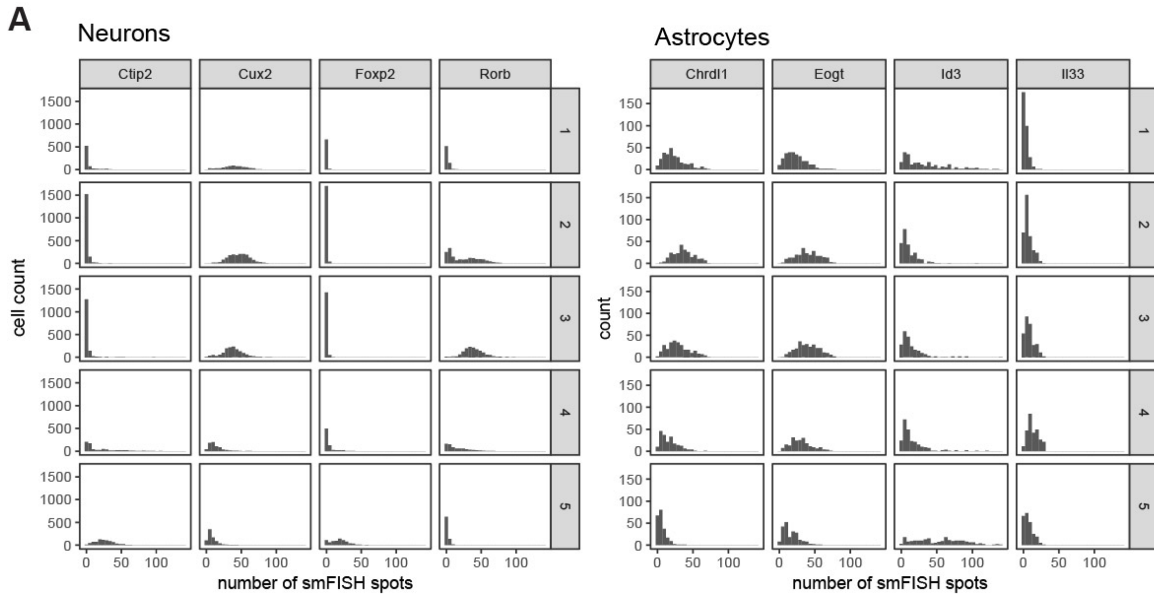


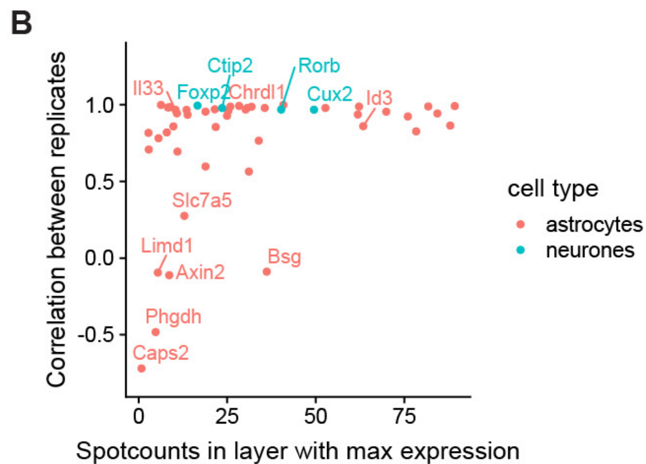


1618
1619
1620 **Supplementary Figure 13: Astrocyte layer gene expression diverges from neuronal laminae.**
1621 **A)** Quantification of astrocyte layer marker expression across cortical depth. Plots show the
1622 single astrocyte expression averaged across ten cortical depth bins in the P14 somatosensory
1623 cortex (n=2 pooled biological replicates).
1624 **B)** Interpolated tile expression plots comparing neuron vs astrocyte layer marker expression
1625 across cortical depth in the P14 barrel cortex (n=3 pooled tissue sections across the
1626 somatosensory cortex from one biological replicate). Astrocyte layer expression domains diverge
1627 from sharply refined neuronal laminae. A Chi2-test comparing the expression of astrocyte to
1628 neuron layer markers cortical depth supported this conclusion (p-value < 1e-16).
1629 **C)** *Il33* expression is enriched in L5 astrocytes but absent from white matter astrocytes at P14.
1630 *n*=2 mice independently assayed, 3 tissue sections imaged.
1631 Scalebar: (C) 10 µm

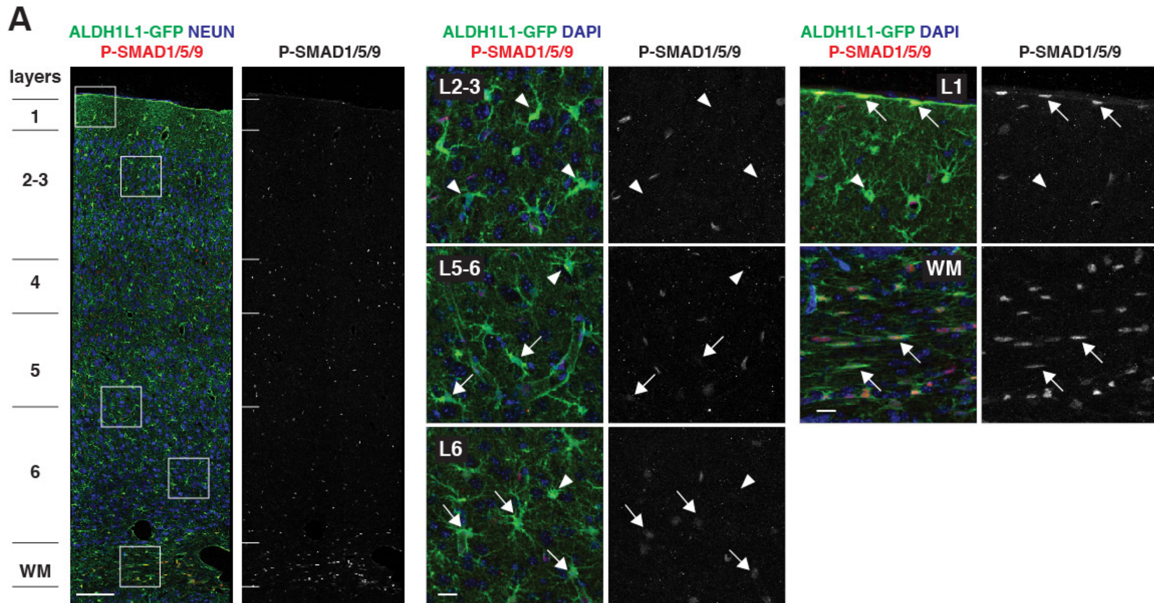


1632
1633
1634 **Supplementary Figure 14: The observed astrocyte layer gene expression patterns are not**
1635 **artifacts of cell size or *Glast* expression level differences.**
1636 Quantification of single astrocyte expression across cortical depth in the P14 barrel cortex (n=3
1637 pooled tissue sections from one biological replicate). The expression of identified layer astrocyte
1638 markers, the white matter astrocyte marker *Gfap* and the pan-astrocyte marker *Aldh1l1* are plotted
1639 as single cell RNA spot counts that are (A) raw, or normalized to (B) astrocyte area, (C) single
1640 cell *Glast* spot counts, and (D) single cell *Glast* signal intensity. *n*=2 mice independently
1641 assayed, 3 tissue sections imaged.
1642
1643
1644
1645

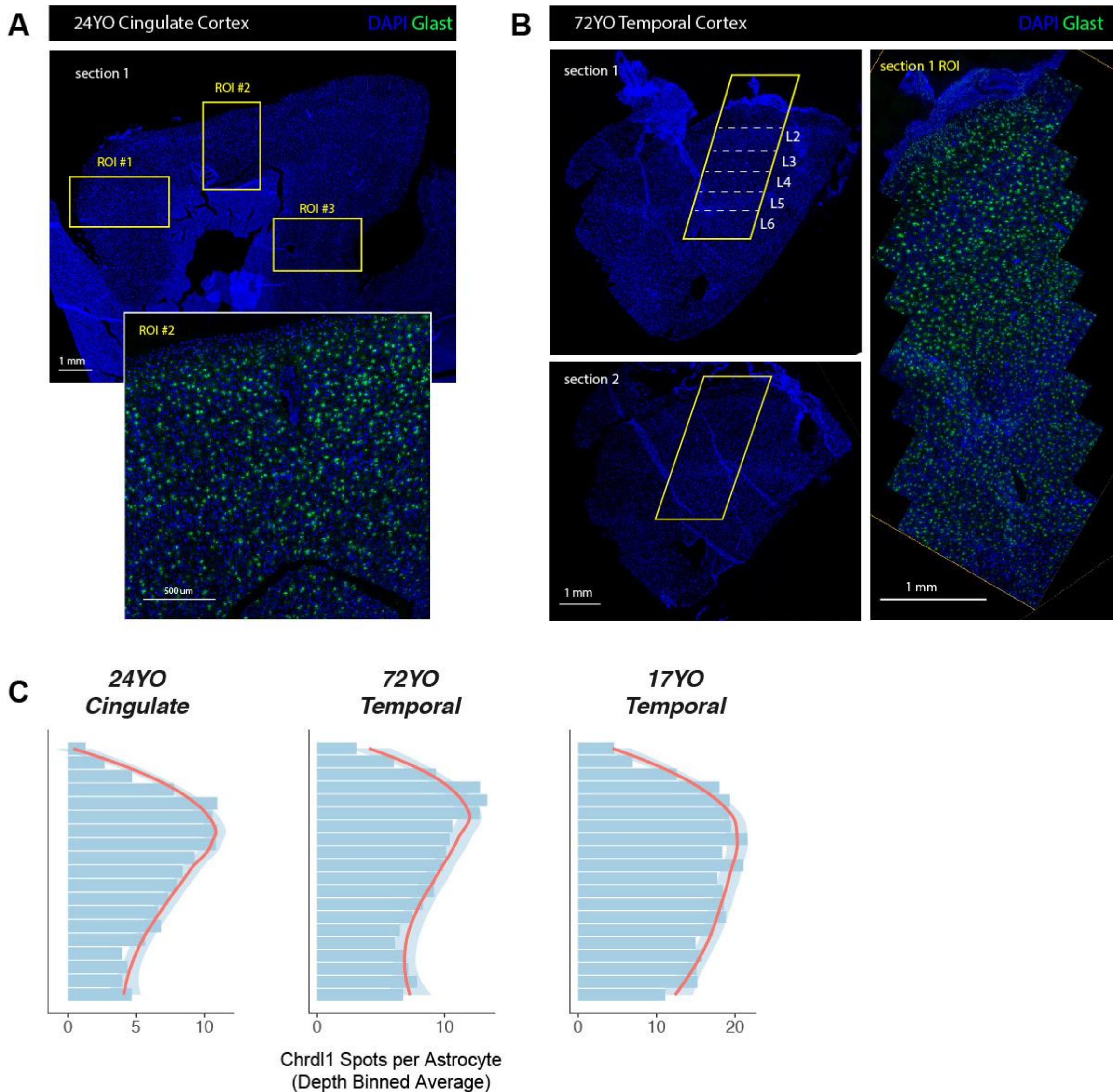




1646
1647
1648 **Supplementary Figure 15: Reproducibility of astrocyte expression patterns across gene**
1649 **expression levels.**
1650 **A)** The best neuronal (left) and astrocyte (right) cortical layer markers have similar
1651 expression level in bins with highest expression. Comparison of expression level (X-axis)
1652 histogram counts (Y-axis) of neuronal and astrocyte cortical layer markers (columns)
1653 across 5 spatial bins (rows).
1654 B) Spatial profiles of most astrocyte and neuron layer markers (color) correlate well
1655 between biological replicates (Y-axis) regardless of the expression level (X-axis,
1656 expression level in bins where genes are maximally expressed). Selected best markers as
1657 well as the genes with the least reproducible patterns are marked. Multiple tissue sections
1658 across the somatosensory cortex were assayed from two P14 mice for the analysis.
1659
1660
1661
1662

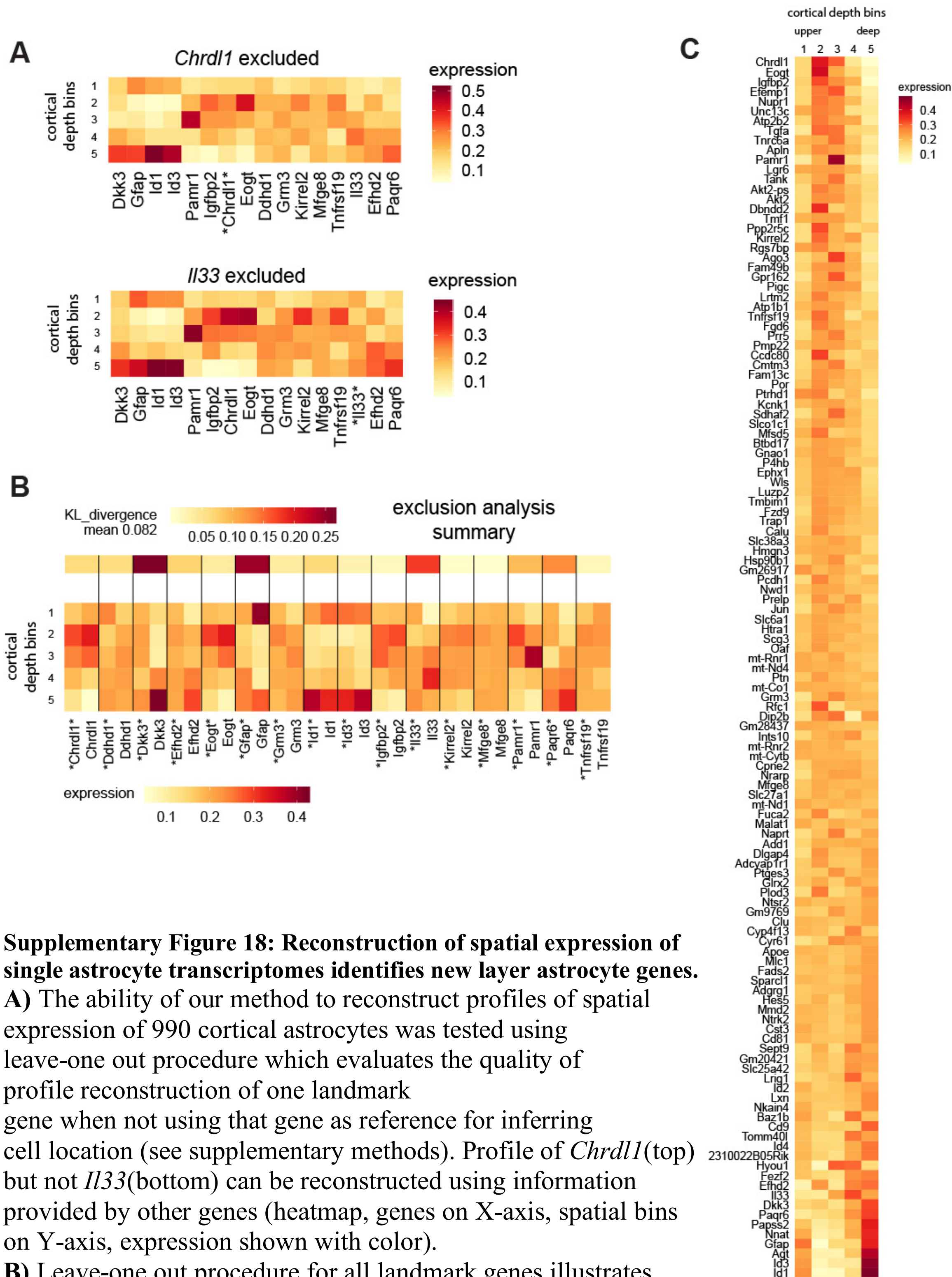


1663
1664
1665 **Supplementary Figure 16: Layer 6 astrocyte enriched Id3 expression is consistent**
1666 **with activated BMP signaling in this population assessed pSmad**
1667 **immunohistochemistry.**
1668 A) L6, WM and L1-subpia astrocytes show higher levels of phosphorylated Smad
1669 (pSmad) immunostaining that indicates Id protein activity. This pattern is consistent with
1670 the expression pattern of *Id3* mRNA detected with LaSTmap smFISH. Images from the
1671 P14 somatosensory cortex of the Aldh1L1-GFP astrocyte reported mice are shown.
1672 Arrows indicate astrocytes with high pSmad, while arrowheads indicate astrocytes with
1673 low pSmad staining. *n*=2 mice assayed in one experiment, 3 tissue sections imaged. b
1674 Scalebar: (large panels) 100 µm, (small panels) 25 µm.
1675
1676
1677
1678
1679
1680
1681
1682
1683



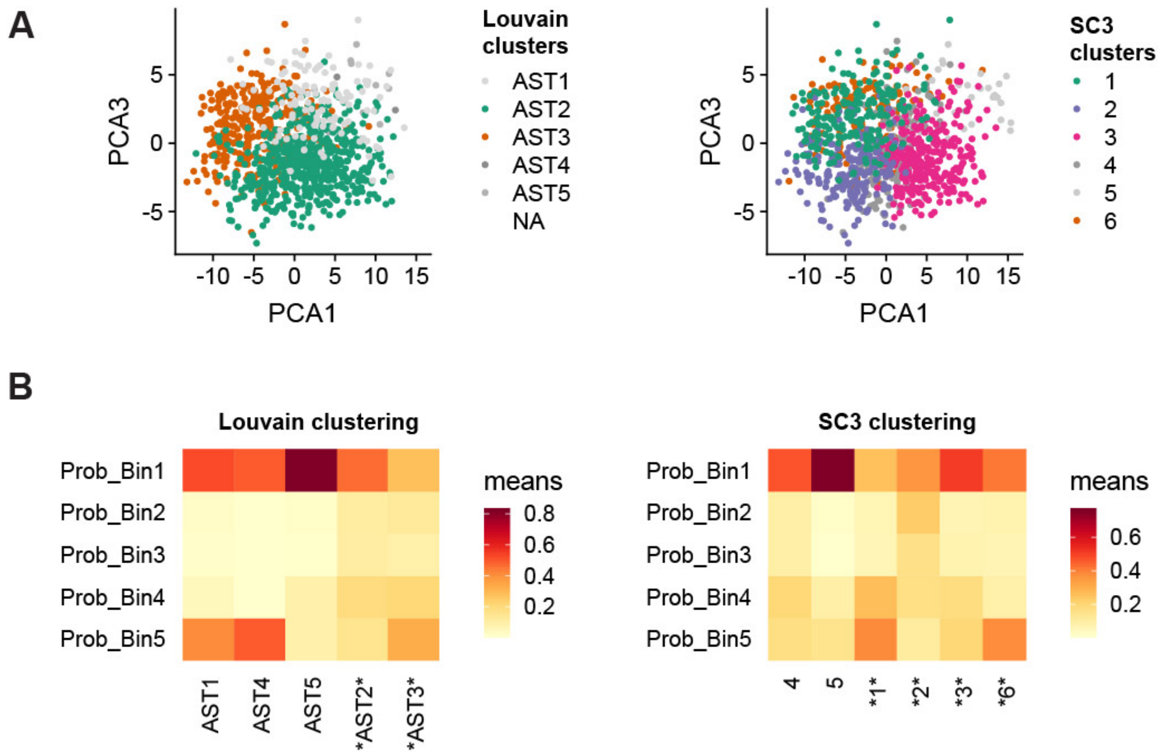
1684
1685

1686 **Supplementary Figure 17: *Chrdl1* expression is enriched in upper layer astrocytes in the**
1687 **adult human cortex.**
1688 **A)** Astrocytes in the 24 year old cingulate cortex. Low magnification images show DAPI staining
1689 of a section through the cingulate cortex (top) and Glast smFISH of the ROI #2 (bottom). Boxed
1690 regions of interest were imaged at 40X to quantify layer astrocyte expression of *Chrdl1*.
1691 **B)** Astrocytes in the 72 year old temporal cortex. Low magnification images show DAPI staining
1692 of two sections through the temporal cortex (left) and Glast smFISH of the ROI on the first
1693 section (right). Boxed regions of interest were imaged at 40X to quantify layer astrocyte
1694 expression of *Chrdl1*.
1695 **C)** Quantification of depth binned average single astrocyte expression of Chrdl1 across the 24YO
1696 cingulate, 72YO temporal and 17YO temporal cortex samples. The three samples were assayed
1697 and imaged independently.
1698 Scalebars: (DAPI) 1 mm, (DAPI-*Glast*) 0.5 mm
1699
1700
1701

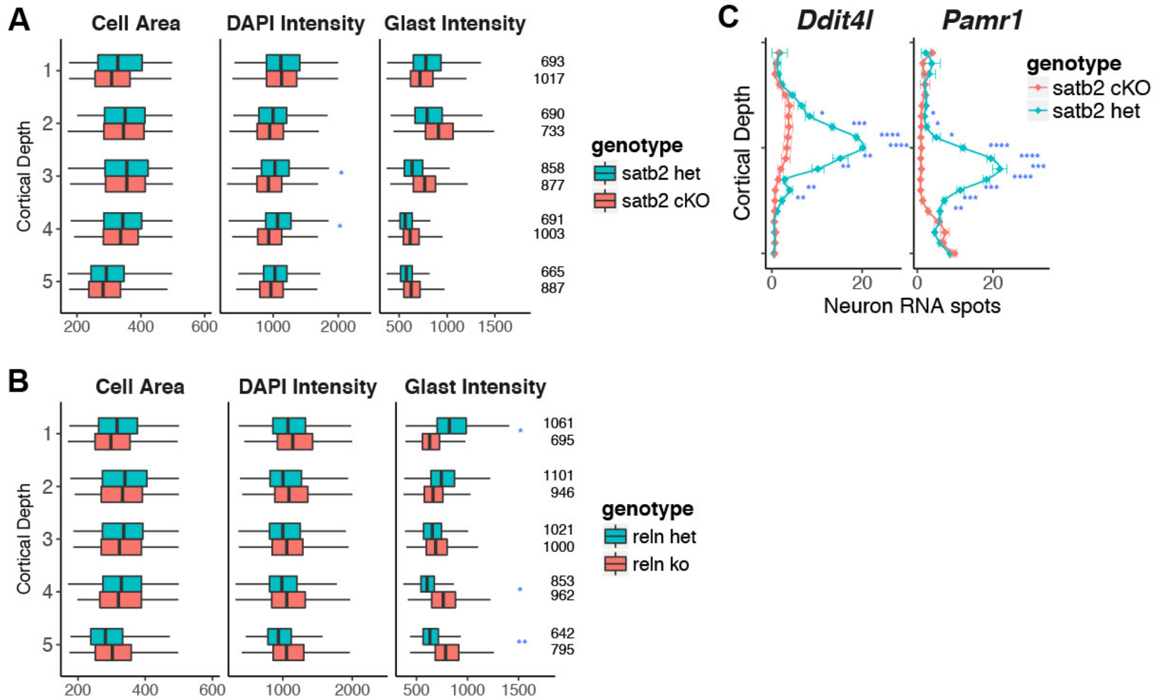


Supplementary Figure 18: Reconstruction of spatial expression of single astrocyte transcriptomes identifies new layer astrocyte genes. **A)** The ability of our method to reconstruct profiles of spatial expression of 990 cortical astrocytes was tested using leave-one out procedure which evaluates the quality of profile reconstruction of one landmark gene when not using that gene as reference for inferring cell location (see supplementary methods). Profile of *Chrdl1*(top) but not *Il33*(bottom) can be reconstructed using information provided by other genes (heatmap, genes on X-axis, spatial bins on Y-axis, expression shown with color).
B) Leave-one out procedure for all landmark genes illustrates which genes have profiles that are easily reconstructed (marked with asterisk) according to visual inspection, and relative entropy (top bar) compared to observed profile.
C) Expression profiles (color) of 125 significantly layer-restricted genes (Y-axis) across 5 spatial bins (X-axis).

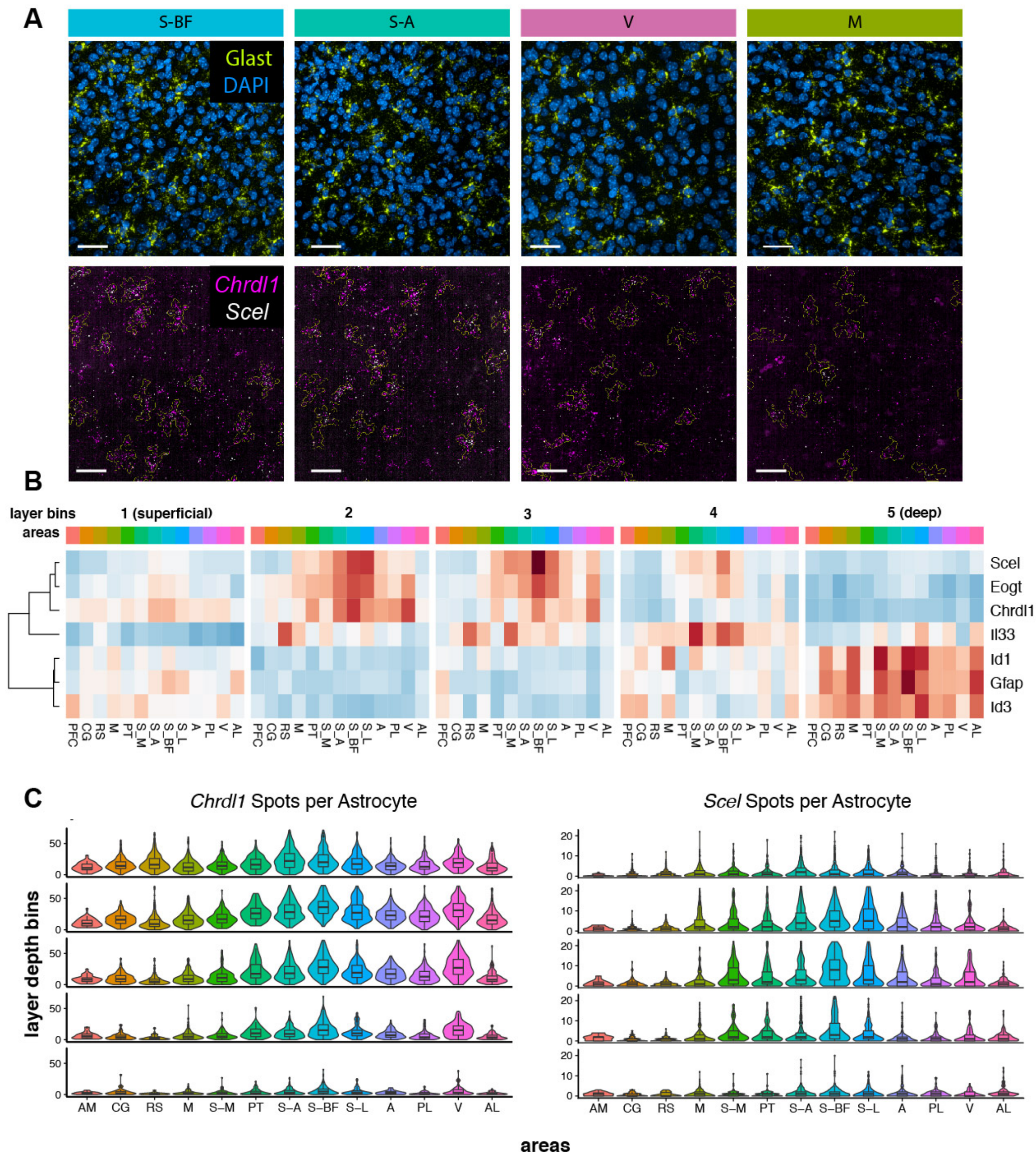
1749



1750
1751
1752
1753 **Supplementary Figure 19: scRNA-seq subtypes identified with unsupervised clustering do**
1754 **not show distinct pattern of layer locations**
1755 **A)** Louvain (left) and SC3 (right) unsupervised clustering methods identify cortical
1756 astrocyte subtypes (color), shown in Principal Component dimensions (PC1 – X-axis,
1757 PC2 – Y-axis) from 990 profiled cells. Louvain clustering is obtained from Batiuk et al
1758 (23) where cluster AST2 and AST3 are cortical astrocytes. SC3 clustering was performed
1759 cortical subset of the data using default parameters. The number of clusters was chosen
1760 based on p-value for the clustering and presence of coherent markers for most clusters.
1761 Clusters 4 and 5 were considered as poor quality due to the lack of markers and that these
1762 are hypothalamic clusters with few cells in the cortex.
1763 **B)** Clusters have overall similar location and cells from most clusters are located in bin 1
1764 and 5 (corresponding to layer 1 and 6). Probability of cell assignment (color) to spatial
1765 bins (Y-axis) average across cells from the same cluster (X-axis). High-quality Louvain
1766 (left) and SC3 (right) clusters are marked with asterisk. SC3 cluster 1 and 2 have higher
1767 location probability in middle layers than AST2 and 3, however, both cells from both
1768 Louvain and SC3 clusters do not follow clear layer pattern.
1769
1770
1771
1772

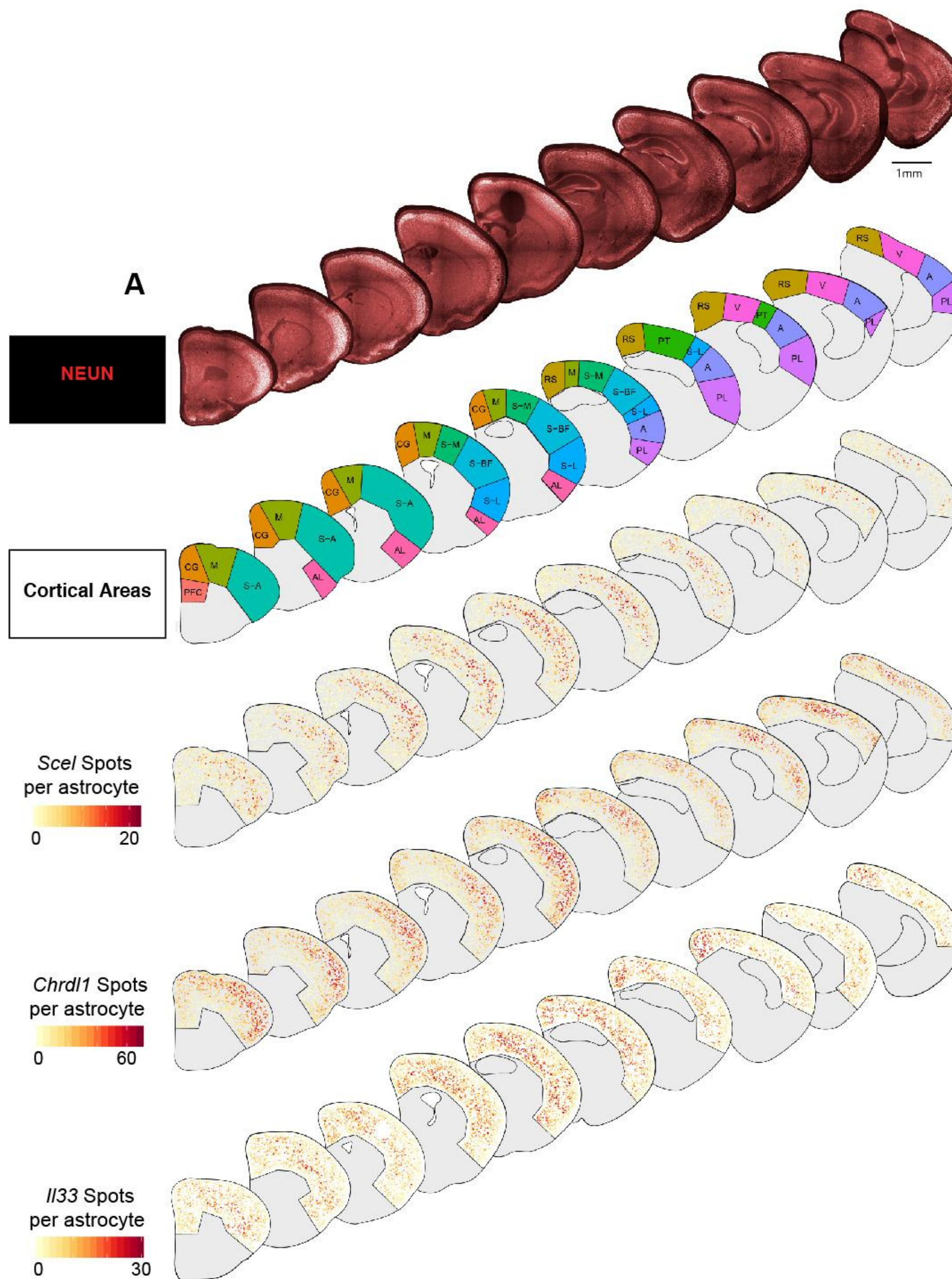


1773
1774
1775 **Supplementary Figure 20: Astrocyte cellular phenotypes in neuronal layer switch**
1776 **experiments and loss of L4 neuron gene expression in *Satb2* cKO.**
1777 **A)** Boxplots comparing astrocyte area, DAPI and Glast intensity between *Satb2* cKO and
1778 littermate controls in the P14 barrel cortex. No significant change is observed in these astrocyte
1779 features upon *Satb2* cKO. Cell counts per genotype per cortical depth bin are shown on the right
1780 (n=3 pooled biological replicates per genotype).
1781 **B)** Boxplots comparing astrocyte area, DAPI and Glast intensity between *Reeler* and littermate
1782 controls in the P14 barrel cortex. In *Reeler*, the difference in *Glast* expression between superficial
1783 and deep astrocytes is inversed, consistent with the inversion of astrocyte layers based on marker
1784 gene expression. Cell counts per genotype per cortical depth bin are shown on the right (n=3
1785 pooled biological replicates per genotype).
1786 **C)** Quantification of cortical depth binned neuronal layer marker expression in cKO vs control
1787 (n=3 pooled biological replicates per genotype). *Satb2* cKO shows loss of L4 neuron gene
1788 expression based on additional L4 markers, *Ddit4l* and *Pamr1*.
1789 Two-tailed student's t-test was used. Data represent mean ± s.d. *P < 0.05,**P < 0.01, ***P <
1790 0.001.
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801



1802
1803
1804 **Supplementary Figure 21: Astrocyte layer gene expression varies across cortical areas.**
1805 **A)** Images showing the enrichment of *Scel* and *Chrdl1* expression in somatosensory areas over
1806 motor and visual cortex. In the bottom panels, the single astrocyte segmentations are shown in
1807 green.
1808 **B)** Expression heatmap showing the expression of layer astrocyte markers across cortical depth
1809 and areas (assayed over n=10 tissue sections from one biological replicate).
1810 **C)** Violin plots showing the quantification of single astrocyte expression of *Chrdl1* and *Scel*
1811 across cortical layers and areas. Number of cells across cortical areas: AM: 141, CG: 858, RS:
1812 1276, M: 1298, S-M: 745, PT: 393, S-A: 2472, S-BF: 1231, S-L: 897, A: 838, PL: 690, V: 668
1813 and AL: 686. *n*=1 mouse, 10 tissue sections independently imaged.
1814 Scalebar: 50 µm

1815



1816
1817
1818 **Supplementary Figure 22: Maps showing the single cell level distribution of select neuronal**
1819 **subtypes.**
1820 **A)** (First row) Low magnification images of P14 hemisections from ten anatomical levels assayed
1821 for NEUN IHC. (Second row) Maps of broad cortical areas included in analysis of regional
1822 astrocyte gene expression. (Bottom rows) Maps showing single astrocyte expression of *Scel*,
1823 *Chrdl1* and *Il33* across the cortex. *n*=1 mouse, 10 tissue sections independently imaged.
1824 Scalebar: 1 mm

Supplementary Table Legends

Supplementary Table 1: Single cortical neuron smFISH dataset. Table listing cellular, anatomical and gene expression measurements of 69,318 single neurons identified across the P14 cortex. The cluster assignments of 46,887 single neurons used for subtype identification across 8 broad cortical areas are also listed. Every row is a single neuron and the table columns are described in the “Supp Table 1 Metadata” sheet.

Supplementary Table 2: Single cortical astrocyte smFISH screen dataset. Table listing cellular, anatomical and gene expression measurements of 41,187 single astrocytes screened in the somatosensory cortex across two biological replicates. 46 candidate layer astrocyte markers as well as the pan-astrocyte marker *Aldh1l1* and the white matter astrocyte marker *Gfap* were multiplexed with the astrocyte marker *Glast* across multiple slides, these are listed under “Supp Table 2 Metadata”. Every row is a single astrocyte and the table columns are described in the metadata sheet.

Supplementary Table 3: RNAscope probes used in this study. Table listing all of the RNAScope probes, their mRNA target regions and ACD catalog numbers.

Supplementary Table 4: Automated histology protocols and reagents. Tables listing the automated 4-plex RNAScope smFISH and IHC protocol used on the Leica BOND RX and the consumable reagents.

Supplementary Table 5: Imaging settings. Tables listing the fluorophores, light sources, exposure times and emission filters used for mouse and human tissue imaging.

Supplementary Table 6: The single neuron segmentation and gene expression quantification pipeline. Table listing all of the steps and settings used in the Harmony software.

Supplementary Table 7: The single astrocyte segmentation and gene expression quantification pipeline. Table listing all of the steps and settings used in the Harmony software.

Supplementary Table 8: List of abbreviations for cortical areas. Table listing all the broad cortical areas examined in this study.

Supplementary Table 9: Cortical layer astrocyte RNAseq data. The RNAseq expression pattern and differential gene expression statistics of 159 candidate layer astrocyte markers. The list of 46 top genes screened with smFISH is also provided. The table columns are described in the “Supp Table 9 Metadata” sheet.

Supplementary Table 10: Reconstruction of spatial expression– layer assignment probability of cells. The matrix of posterior probability of cell assignment which shows the probability of each cell belonging to each spatial bin (see methods equation 6).

Supplementary Table 11: Reconstruction of spatial expression – average profiles of later-restricted genes. The table shows average reconstructed expression levels for 4963 genes that are detected in at least 145 scRNA-seq cells, including Kruskal–Wallis test p-value and q-value corrected with q value method.