SUPPLEMENTARY TEXT

Supplementary Methods

Validation of the Percoll PLUS equilibrium density centrifugation step for isolation of ultrapure astrocytes using flow cytometry. Post-natal day (P)56 CD1 wild type or Aldh1l1-eGFP transgenic mice were used to prepare cell suspensions as described in the Online Methods. Following equilibrium density centrifugation (with Percoll PLUS density medium), cell pellets were re-suspended in 0.5% BSA/PBS (without Ca^{2+}/Mg^{2+}) (Sigma). The remaining supernatants were diluted 6 times with 1x HBSS (with Ca^{2+} and Mg^{2+}) (Sigma) and were spun down at $300g_{Av}$ for 10 mins at room temperature. Pelleted debris, with any remaining cells, was re-suspended in 0.5% BSA/PBS (without Ca^{2+}/Mg^{2+}). All samples were filtered through a 20 μm Nitex filter (SEFAR). Dead cells were excluded with the use of the viability dye 7-AAD (1:20 dilution) (eBioscience).

Flow cytometry was performed using a BD FACSCanto I (BD Biosciences), operated by FACSDiva software (version 6.1.3). Forward scatter (FSC) and Side Scatter (SSC) gating was used to eliminate debris. Compensations were set on single color controls. Gates were based on negative control samples. Dead cells were excluded based on 7-AAD staining. Sample fractions from the same preparation were adjusted to be equal in volume and sent through the flow cytometer at a constant speed to estimate the quantity of cells per fraction. At least 380,000 events were recorded per run. The experiment was repeated twice with identical results.

Validation of the FACS sorting strategy using flow cytometry. P56 - P70 CD1 wild type or Aldh1l1-eGFP transgenic mice were used. The protocols used to prepare the cell suspension and perform cell surface antibody labelling are described in the Online Methods. Flow cytometry was performed as described above. Forward scatter plots (FSC-W/FSC-A) were used to separate single cells from cell doublets and tissue clumps. At least 20,000 cells were recorded per sample. The experiment was repeated twice with identical results.

Quality control of astrocytes after FACS isolation.
Confirmation of single cell isolation: P49 - P63 C57BL/6J males were used in trial experiments using FACS to isolate single astrocytes. The protocols used to prepare the cell suspension and perform FACS are described in the Online Methods. Following FACS, cells were plated onto glass-bottom μ-dishes (Ibidi), coated with poly-D-lysine (0.1 mg/ml, Sigma), centrifuged at $300g_{Av}$ for 10 min at 4°C. This was followed by fixation using 4% PFA solution (Sigma) for 10 min at room temperature. DAPI (2 μg/ml final concentration) was added and cells were

imaged using an IX81 microscope (Olympus) with a 20x objective and Hamamatsu ImagEM (C9100-13) camera, controlled by Xcellence acquisition software (Olympus). Dishes were imaged using both phase contrast and fluorescence microscopy. Figures were prepared using FIJI and Adobe Illustrator CS6 16.0.3.

RNA quality: Immediately following FACS, cells were pelleted at $300g_{Av}$ for 5 min at 4°C. RNA was then isolated using a RNeasy Plus Micro kit (Qiagen), according to the manufacturer's instructions. RIN (RNA integrity number) was measured using an Agilent RNA 6000 Pico Kit (Agilent Technologies) on a 2100 Bioanalyzer (Agilent Technologies) controlled using 2100 Expert Software, according to standard protocols.

Supplementary Results

Validation of the Percoll PLUS equilibrium density centrifugation step for isolation of ultrapure astrocytes using flow cytometry. Percoll PLUS was used, as previously described[1], for robust removal of myelin (and other debris) that contaminant the cell suspension, using equilibrium density centrifugation. This step was critical as we found myelin and debris to interfere with ACSA-2 staining and FACS sorting. The different buoyant densities of cells and myelin (plus other debris) meant that cells were pelleted, while the majority of myelin and debris remained in the supernatant. Care was taken to ensure that all astrocytes were pelleted, to avoid introducing any bias into the subsequent analysis. Supplementary Figure 1 shows the results of an experiment designed to assess the distribution of viable astrocytes between the cell pellet and supernatant following centrifugation. We found that 95%+ of astrocytes were found in the pellet (irrespective of the brain region from which they were isolated). This experiment was repeated twice with identical results (for both brain regions).

Validation of the FACS sorting strategy using flow cytometry. FACS gates were optimized to allow recovery of all astrocytes, including those with relatively weak ACSA-2 staining, while avoiding isolation of oligodendrocytes, a fraction of which are also positive for ACSA-2 (Supplementary Figure 2a). For this reason, we excluded oligodendrocytes during FACS by including a stain for O1, a validated marker of mature oligodendrocytes[2], as there was no O1 labelling of eGFP-positive cells isolated from the Aldh1l1-eGFP mouse (which is considered to be the current gold-standard pan-astrocyte marker line)[3] (Supplementary Figure 2b).

To avoid introducing bias into our analysis through isolation of a subpopulation of astrocytes, we validated our use of the ACSA-2 staining method using the Aldh1l1-eGFP mouse line. In our hands, ACSA-2 labelled all eGFP-positive astrocytes isolated from cortex and hippocampus (Supplementary Figure 2c). Furthermore, the vast majority (>96%) of identified astrocytes, in a previously published single cell dataset[4], expressed *Atp1b2* (Supplementary Figure 3), which is the epitope for the ACSA-2 antibody[1].

Quality control of astrocytes after FACS isolation. Single astrocytes were FACS isolated using the strategy depicted in Supplementary Figure 4. The quality of the single cell preparation was quality controlled in a number of ways, prior to use in library preparation.

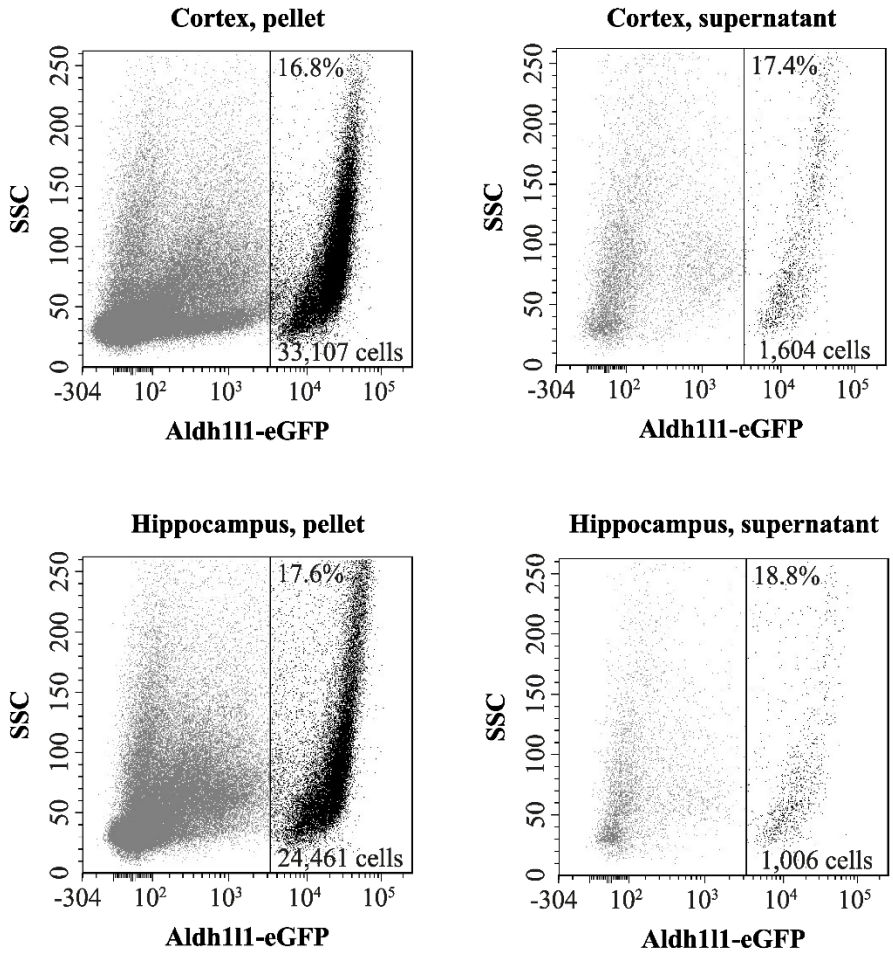
Confirmation of single cell isolation: The quality of the cell preparation was checked to assess the number of single cells in the FACS sorted sample, relative to clumps of cells or cell debris. Adjustment of FSC/SSC gating was made to ensure that all astrocytes, irrespective of size or granularity, were retained, while avoiding extensive amounts of small debris (Supplementary

Figure 4). This procedure isolated cells of high viability (85-92%) as checked using vital dye (Trypan Blue) exclusion. Cells were then attached onto plates, fixed and imaged in phase contrast, as well as with DAPI dye staining (Supplementary Figure 5a). For these experiments, single cells were defined as having a clearly demarcated membrane under phase contrast and a single DAPI stained nucleus. Cell debris seen under phase contrast did not show DAPI staining, whereas cell clusters contained multiple DAPI signals. Cell suspensions prepared from cortex contained on average 77.9% single cells, 20.5% debris and only 1.6% clustered cells. Likewise, cell suspensions prepared from hippocampus contained on average 75.2% single cells, 21.6% debris and only 3.2% clustered cells. Samples for this set of experiments were prepared in 4 separate FACS isolation procedures (2 separate isolations per brain region).

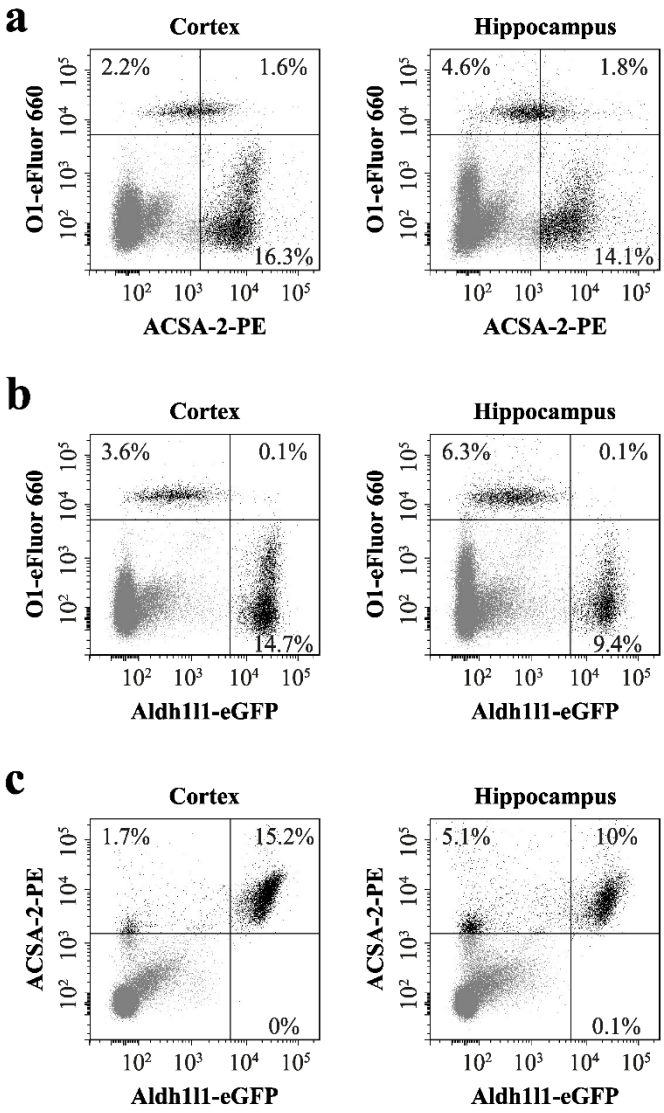
RNA quality: The quality of RNA extracted from sorted astrocytes was checked, as an indicator of general cell health. Bioanalyzer analysis showed that RNA extracted from isolated astrocytes was of good quality with the RNA integrity number (RIN) between 7.9 and 8.6 (where 1 is fully degraded RNA and 10 is fully intact RNA) (Supplementary Figure 5b).

Adaptation of the Smart-seq2 protocol for library preparation from adult astrocytes. The standard Smart-seq2 protocol[5] we originally used did not produce cDNA libraries of good quality, likely due to the low amount of RNA present in adult astrocytes[6]. To improve library quality, we systematically varied steps in the published protocol. First, we reduced the concentration of the template switching oligonucleotide (TSO) used in the final reaction mixture to 0.2 μM, thus reducing the amount of small primer concatamers present. Second, the number of cycles used in the PCR pre-amplification step was increased to 22. Finally, amplified concatamers were depleted using Agencourt Ampure XP beads in a modified ratio of beads to DNA of 0.8:1. Representative fragment analyzer electropherograms of cDNA produced from a single cell, or a pool of libraries, prepared using this modified method are shown in Supplementary Figure 6.

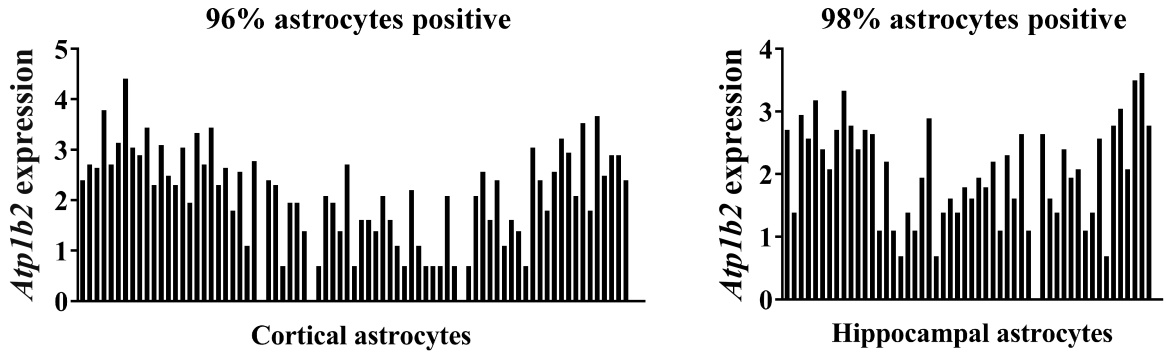
Supplementary Figures.



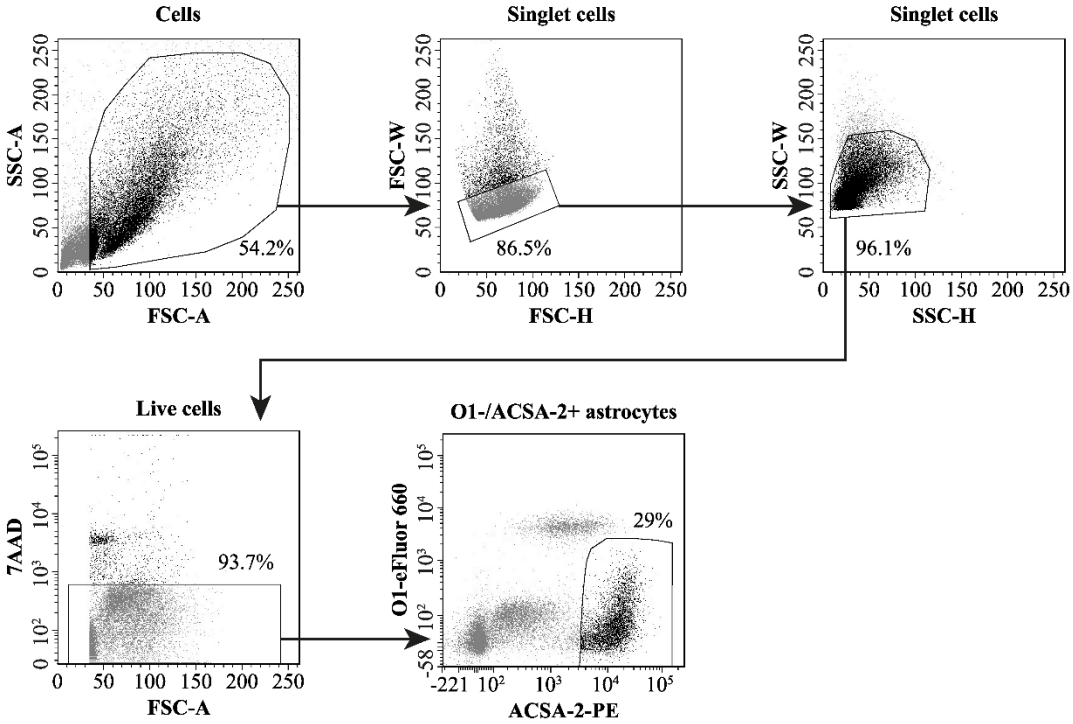
Supplementary Figure 1. Validation of the Percoll PLUS equilibrium density centrifugation step for isolation of ultrapure astrocytes using flow cytometry. Equilibrium density centrifugation with Percoll PLUS efficiently removed myelin and cell debris from the single cell suspension but did not affect astrocyte recovery. Following centrifugation, the resulting cell pellet and supernatant (containing myelin and debris) were brought to the same volume and analyzed using flow cytometry. Aldh1l1-eGFP-positive astrocytes were efficiently pelleted (>95% recovery), and were not retained in the supernatant, irrespective of whether tissue from cortex or hippocampus was used in the initial dissociation. The experiment was repeated twice with identical results. SSC: side scatter. The vertical line represents the gate separating Aldh1l1-eGFP-positive astrocytes from eGFP-negative cells. Numbers within the gate indicate the number of cells recorded during the run, both in absolute numbers and as a percentage of the total cell population.



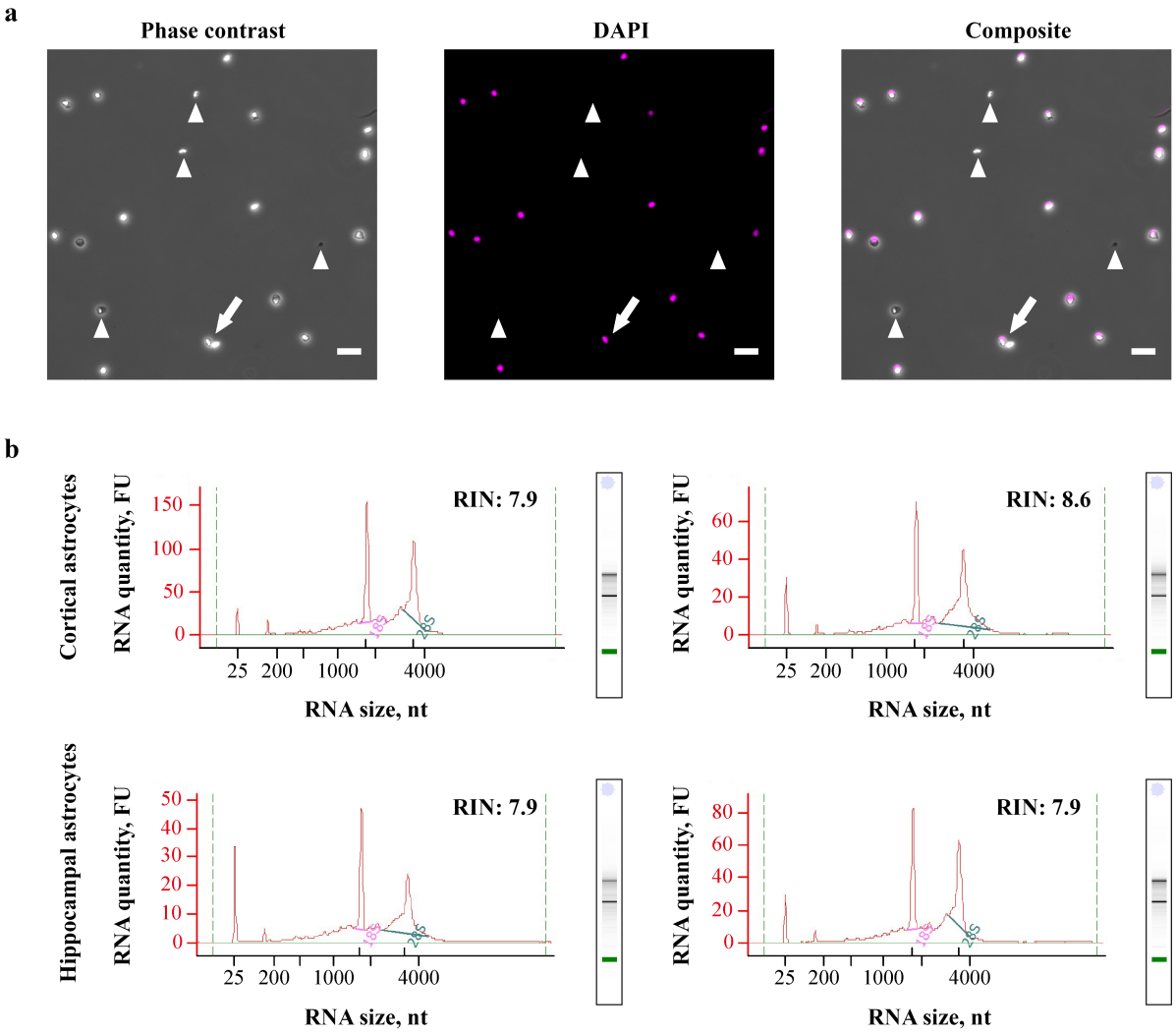
Supplementary Figure 2. Validation of the FACS sorting strategy using flow cytometry. (a) Cortical and hippocampal cell suspensions were stained with oligodendrocyte-specific (O1-eFluor660) and astrocyte-specific (ACSA-2-PE) antibodies. Flow cytometry analysis revealed that a fraction of ACSA-2 labelled cells co-stained with O1. On average, 13% of ACSA-2-positive cells were oligodendrocytes in the cortical cell suspension, and 18% in the hippocampal cell suspension. O1-positive oligodendrocytes were excluded in astrocyte sorting experiments. (b) eGFP-positive cells isolated from the Aldh1l1-eGFP mouse (which is considered to be a pan-astrocyte marker line) were negative for O1-staining. (c) The ACSA-2 antibody labelled all Aldh1l1-eGFP-positive astrocytes in the cell suspension. Flow cytometry experiments were repeated twice, with independent tissue preparations, and gave essentially identical results. Percentages represent the fraction of total cells present within the gate.



Supplementary Figure 3. Expression profile of *Atp1b2* in single astrocytes. *Atp1b2* encodes the epitope for the ACSA-2 antibody. To avoid introducing bias into our analysis through isolation of a subpopulation of astrocytes, we assessed the expression profile of *Atp1b2* across astrocytes using a previously published single cell dataset[4]. The vast majority of cells, in both cortex and hippocampus, expressed *Atp1b2*, confirming its utility as a general astrocyte marker (see also Supplementary Figure 2). Data is shown in ln-scale.

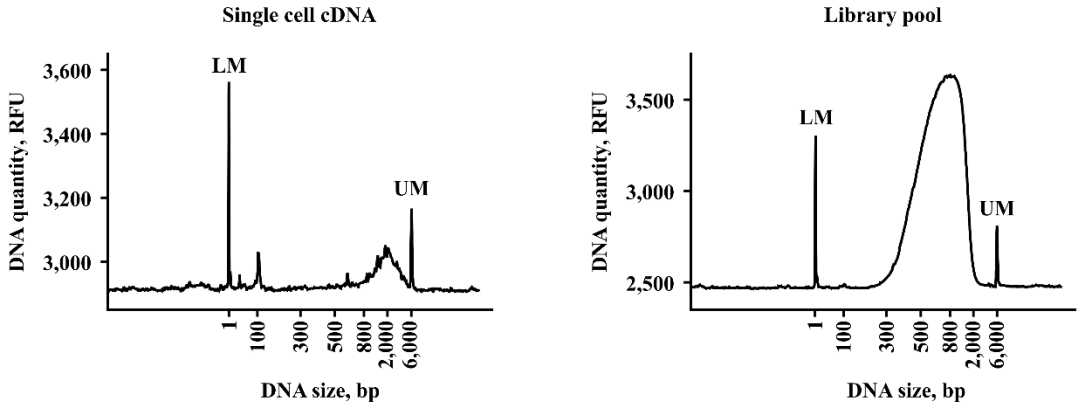


Supplementary Figure 4. FACS sorting strategy for adult astrocytes. Cells were initially selected using forward and side scatter area plots (FSC-A and SSC-A). Cells were selected from inside the indicated gate. To minimize the amount of small debris collected, care was taken to adjust the lower limit on the forward scatter (measure of size) axis, although the gate remained wide enough that smaller cells were still captured. Cell doublets were excluded using forward/side scatter width vs height plots (FSC-W/FSC-H and SSC-W/SSC-H). Dead cells were excluded based on 7-AAD staining. Astrocytes were isolated as ACSA-2-PE-positive/O1-eFluor660-negative cells. Percentages represent the fraction of total cells present within the gate.

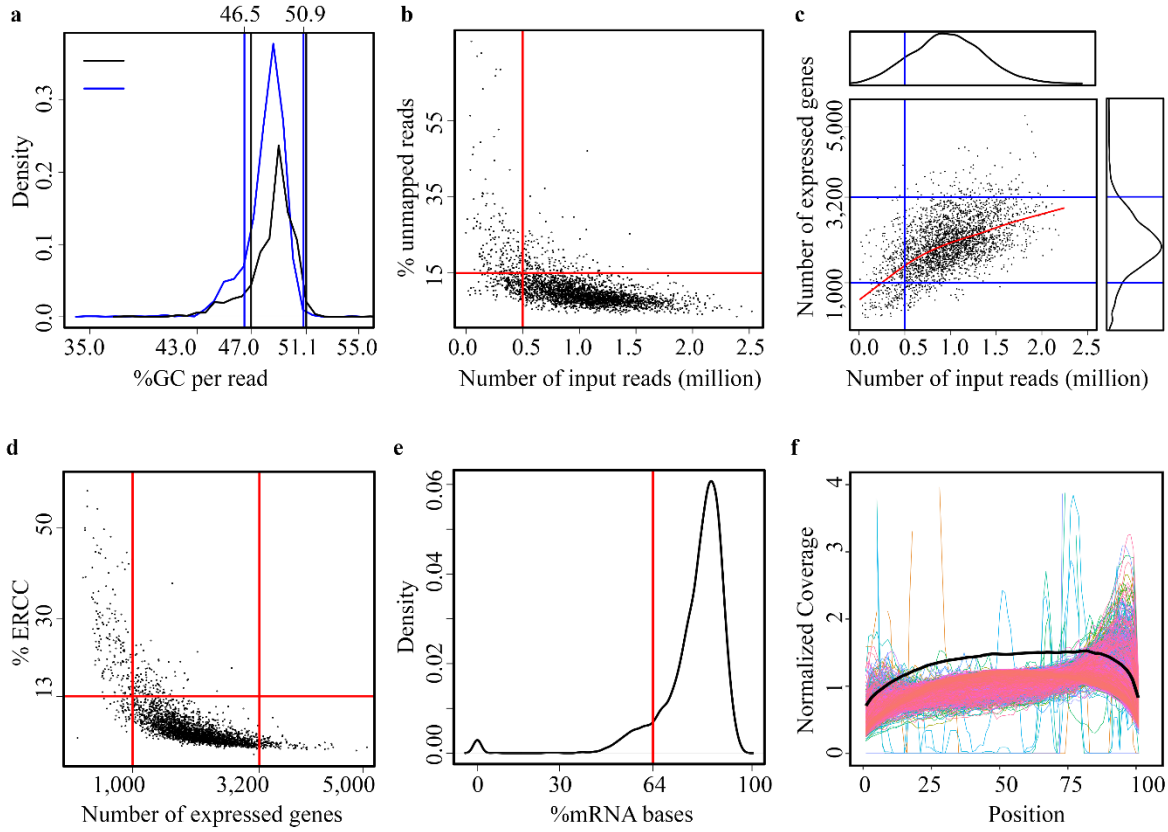


Supplementary Figure 5. Quality control of astrocytes after FACS isolation. (a) The amount of debris and cell clumps in the FACS sorted sample was checked by phase contrast and fluorescence (DAPI) microscopy. Objects without a DAPI signal were considered debris. Objects appearing as several cells clumped together in the phase contrast image and/or with multiple DAPI signals were counted as cell doublets. More than 75% of sorted objects were single cells. Arrow, cell doublet; Arrowhead, debris. Experiments were repeated twice for each brain region with essentially identical results. Scale bar, 30 µm.

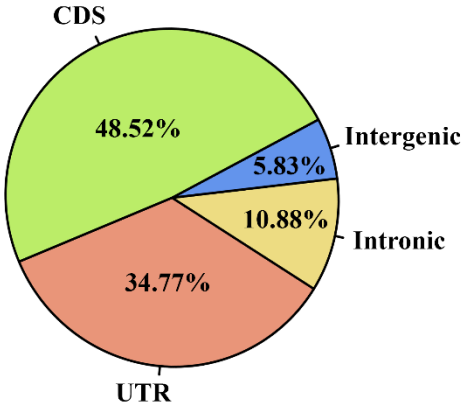
(b) The quality of RNA recovered from an astrocyte population post-sorting was checked using a Bioanalyzer. The analysis confirmed that the RNA recovered from both hippocampal and cortical astrocytes was of high quality (RIN 7.9-8.6).



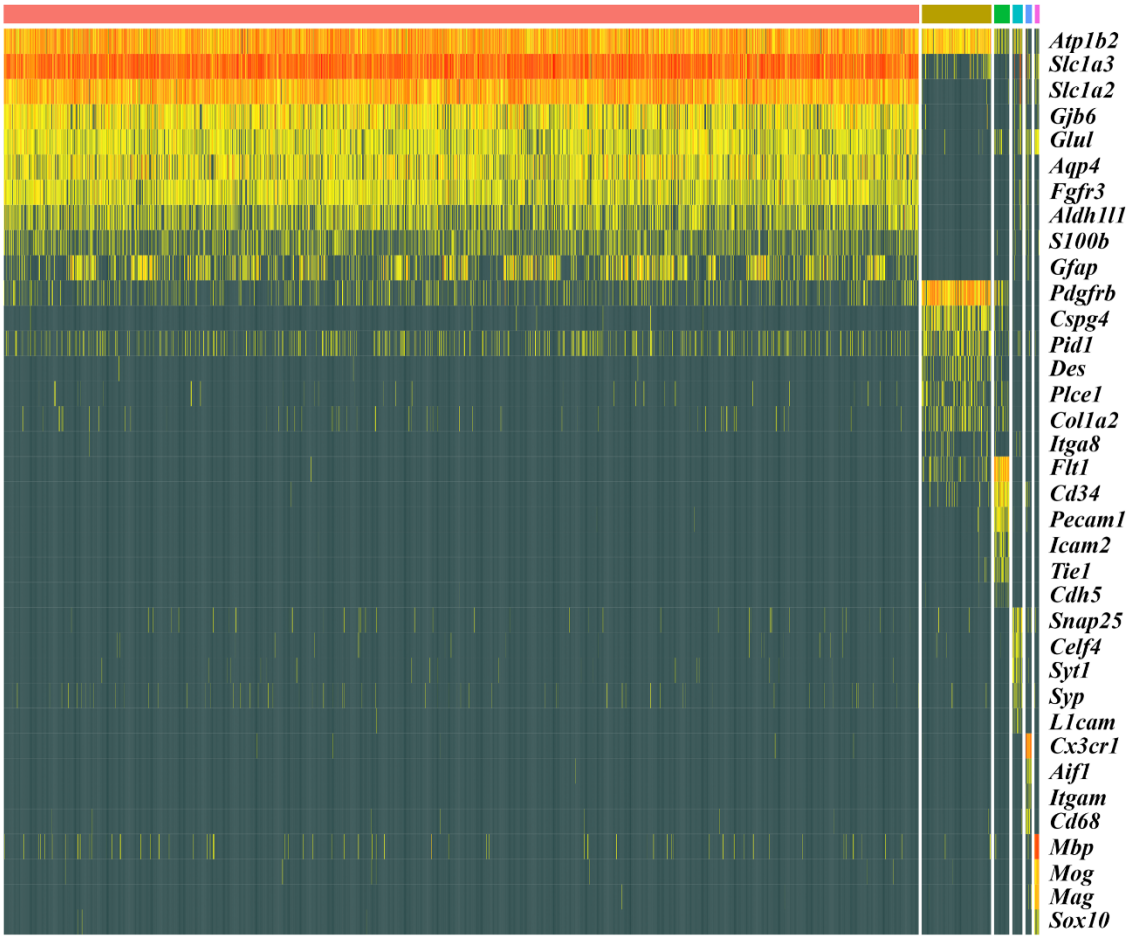
Supplementary Figure 6. Representative electropherograms obtained after Bioanalyzer analysis. Left: Pre-amplified cDNA from a single astrocyte. Conditions for cDNA preparation were adjusted for experiments using small cells (<12 μm diameter), to simultaneously reduce the level of primer concatamers (approx. 100 bp) and improve the quantity of *bona fide* pre-amplified cDNA obtained (size between 800 – 6000 bp). Right: Pool of 384 single cell RNA sequencing libraries produced from cDNA using a Nextera XT DNA Library Preparation Kit. LM, lower marker; UM, upper marker; RFU, relative fluorescence units.



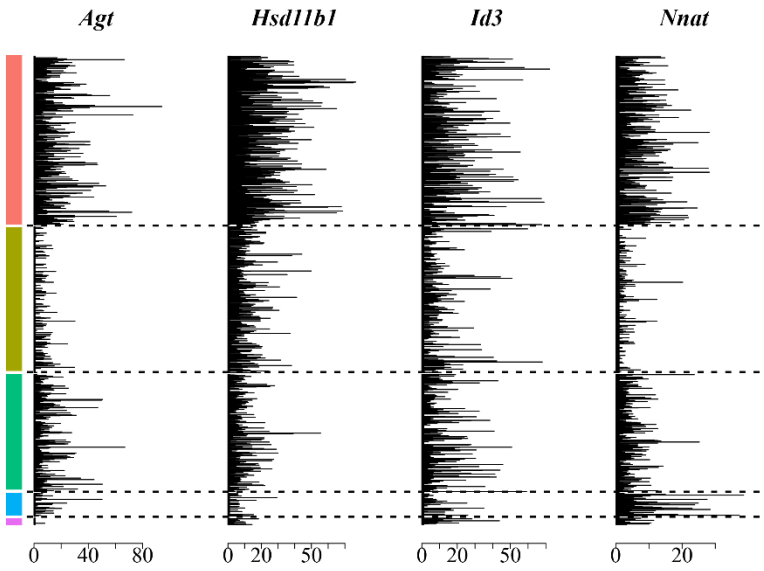
Supplementary Figure 7. Computational removal of low quality single cell libraries. Low quality libraries were removed based on: (a) the proportion of GC reads per transcript (falling outside the mean range of 46.5% - 50.9%, or the median range of 47.0% - 51.1%); (b) the amount of input reads per library (<0.5 million), or fraction of unmapped reads per library (>15%); (c) the number of genes expressed per cell (<1000 or >3200); (d) the proportion of ERCC reads relative to total (>13%); (e) the fraction of mRNA reads per cell (<64%); (f) 5' or 3' mapping bias (0: beginning of 5' UTR, 100: end of 3' UTR). Black, blue and red curves in (a-f) show the cut-offs separating high and low-quality libraries. The red curve in **(c)** represents the best fit of the data. 2,031 high quality cDNA libraries passed quality control and were retained for further analysis.



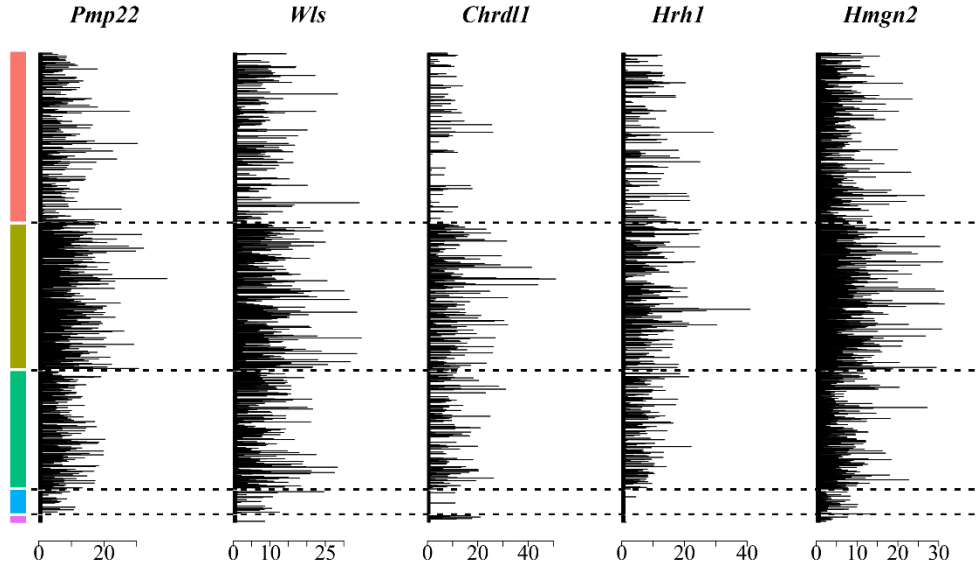
Supplementary Figure 8. Average distribution of reads across the mouse genome for the libraries retained for analysis. In the 2,031 high quality sequencing libraries retained for analysis (see Supplementary Figure 7), a high fraction of reads mapped to exons (CDS, coding sequence; UTR, untranslated region). Conversely, a low fraction of reads mapped to intronic and intergenic regions.



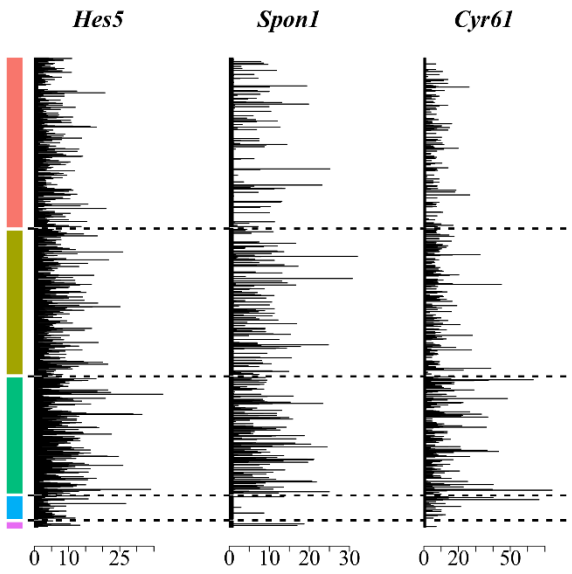
Supplementary Figure 9. Gene expression heatmap for cell type specific markers expressed in the high-quality cells retained for analysis. 2,015 cells (columns) were grouped based on clusters identified by high level Seurat clustering (Figure 1a). Expression of standard marker genes (rows) was restricted to specific clusters, identifying the major CNS cell types. Red, astrocytes; brown, mural cells; green, endothelial cells; turquoise, neurons; blue, microglia; pink, oligodendrocytes. Normalized gene expression data is shown in ln-scale. Grey, no expression; yellow, low expression; red, high expression.



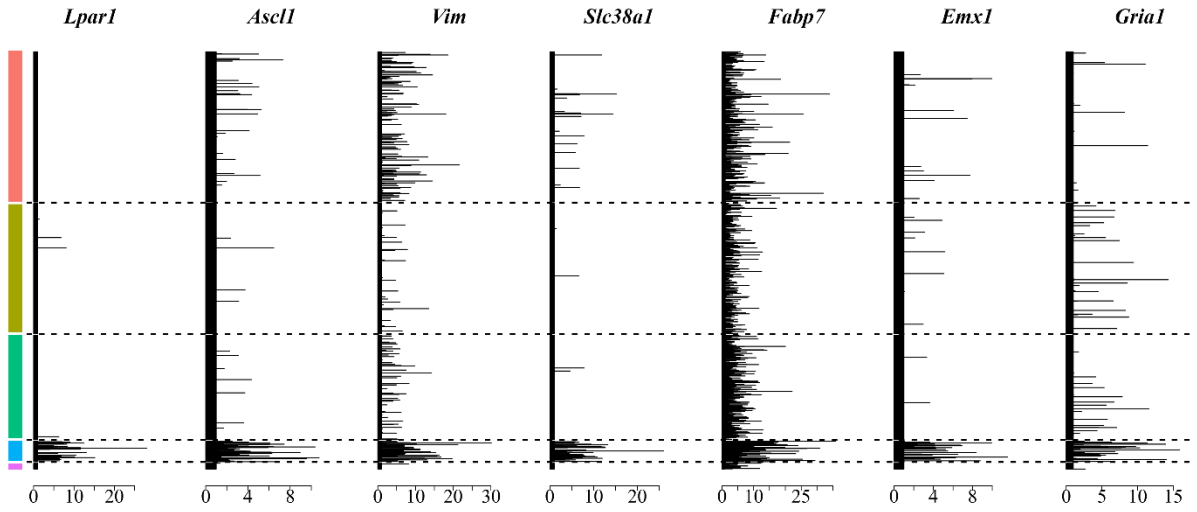
Supplementary Figure 10. Selected genes overexpressed in AST1. AST1 (red), AST2 (green), AST3 (turquoise), AST4 (blue) and AST5 (purple) are delineated with dotted lines. Expression data (Seurat normalized counts) is shown on a linear scale.



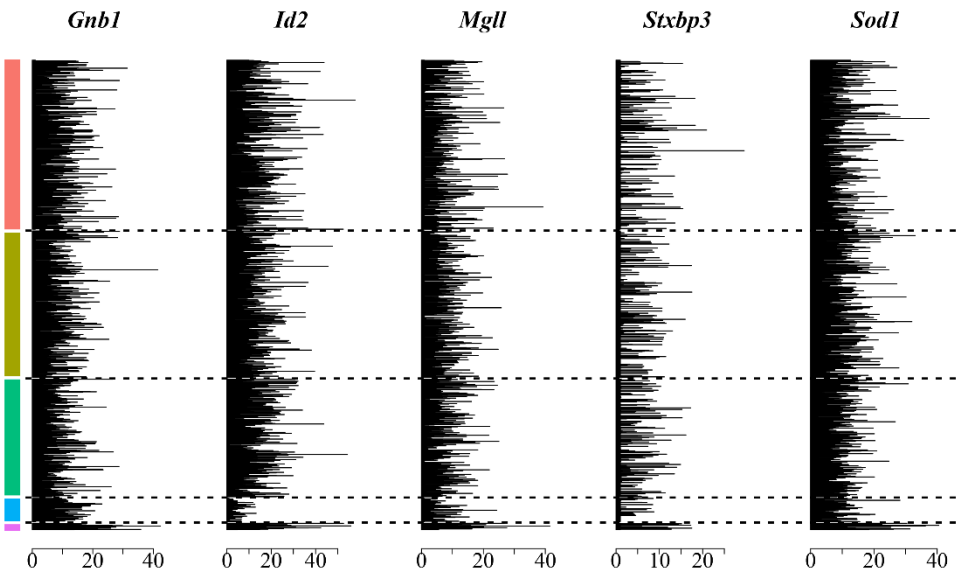
Supplementary Figure 11. Selected genes overexpressed in AST2. AST1 (red), AST2 (green), AST3 (turquoise), AST4 (blue) and AST5 (purple) are delineated with dotted lines. Expression data (Seurat normalized counts) is shown on a linear scale.



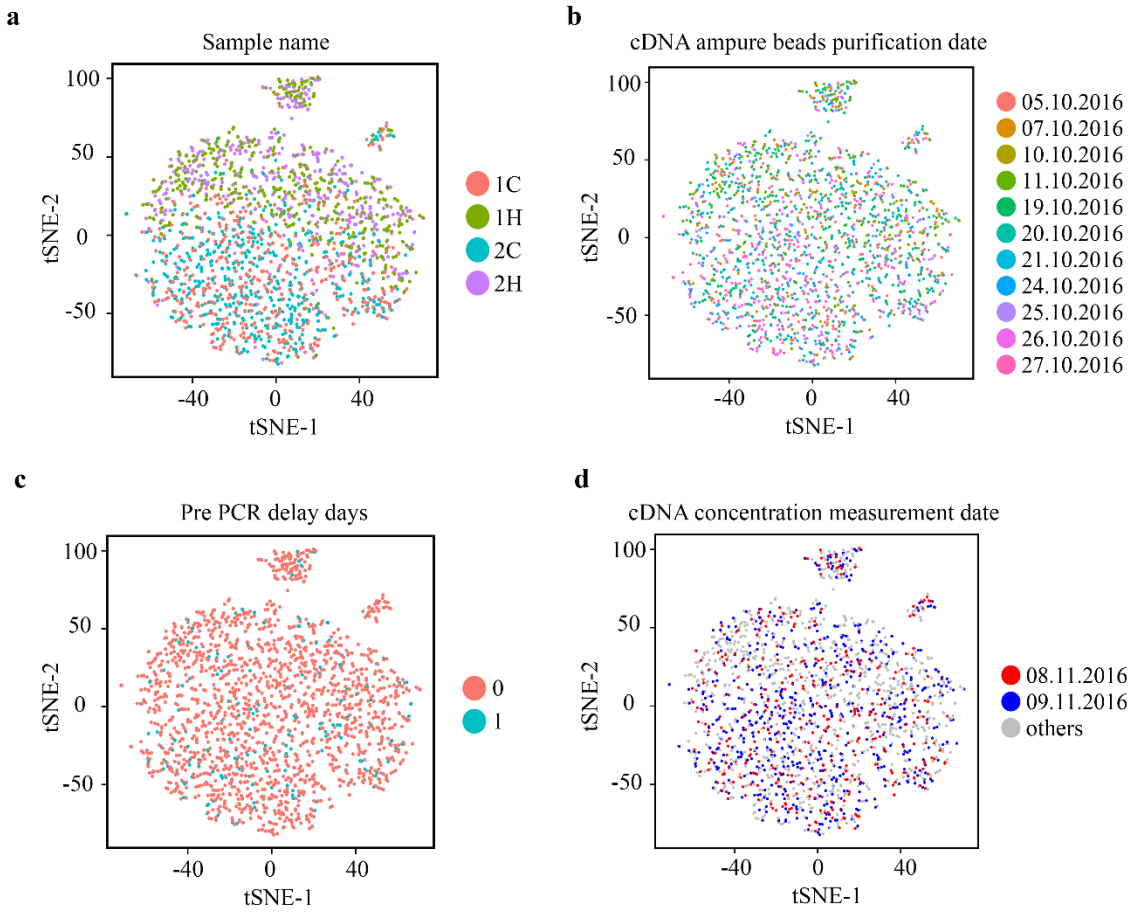
Supplementary Figure 12. Selected genes overexpressed in AST3. AST1 (red), AST2 (green), AST3 (turquoise), AST4 (blue) and AST5 (purple) are delineated with dotted lines. Expression data (Seurat normalized counts) is shown on a linear scale.



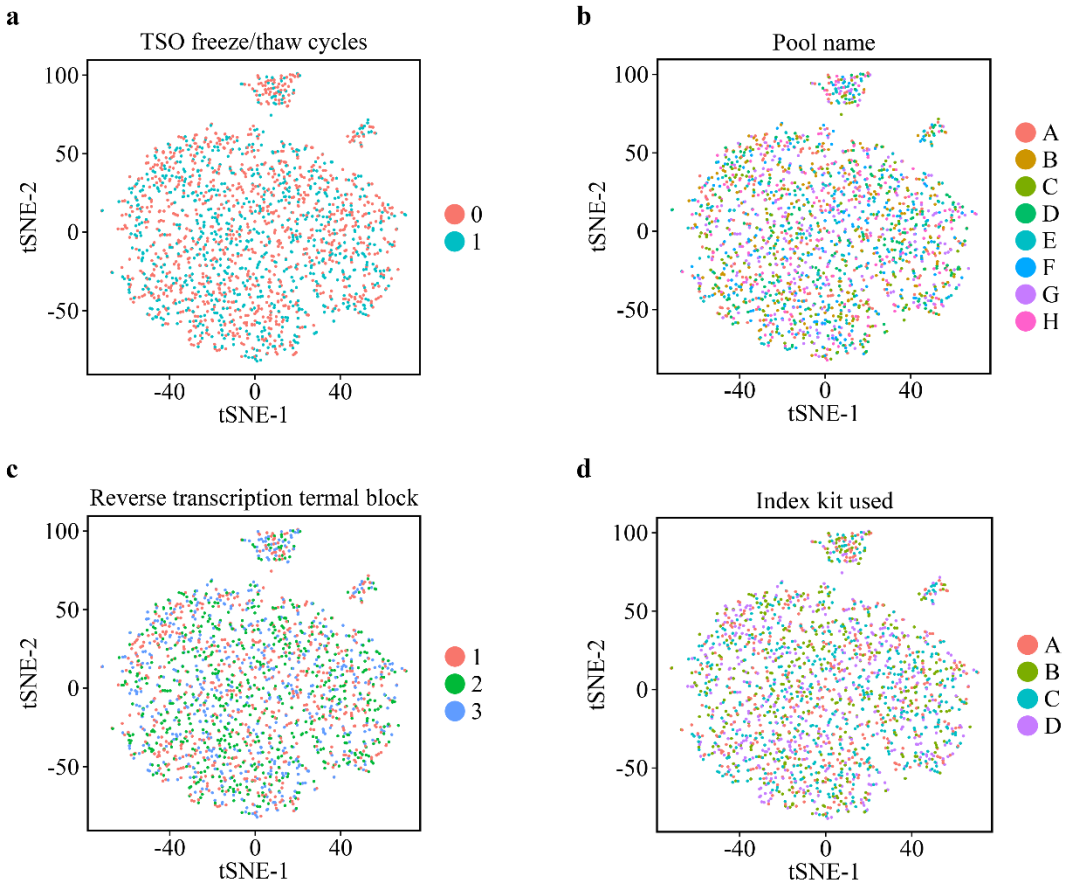
Supplementary Figure 13. Selected genes overexpressed in AST4. AST1 (red), AST2 (green), AST3 (turquoise), AST4 (blue) and AST5 (purple) are delineated with dotted lines. Expression data (Seurat normalized counts) is shown on a linear scale.



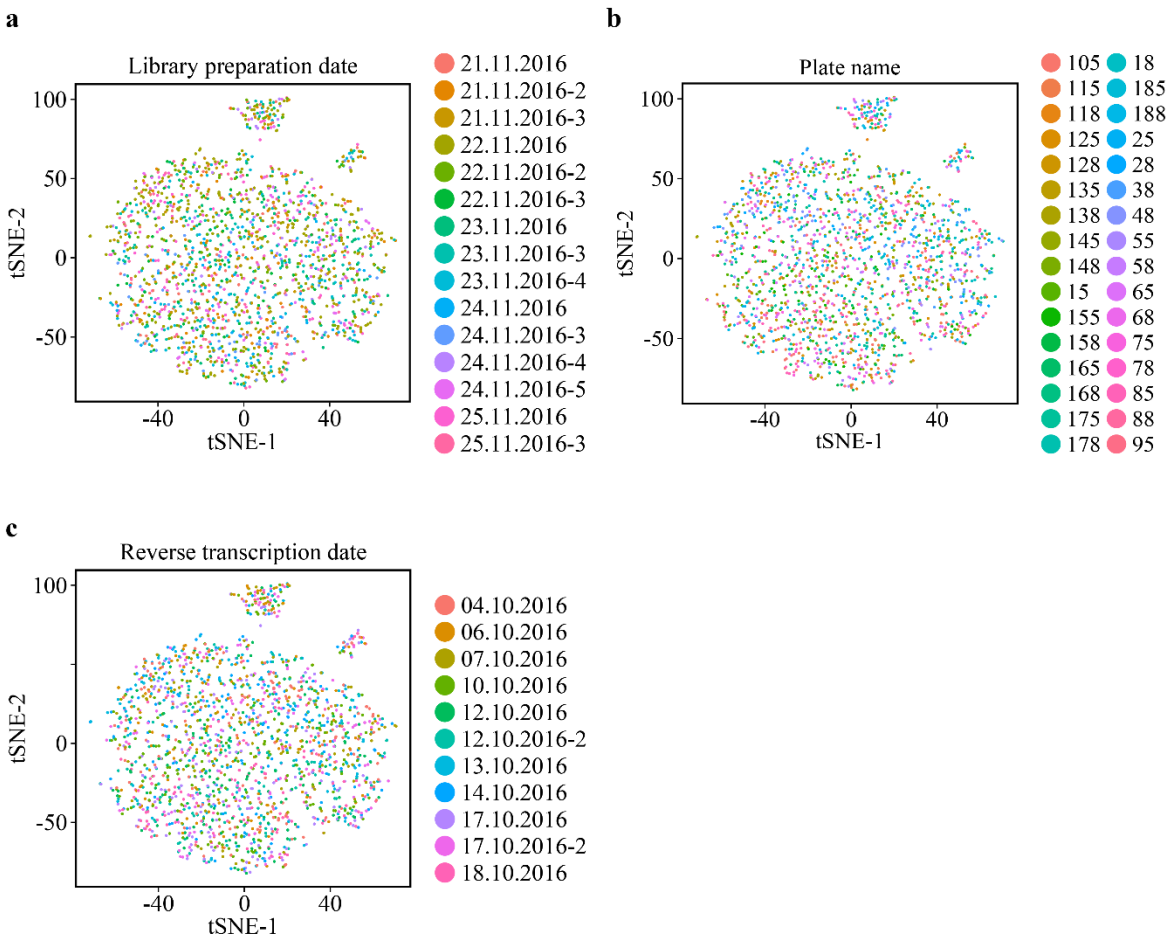
Supplementary Figure 14. Selected genes overexpressed in AST5. AST1 (red), AST2 (green), AST3 (turquoise), AST4 (blue) and AST5 (purple) are delineated with dotted lines. Expression data (Seurat normalized counts) is shown on a linear scale.



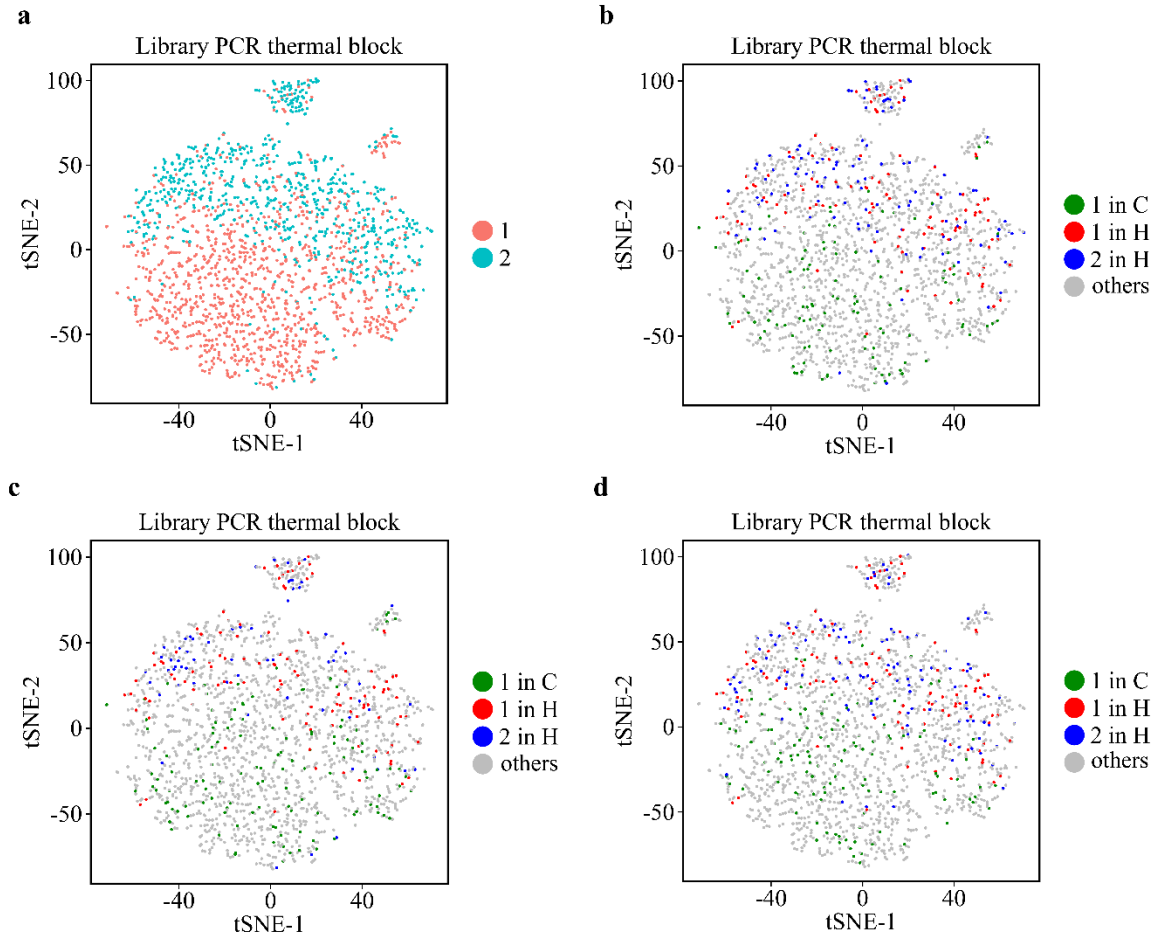
Supplementary Figure 15. Controlling for experimental batch effects on Seurat clustering of astrocyte data (part 1): Batch effects were not driven by (a) the use of independent tissue preparations from cortex or hippocampus (first cortical sample, 1C; second cortical sample, 2C; first hippocampal sample, 1H; second hippocampal sample, 2H). However, a prominent segregation driven by the region of sample origin was observed (Figure 1e). Batch effects driven by (b) the date of cDNA clean-up with Ampure beads, (c) the delay in days between reverse transcription and pre-amplification or (d) the date on which the concentration of the cDNA was measured were not observed.



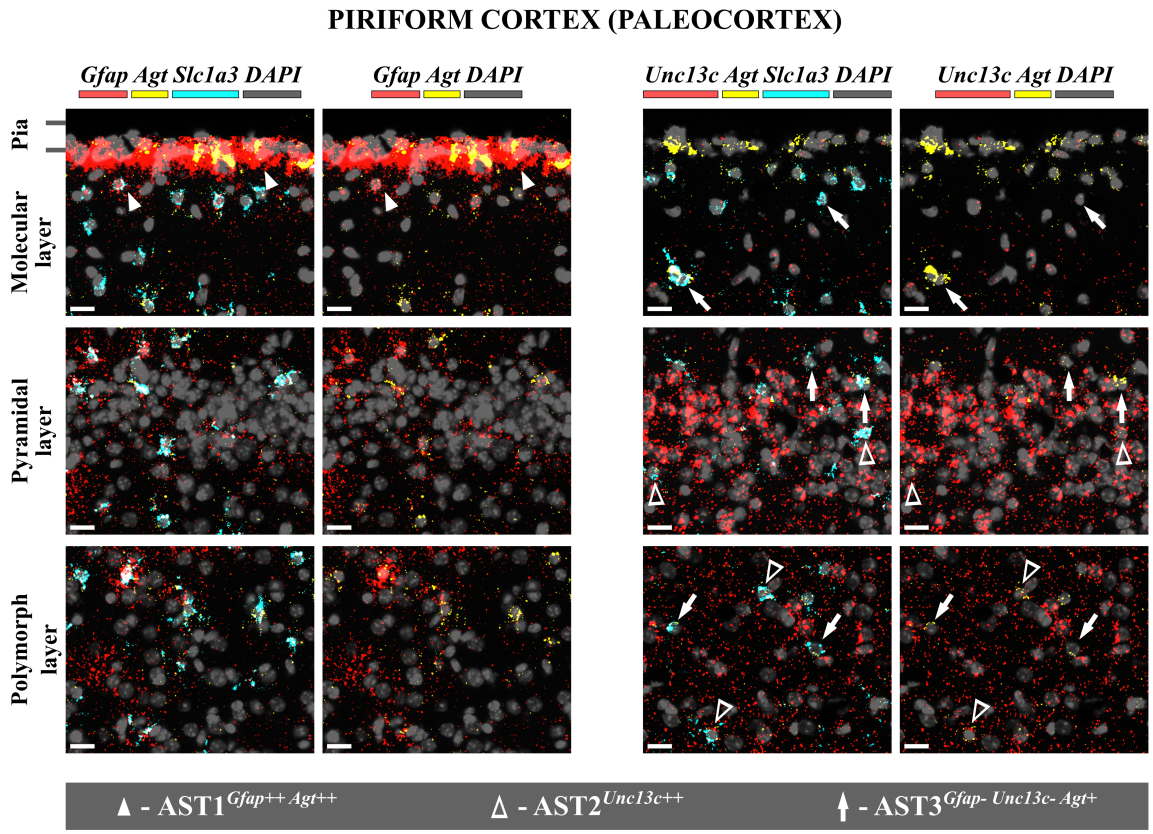
Supplementary Figure 16. Controlling for experimental batch effects on Seurat clustering of astrocyte data (part 2): Batch effects were not driven by (a) the number of freeze-thaw cycles experienced by the template switching oligonucleotide (TSO), (b) the independent clean up (Ampure beads) and sequencing (NextSeq500) of pooled libraries, (c) the use of separate PCR thermal blocks for reverse transcription or (d) the use of independent Nextera XT index kits.



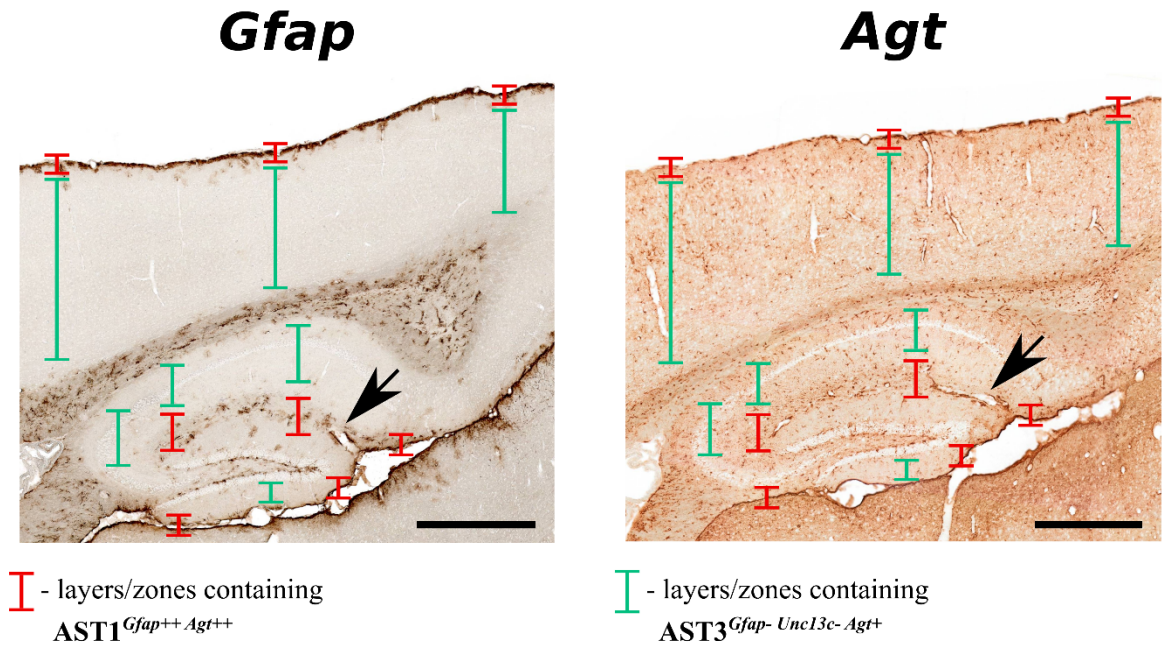
Supplementary Figure 17. Controlling for experimental batch effects on Seurat clustering of astrocyte data (part 3): Batch effects were not driven by (a) the date on which the library preparation was performed, (b) the unique plate in which library preparation was performed (including all procedures from FACS until library pooling) or (c) the date of reverse transcription.



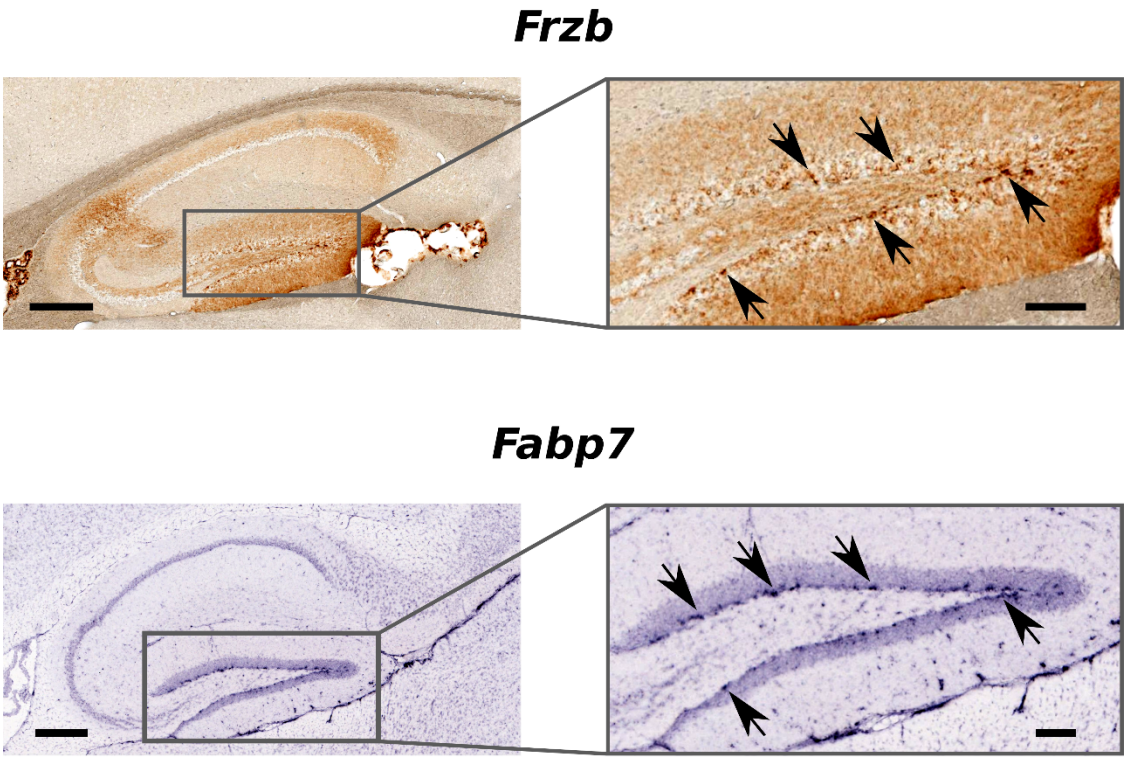
Supplementary Figure 18. Controlling for experimental batch effects on Seurat clustering of astrocyte data (part 4): Apparent sample segregation correlated with the use of independent PCR thermal blocks (1 or 2) for library preparation (a) was actually driven by the brain regions from which the samples originated (Figure 1e) and not due to technical reasons. The distribution of libraries from random sets of plates prepared from hippocampal (H) or cortical (C) astrocytes were plotted (b-d) with no obvious pattern linked to the use of different PCR thermal blocks emerging.



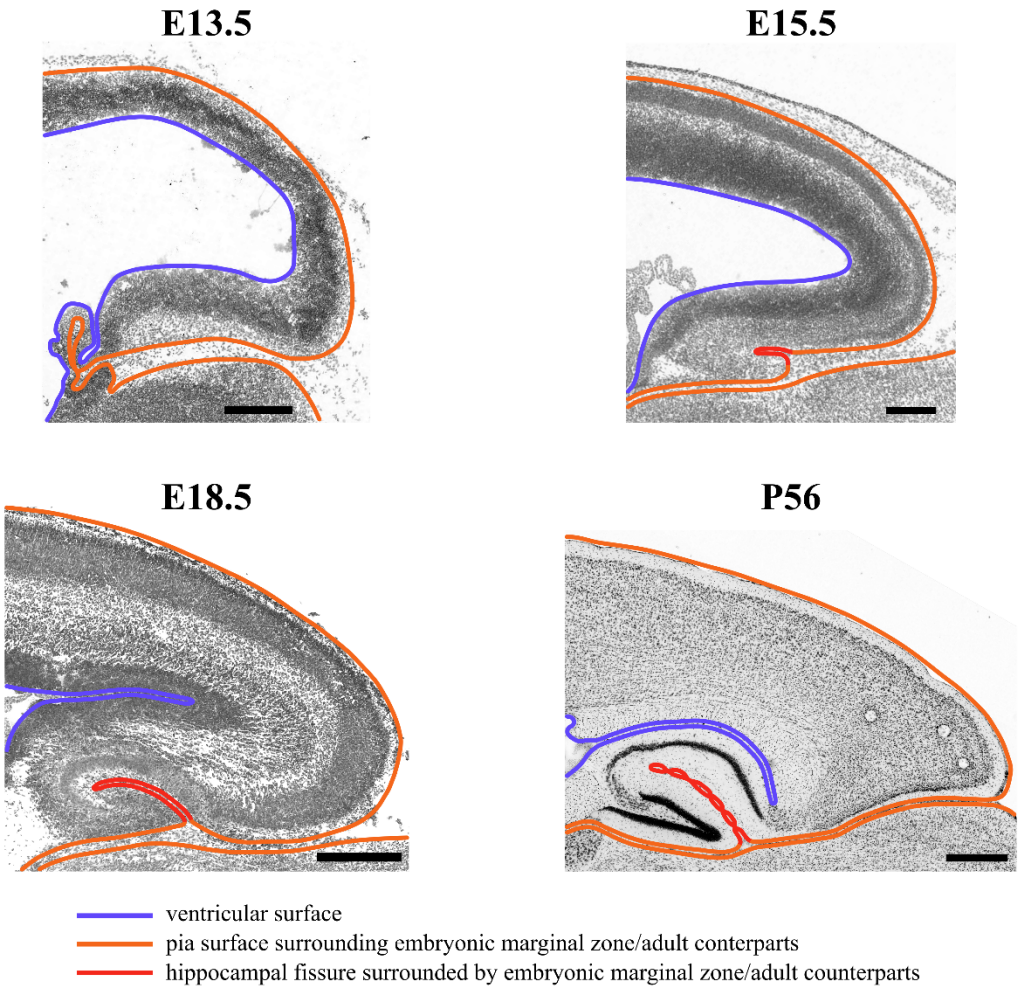
Supplementary Figure 19. Differential positioning of AST1, AST2 and AST3 in the piriform cortex. *In situ* hybridization (ISH) experiments to determine the spatial distribution of astrocyte subtypes used *Slc1a3* (as a validated marker for astrocytes in the cortex and hippocampus) (Figure 1). Astrocytes belonging to AST1 were defined by high expression of *Gfap* (++) and *Agt* (++) (closed arrowheads, left panels). *Gfap* expression was largely limited to the pial surface, whereas expression of *Agt* in astrocytes was broader, spreading from the molecular layer through to pyramidal and polymorph layers. Astrocytes belonging to AST2 were detected in the pyramidal and polymorph layers as cells expressing higher levels of *Unc13c* (++) (open arrowheads, right panels). Astrocytes belonging to AST3 were detected in the molecular, pyramidal and polymorph layers as cells expressing the marker gene *Agt* and showing little (or no) expression of *Unc13c* or *Gfap* (arrows, right panels). Scale bars, 20 µm.



Supplementary Figure 20. Mapping the spatial distribution of AST1 and AST3 subtypes using the GENSAT database. The GENSAT database[7, 8] contains images of tissue sections from BAC transgenic animals, in which gene expression patterns are revealed by eGFP expression under endogenous promoter control. Low resolution sagittal sections were retrieved showing the eGFP expression pattern from the *Gfap (left)* and *Agt* (right) promoters, respectively. Based on high expression levels of *Gfap* and expression of *Agt,* AST1 was mapped to layer 1 of cortex, the subpial zone in the hippocampus and the stratum lacunosum-moleculare around the hippocampal fissure (red bars). The initial part of the hippocampal fissure is indicated (black arrow). AST3 was mapped to cortical layers 1-6, and stratum radiatum, stratum pyramidale, stratum oriens and dentate gyrus in the hippocampus. Note that both markers (*Gfap* and *Agt*) are also highly expressed in white matter astrocytes, which were not sequenced in this study. Scale bars, 500 µm.



Supplementary Figure 21. AST4 maps to the hippocampus. *Frzb* and *Fabp7* were identified as markers of AST4. Based on gene expression data (GENSAT atlas), *Frzb* expression maps to several regions of the hippocampus, with the highest expression in cells of the subgranular zone of the dentate gyrus (top). Based on *in situ* hybridization data from the Allen Mouse Brain Atlas, *Fabp7* is highly expressed in the subgranular zone of the dentate gyrus (bottom). Based on this expression pattern, and the known use of *Fabp7* as a marker of adult neural stem cells, we suggest AST4 represents a population of such cells. Arrows indicate cells expressing the genes of interest. Scale bars, 300 µm (left); right, 100 µm (right).



Supplementary Figure 22. Development of the hippocampal fissure and its marginal zone. Cortex and hippocampus both develop from the embryonic pallium. At E13.5, the hippocampal fissure is absent. The first signs of cell migration and fissure formation appear at E15.5. During later development (E18.5 and into adulthood) the hippocampal fissure separates into the CA1 and CA2 regions, as well as the dentate gyrus. Sagittal sections are displayed. E, embryonic day; P, post-natal day. Scale bars, E13.5 – 280 µm; E15.5 – 220 µm; E18.5 – 330 µm; P56 – 525 µm. Adapted from the Allen Developing Mouse Brain Atlas.

Supplementary Tables.

Cell type	Cortex, % of cells	Hippocampus, % of cells	Cortex and hippocampus, % of cells
Astrocytes	**94.2**	**85.2**	**90.0**
Neurons	0.0	1.8	0.8
Oligodendrocytes	0.2	0.7	0.4
Microglia	0.6	0.5	0.5
Endothelium	1.1	1.9	1.5
Mural cells	3.9	9.9	6.8

Supplementary Table 1. Representation of different brain cell types in FACS sorted cells. Cell types were identified by high level Seurat clustering and expression of cell type specific markers (listed in Supplementary Figure 9). The vast majority of isolated cells were astrocytes. Mural cells represented the single biggest contaminating cell type.

AST	Number of cells	% of total cells	% subtype specific cells obtained from cortex	% subtype specific cells obtained from hippocampus
1	661	36.5	17.6	82.4
2	573	31.6	**94.4**	5.6
3	460	25.4	68.0	32.0
4	91	5.0	2.2	**97.8**
5	26	1.5	69.2	30.8

Supplementary Table 2. Distribution of astrocytes between the identified subtypes.

AST	Category	Term	Gene count	p-value	Benjamini
1	-	-	-	-	-
2	GOTERM_BP_DIRECT	GO:0008202~steroid metabolic	6	3.71E-04	0.15
	GOTERM_BP_DIRECT	GO:0006629~lipid metabolic process	12	7.36E-04	0.19
3	GOTERM_CC_DIRECT	GO:0005902~microvillus	4	2.46E-04	0.02
	GOTERM_BP_DIRECT	GO:0032870~cellular response to	5	6.45E-07	2.67E-04
4	GOTERM_BP_DIRECT	GO:0006412~translation	48	3.40E-14	7.21E-11
	GOTERM_MF_DIRECT	GO:0098641~cadherin binding	35	8.72E-10	2.71E-07
	GOTERM_BP_DIRECT	GO:0001731~formation of translation	8	5.63E-05	0.02
	GOTERM_CC_DIRECT	GO:0070062~extracellular exosome	171	1.69E-28	8.27E-26
	GOTERM_MF_DIRECT	GO:0044822~poly(A) RNA binding	105	2.90E-17	1.80E-14
	GOTERM_CC_DIRECT	GO:0043209~myelin sheath	30	4.73E-11	5.78E-09
	GOTERM_CC_DIRECT	GO:0005925~focal adhesion	42	1.05E-10	1.03E-08
	GOTERM_BP_DIRECT	GO:0006457~protein folding	22	1.07E-08	1.14E-05
	GOTERM_CC_DIRECT	GO:0031012~extracellular matrix	24	1.13E-07	4.25E-06
	GOTERM_CC_DIRECT	GO:0042470~melanosome	15	1.01E-05	3.28E-04
	KEGG_PATHWAY	mmu04145:Phagosome	13	7.38E-04	0.07
	GOTERM_BP_DIRECT	GO:0008284~positive regulation of	24	7.59E-04	0.14
	GOTERM_CC_DIRECT	GO:0043005~neuron projection	25	1.05E-03	0.03
	GOTERM_CC_DIRECT	GO:0043197~dendritic spine	13	1.68E-03	0.04
	GOTERM_CC_DIRECT	GO:0043025~neuronal cell body	27	2.31E-03	0.04
	GOTERM_CC_DIRECT	GO:0005874~microtubule	21	2.66E-03	0.05
	GOTERM_CC_DIRECT	GO:0000785~chromatin	11	8.74E-03	0.12
5	GOTERM_CC_DIRECT	GO:0005743~mitochondrial inner	34	5.30E-14	3.63E-12
	KEGG_PATHWAY	mmu00190:Oxidative phosphorylation	17	6.71E-07	2.95E-05
	KEGG_PATHWAY	mmu00010:Glycolysis /	9	3.21E-05	9.40E-04
	GOTERM_CC_DIRECT	GO:0000421~autophagosome	4	7.68E-03	0.12
	GOTERM_CC_DIRECT	GO:0005739~mitochondrion	86	2.80E-22	7.68E-20
	GOTERM_CC_DIRECT	GO:0070062~extracellular exosome	89	2.67E-18	3.66E-16
	GOTERM_CC_DIRECT	GO:0043209~myelin sheath	26	1.54E-15	1.42E-13
	GOTERM_BP_DIRECT	GO:0055114~oxidation-reduction	31	3.77E-08	4.49E-05
	GOTERM_CC_DIRECT	GO:0005777~peroxisome	10	3.64E-04	0.01
	GOTERM_CC_DIRECT	GO:0005925~focal adhesion	16	1.61E-03	0.04
	GOTERM_CC_DIRECT	GO:0015629~actin cytoskeleton	12	1.71E-04	5.83E-03
	GOTERM_CC_DIRECT	GO:0031012~extracellular matrix	11	9.88E-04	0.02

Supplementary Table 3. Selected categories obtained from functional enrichment analysis of overexpressed genes using DAVID for the identified astrocyte subtypes. A cutoff ≤ 0.1 was applied to the p-value (EASE score, modified Fischer's exact test): An arbitrary cutoff ≤ 0.2 was used for Benjamini-Hochberg (false discovery rate; FDR) corrected p-values. The full list of categories identified is reported in Supplementary File 3.

AST	Category	Genes	N genes
1	Synaptogenesis and neurite outgrowth	*Nrxn1, Prex2, Plekhb1*	3
	Neural tissue development	*Id3, Nnat, Plekhb1, Selenop*	4
	Iron metabolism	*B2m*	1
	Synaptic plasticity	*Agt*	1
	Transcription activators, repressors and	*Id3*	1
	Ubiquitination and proteasome	*Fbxo2*	1
	Circadian clock	*Id3*	1
	Glutamatergic neurotransmission	*Arl6ip1*	1
2	Transcription activators, repressors and regulators	*Xbp1, Zscan26, Tagln3, Ccnt2, Per3, Flcn, Srsf2, Wfs1, Hmgn3, Tsc22d3, Abhd14b, Chd9*	12
	Glutamatergic neurotransmission	*Slc3a2, Slc7a10, Gria2, Slc38a3, Slc25a18*	5
	Insulin signaling	*Akt2, Ogt, Rab31*	3
	Insulin-like growth factor signaling	*AI464131, Igfbp2, Akt2, Itga6, Igfbp2*	5
	Thyroid hormone metabolism and	*Dio2, Slc3a2, Slco1c1*	3
	Synaptogenesis and neurite outgrowth	*Slitrk2, Slc39a12, Lrtm2, Akt2, Tnik, Sema4b, Pmp22*	7
	Synaptic plasticity and memory	*Cd47, Hrh1, Egfr, Ndrg4*	4
	Neural tissue development	*Ndrg4, Bcar3, Wls, Chrdl1, Mfsd2a, Hmgn3, Akt2*	7
	Circadian clock	*Hrh1, Ogt, Per3*	3
	Toll-like receptor signaling	*Tril, Mertk, Tlr3*	3
	Ubiquitination and proteasome	*Fbxo44, Ttc3, Znrf3, Tmem129*	4
	Ion channels	*Kcnk1, Tmc7, Ttyh2*	3
	Wnt signaling	*Wls, Znrf3, Tnik*	3
	BMP signaling	*Bmp2k, Chrdl1*	2
	Notch signaling	*Bmp2k*	1
	Purinergic signaling	*Adora2b*	1
	Blood brain barrier	*Mfsd2a*	1
3	Transcription activators, repressors and	*Etv5, Hes5, Il33, Cyr61*	3
	Development	*Adgrg1, Fam20c, Hes5*	3
	Cell adhesion	*Adgrg1, Cyr61, Fam20c, Vcam1, Spon1*	5
	Synapse organization and neurite	*Lrtm2, Etv5, Spon1*	3
	GABAergic neurotransmission	*Gabrg1*	1
	Carbohydrate metabolism	*Pygm, Pgm1, Ppp1r3g, Chst10*	4
	Insulin-like growth factor signaling	*Igfbp2, Cyr61*	2
	Ion channels	*Kcnj16*	1
	Synaptic function	*Trim9*	1
	Lipid metabolism	*Crot*	1
	Ubiquitination and proteasome	*Trim9*	1
	Mitosis and cell cycle regulation	*Sept9, Stat3*	2
	Immediate early genes	*Fos, Jun, Dusp1*	3

Supplementary Table 4. Manually assigned functions for genes overexpressed in AST1, AST2 and AST3.

AST	Category	Genes	N genes
4	Neural tissue development	*Lpar1, Sirt2, Lamb2, Id4, Ptprs, Dab1, Kif3a, Ascl1, Gfap, Ywhae, Ywhaq, Ywhaz, Cd9, Ndn, Fabp7, Fezf2, Ppp1ca, Cdh4, Padi2, Atxn10, Emx1, Smarcb1, H3f3b, Ywhah, Itm2c, Ephb1, Arf4, Eif5a*	28
	Mitosis and cell cycle regulation	*Sirt2, Sept2, Emp2, Sdcbp, Ran, Nudc, Mapre1, Hsp90ab1, Rhoa, Rack1, Ranbp1, Rps3, Rbbp4, Ccnd2, Calm2, Ndn, Pebp1, Ppp1ca, Btg2, Smarcb1, H2afz, Sumo1, Wdr1, Nap1l1, Pdpn, Hmgb1, Maged1, Eif5a, Ybx1, Phgdh, Cdk4, Gnai2, Nfia, Smc3, Arl3*	35
	Transcription activators, repressors and regulators	*App, Hmgn1, Sirt2, Id4, Ascl1, Puf60, Hnrnpk, Arl2bp, Hes6, Hmgb1, Cxxc5, Maged1, Zbtb20, Ybx1, Hmgn2, Nfia, Hsp90ab1, Ywhaq, Ywhaz, Rps14, Rps3, Rbbp4, Chchd2, Fezf2, Nono, Med25, Emx1, Btg2, Smarcb1, Sri, Tcf4, Ywhah, Txn1, Maf1, Sumo1*	35
	Telomere and telomerase function and maintenance	*Hsp90ab1, Cct8, Hnrnpc, Hnrnpa2b1, Hsp90aa1, Hmgb1, Ptges3*	7
	DNA repair	*Hmgn1, Nono, H2afz, Hmgb1, Ybx1*	5
	Neurogenesis and neuron differentiation	*Lpar1, Id4, Dab1, Ascl1, Bcl2, Nudc, Ctnna1, Rhoa, Ndn, Fabp7, Btg2, Smarcb1, Tcf4, Itm2c*	14
	Genes linked to dentate gyrus and hippocampus	*Lpar1, Id4, Kif3a, Ran, Ywhae, Fezf2, Pebp1, Btg2, Ephb1*	9
	Astrocyte development	*Mapt, Dab1, Gfap, Vim, Phgdh*	5
	Oligodendrocyte development	*Lpar1, Sirt2, Ascl1, Gsn, Cd9, Kcnj10, Emx1, Wdr1, Phgdh*	9
	Cell migration	*Lpar1, Lama2, App, Arpc2, Dab1, Emp2, Ddr1, Twf1, Sdcbp, Tnc, Nudc, Ywhae, Ctnna1, Rhoa, Rack1, Cd9, Ndn, Adam15, Usp9x, Wdr1, Ephb1, Arf4, Pdpn, Ybx1*	24
	Actin cytoskeleton organization	*Lpar1, Arpc2, Emp2, Twf1, Cnn3, Dbnl, Ctnna1, Sorbs1, Gsn, Rhoa, Marcks, Capzb, Tmem47, Ywhah, Wdr1, Cdc42se2, Eif5a*	17
	Cilia organization	*Sept2, Kif3a, Gsn, Cct8, Atxn10, Dnal4, Arl3*	7
	RNA Splicing	*Puf60, Hnrnpc, Nono, Hnrnpa2b1, Hnrnpf, Srsf6, Hnrnpk, Ybx1*	8
	Circadian clock	*Ubb, Ppp1ca, Nono, Hnrnpr, Maged1, Cdk4*	6
	Synaptic plasticity, memory and learning	*Mapt, Gfap, Tnc, Ccnd2, Ppp1ca, Kcnj10, Btg2, Gria1, Arf4*	9
	Synaptic function	*Lama2, Pebp1, Kcnj10, Nono, Gria1*	5
	Ubiquitination and proteasome	*Ndfip1, Rack1, Ubb, Nedd4, Usp9x, Cacybp, Sumo1*	7
	Autophagy	*Park7, Sirt2, Mapt, Ppp1ca, Pink1, Ctsl, Anxa7*	7
	Glutamatergic neurotransmission	*Gria1, Slc38a1*	2
	Ion channels	*Gria1, Kcnj10*	2
	Immediate early genes	*Fos, Jun*	2

Supplementary Table 5. Manually assigned functions for genes overexpressed in AST4.

AST	Category	Genes	N genes
5	Transcription activators, repressors and regulators	*Ss18l2, Arl2bp, Rrn3, S100a1, Sub1, Smarca2, Taf13, Tsc22d4, Ybx1, Aip, Chchd2, Chchd3, Eef1a1, Glo1, Hsbp1, Hdgf, Id2, Med21, Nedd8, Prdx5, Ptma, Rps3, Tp53inp2, Tmbim6, Znf706, Zhx3*	26
	Neural tissue development	*Arf4, S100a1, Abi2, Cfl1, Cryab, Gpm6b, Gnb1, Id2, Pebp1, Rida, Stxbp3, Ywhae, Uqcrq*	13
	Neurite outgrowth	*Arf4, Rab10, P33monox, Rrn3, Abi2, Cfl1, Eef1a1, Gsk3b, Mgll, Plekhb1, Rhoa*	11
	Learning and memory	*Arf4, Abi2, Mgll, Dbi*	4
	Synaptic function	*Calm3, Baalc, Cript, Dbi, Gpm6b, Mgll*	6
	Formation of lamellipodia and pseudopodia, cell migration	*Carmil1, Cdc42ep4, Arpc2, Cfl1, Mapre2, Crk, Rhoa, Sdcbp, Ywhae*	9
	Lipid metabolism	*Acsbg1, Acox1, Dbi, Fabp5, Fads1, Nudt19, Pla2g16, Scp2, Mgll*	9
	Ubiquitination and proteasome degradation	*Dda1, Skp1, Nedd8, Psmb1, Psmb5, Psmb6, Ubb*	7
	Wnt signaling	*Fermt2, Gsk3b, Snx3*	3
	Cell signaling adapter proteins	*Sdcbp, Tom1l1, Ywhae*	3
	Circadian clock	*Gsk3b, Id2*	2
	GABAergic neurotransmission	*Dbi, Gabarap*	2
	pH regulation	*Ahcyl1, Ca2*	2
	Iron metabolism	*Fth1, Ftl1*	2
	EGF signaling	*Arf4*	1
	Endocanabinoid signaling	*Mgll*	1
	Molecular chaperones	*Dnaja1, Dnaja2, Hspb8, Hspa9, Hsp90aa1*	5
	Neuron differentiation and neurogenesis	*Id2, Uqcrq, Ubb, P33monox, Gng5, Ppia, Rhoa, Sod2*	8
	Mitosis and cell cycle regulation	*Mapre2, Rps3, Tom1l1, Pebp1, Chmp2a, Cfl1, Gnl3l, Rhoa, Sh3glb1, Calm3, Sdcbp, Id2, Rida*	13
	DNA repair	*Hmgb1, Ybx1, Rps3, Hist1h2bc*	4
	Chromatin organization	*Hmgb1, Hist1h2bc*	2
	Telomere and telomerase function and	*Gnl3l, Hmgb1*	2
	Cerebral cortex development	*Rhoa, Ywhae*	2
	Hippocampus development	*Ywhae, Pebp1, Uqcrq*	3

Supplementary Table 6. Manually assigned functions for genes overexpressed in AST5.

Gene name	ACD product code	Fluorescence channel	mRNA accession number
Agt	426941-C2	C2	NM_007428.3
Frzb	404861	C1	NM_011356.4
Ogt	423601	C1	NM_139144.4
Gfap	313211	C1	NM_001131020.1
Nes	313161-C2	C2	NM_016701.3
Slc1a3	430781-C3	C3	NM_148938.3
Unc13c	519021	C1	NM_001081153.1
Fam107a	519011-C2	C2	NM_183187.3

Supplementary Table 7. RNAScope probes used for *in situ* hybridisation. Details of the probes used, and the mRNA sequences against which they were designed, are given.

Supplementary Files.

Supplementary File 1. Full list of astrocyte subtype markers identified by Seurat.

Supplementary File 2. Transcript count table. Counts from 49,568 genes and 92 ERCC spike-in controls across 2,031 high quality cells that passed quality control (FASTQC and additional filtering steps shown in Supplementary Figure 7). Columns represent cells. Rows represent genes and ERCC controls.

Supplementary File 3. Full gene enrichment analysis performed using DAVID.

References:

1. Batiuk, M.Y.*, et al. J Biol Chem* **292**, 8874-8891 (2017).
2. Yang, Y., Lewis, R. & Miller, R.H. *Dev Biol* **350**, 127-138 (2011).
3. Cahoy, J.D.*, et al. J Neurosci* **28**, 264-278 (2008).
4. Zeisel, A.*, et al. Science* **347**, 1138-1142 (2015).
5. Picelli, S.*, et al. Nat Protoc* **9**, 171-181 (2014).
6. Rusnakova, V.*, et al. PLoS One* **8**, e69734 (2013).
7. Heintz, N. *Nat Neurosci* **7**, 483 (2004).
8. Gong, S.*, et al. Nature* **425**, 917-925 (2003).