Science
AAAS

Supplementary Materials for

Molecular and cellular evolution of the primate dorsolateral prefrontal cortex

Shaojie Ma†, Mario Skarica†, Qian Li, Chuan Xu, Ryan D. Risgaard, Andrew T.N. Tebbenkamp, Xoel Mato-Blanco, Rothem Kovner, Željka Krsnik, Xabier de Martin, Victor Luria, Xavier Martí-Pérez, Dan Liang, Amir Karger, Danielle K. Schmidt, Zachary Gomez-Sanchez, Cai Qi, Kevin T. Gobeske, Sirisha Pochareddy, Ashwin Debnath, Cade J. Hottman, Joshua Spurrier, Leon Teo, Anthony G. Boghdadi, Jihane Homman-Ludiye, John J. Ely, Etienne W. Daadi, Da Mi, Marcel Daadi, Oscar Marín, Patrick R. Hof, Mladen-Roko Rasin, James Bourne, Chet C. Sherwood, Gabriel Santpere, Matthew J. Girgenti, Stephen M. Strittmatter*, André M.M. Sousa*, Nenad Sestan*

†These authors contributed equally
Correspondence to: stephen.strittmatter@yale.edu (S.M.S.), andre.sousa@wisc.edu (A.M.M.S.) and nenad.sestan@yale.edu (N.S.)

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Materials and Methods

Human, chimpanzee, macaque and marmoset postmortem tissue

Human, chimpanzee, and macaque tissue samples used for snRNA-seq were obtained from the collections of the Sestan laboratory (table S1), human samples used for sn-multiome were obtained from the Girgenti laboratory (table S1) and marmoset samples were obtained from the Bourne laboratory (Monash University, Australia) (table S1). All clinical histories, tissue specimens, and histological sections were evaluated to assess for signs of disease, injury, and gross anatomical and histological alterations. Human tissues were collected following the guidelines provided by the Yale Human Investigation Committee for the Sestan or Girgenti laboratory collection. Human tissues were collected and handled in accordance with ethical guidelines and regulations for the research use of human brain tissue set forth by the NIH (http://bioethics.od.nih.gov/humantissue.html) and the WMA Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html). Appropriate informed consent was obtained, and all available non-identifying information was recorded for each specimen. No obvious signs of neuropathological alterations were observed in any of the human specimens considered and analyzed in this study. For all specimens, regions of interest were sampled from frozen tissue slabs or whole specimens stored at -80 ºC.

All experiments using nonhuman primates were carried out in accordance with a protocol approved by Yale University's Committee on Animal Research and NIH guidelines. All chimpanzees suffered sudden death with no prolonged agonal state and for reasons other than their participation in this study, and without any relation to the tissue used. The study complied with the definition of research from "NIH Research Involving Chimpanzees (NOT-OD-16-095)" regarding the use of chimpanzees or chimpanzee biomaterials. For marmosets, experiments were conducted

according to the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and were approved by the Monash University Animal Ethics Committee. Animals were obtained and housed at the National Nonhuman Primate Breeding and Research Facility (Monash University, Australia). To ensure consistency between specimens, all dissections from the same species were performed by the same person.

Tissue dissections and processing

Samples of dlPFC were dissected from four frozen adult human, chimpanzee, and macaque brains and four fresh marmoset brains (Fig. 1A and table S1) and stored at -80 ºC until used for nuclei isolation. The collected samples contained all six cortical layers (L1–L6) and a small amount of underlying white matter. Tissue samples from the brain specimens analyzed in the study were fixed in 4% paraformaldehyde and processed for histological examination using Nissl stain and immunohistochemistry. No obvious signs of neuropathological alterations were observed in any of the human and nonhuman specimens analyzed in this study. Nissl staining was also used to microscopically evaluate tissue and to histologically verify the identity of the dissected samples, which corresponded to Brodmann area 46.

Single nucleus RNA-seq nuclei isolation, microfluidic capture and cDNA synthesis

The brain cell nuclei were isolated according to our previous protocol (*10, 13, 41*). To avoid anatomical experimental bias for cell nuclei isolation, frozen samples were next finely pulverized to powder in liquid nitrogen with mortar and pestle (Coorstek #60316, #60317). All buffers were ice-cold and all reagents used for consequent nuclear isolation were molecular biology grade unless stated otherwise. 20-30 mg of pulverized tissue was added into 5 ml of ice-cold “lysis

buffer": 320 mM sucrose (Sigma #S0389), 5 mM CaCl2 (Sigma #21115), 3 mM $Mg(Ace)_2$ (Sigma #63052), 10mM Tris-HCl (pH 8; AmericanBio #AB14043), protease inhibitors without EDTA (Roche #11836170001), 0.1 mM EDTA (AmericanBio #AB00502), RNAse inhibitor (80U/ml) (Roche #03335402001), 1mM DTT (Sigma #43186), 0.1% TX-100 (v/v) (Sigma #T8787) nuclease-free water (AmericanBio #AB02123). Reagents: DTT, RNAse Protector, protease inhibitors, TX-100 were added immediately before use. The suspension was transferred to Dounce tissue grinder (15ml volume, Wheaton #357544; autoclaved, RNAse free, ice-cold) and homogenized with loose and tight pestles, 30 cycles each, with constant pressure and without introduction of air. The homogenate was strained through 40-µm tube top cell strainer (Corning #352340) pre-wetted with 1ml "isolation buffer": 1800 mM sucrose (Sigma #S0389), 3 mM Mg(Ace)2 (Sigma #63052), 10mM Tris-HCl (pH 8) (AmericanBio #AB14043), protease inhibitors without EDTA (Roche #11836170001), RNAse inhibitor (80U/ml) (Roche #03335402001), 1mM DTT (Sigma #43186), nuclease-free water (AmericanBio #AB02123). Additional 9 ml of "isolation buffer" was added to wash the strainer. Final 15 ml of solution was mixed by inverting the tube 10x and carefully pipetted into 2 ultracentrifuge tubes (Beckman Coulter #344059) onto the "isolation buffer" cushion (5 ml) without disrupting the phases. The tubes were centrifuged at 30000 x g, for 60 min at 4 ˚C on the ultracentrifuge (Beckman L7-65) and rotor (Beckman SW41-Ti). Upon end of ultracentrifugation, the supernatant was carefully and completely removed and 100 ul of resuspension buffer (250 mM sucrose (Sigma #S0389), 25 mM KCl (Sigma #60142), 5mM MgCl2 (Sigma #M1028), 20mM Tris-HCl (pH 7.5) (AmericanBio #AB14043; Sigma #T2413), protease inhibitors w/o EDTA (Roche #11836170001), RNAse inhibitor (80U/ml) (Roche #03335402001), 1mM DTT (Sigma #43186), nuclease-free water (AmericanBio #AB02123)) was added dropwise on the pellet in each tube and incubated on ice

for 15 minutes. Pellets were gently dissolved by pipetting 30x with 1ml pipette tip, pooled and filtered through 35-µm tube top cell strainer (Corning #352235). Finally, nuclei were counted on hemocytometer and diluted to 1 million/ml with "sample-run buffer": 0.1% BSA (Gemini Bio-Products #700-106P), RNAse inhibitor (80U/ml) (Roche #03335402001), 1mM DTT (Sigma #43186) in DPBS (Gibco #14190).

The nuclei samples were placed on ice and taken to Yale Center for Genome Analysis core facility and processed within 15 minutes for snRNA-seq with targeted nuclei recovery of 10,000 nuclei (4 reactions/technical replicates for sum of 40,000 per biological sample), respectively, on microfluidic Chromium System (10x Genomics) by following the manufacturer's protocol (10x Genomics, CG000183_Rev_A), with Chromium Single Cell 3′ GEM, Library & Gel Bead Kit v3, (10x Genomics #PN-1000075) and Chromium Single Cell B Chip Kit (10x Genomics #PN-1000074), Chromium i7 Multiplex Kit (10x Genomics #PN-120262) on Chromium Controller (10x Genomics). Due to limitations imposed by source RNA quantity, cDNA from nuclei was amplified for 14 cycles.

Single nucleus RNA-seq library preparation and sequencing

Post cDNA amplification cleanup and construction of sample-indexed libraries and their amplification followed manufacturer's directions (10x Genomics, CG000183_Rev_A), with the amplification step directly dependent on the quantity of input cDNA.

In order to reach sequencing depth sufficient for characterization of all cells, we targeted 30,000 raw reads per nucleus. Single nucleus libraries were run using paired end sequencing with single indexing on the HiSeq 4000 platform (Illumina) by following manufacturer's instructions

(Illumina; 10x Genomics, CG000183_Rev_A). To avoid lane bias, multiple uniquely indexed libraries were multiplexed and distributed over several lanes.

Single nucleus multiome: nuclei isolation, microfluidic capture and cDNA synthesis

The brain cell nuclei isolation is a comparable modification of our previous protocol (*10, 13, 41*) with yield improvement for limited tissue and additional removal of debris. It has been additionally tested and adapted for Chromium Single Cell Multiome ATAC + Gene Expression (10x Genomics PN 1000283). Small amounts of sectioned dlPFC tissue (up to 20 mg) were added into 1 ml of ice-cold lysis buffer ("Buffer A" is 250 mM sucrose (Sigma #S0389), 25 mM KCl (Sigma #60142), 5 mM MgCl2 (Sigma #M1028), 10mM NaCl (Sigma #71386), 10 mM Tris-HCl (pH 7.4) (AmericanBio #AB14043; Sigma #T2413), protease inhibitors w/o EDTA (Roche #11836170001), RNAse inhibitor (80 U/ml) (Roche #03335402001), 1mM DTT (Sigma #43186) , 1% BSA (m/v) (Gemini Bio-Products #700-106P), 0.1% NP-40 (v/v) (Sigma #74385), 0.1% Tween-20 (v/v) (Bio-Rad #166-2404), 0.01% Digitonin (m/v) (Thermo Fisher #BN2006), nuclease-free water (AmericanBio #AB02123). DTT, RNAse Protector, protease inhibitors, and all detergents were added immediately before use. The suspension was transferred to 2ml Dounce tissue homogenizer (Sigma #D8938) and lysed with constant pressure and without introduction of air, with pestle A (30x) and pestle B (30x). The homogenate was strained through pre-wetted 40 um tube top cell strainer (Corning #352340). All subsequent centrifugation was performed in a refrigerated, bench-top centrifuge with swing-out rotor (Eppendorf #5804R, #S-4-72). All lysates were centrifuged at 1000g, 10 min, 4 ˚C, pellets were saved, and resuspended in 0.4 ml "Buffer B" ("Buffer B" is "Buffer A" w/o detergents).

Final 0.4 ml of solution was mixed with 0.4 ml (1:1) of “Buffer C” ((iodixanol 50% (v/v) (Axis-Shield #1114542), 25 mM KCl (Sigma #60142), 5 mM MgCl2 (Sigma #M1028), 10mM NaCl (Sigma #71386), 10 mM Tris-HCl (pH 7.4) (AmericanBio #AB14043; Sigma #T2413), protease inhibitors w/o EDTA (Roche #11836170001), RNAse inhibitor (80 U/ml) (Roche #03335402001), 1mM DTT (Sigma #43186), 1% BSA (m/v) (Gemini Bio-Products #700-106P), nuclease-free water (AmericanBio #AB02123)). The suspension (25% iodixanol final) was mixed 10x head over head and overlayed on 0.6 ml of 29% iodixanol cushion (appropriate mix of Buffer “B” and “Buffer “C”). The tubes were then centrifuged at 3000g, for 30 min at 4 ˚C.

Following centrifugation, the supernatants were then carefully and completely removed, 0.5 ml of “Buffer D” ((10mM Tris-HCl (pH 7.4) (AmericanBio #AB14043; Sigma #T2413), 10mM NaCl (Sigma #71386), 3mM MgCl2 (Sigma #M1028), 1mM DTT (Sigma #43186), RNAse inhibitor (1000U/ml) (Roche #03335402001), 0.01% (v/v) NP-40 (Sigma #74385), 0.01% (v/v) TWEEN-20 (Bio-Rad #166-2404), 0.001% (m/v) Digitonin (Thermo Fisher #BN2006), 1% (m/v) BSA (Gemini Bio-Products #700-106P), nuclease-free water (AmericanBio #AB02123)) with P1000 tip was added and resuspend gently 10x with 400 microliter set-up on P1000, without introduction of air. Leave on ice, undisturbed, for 5 minutes.

Additional 1.0 ml of “Buffer E” was added (10mM Tris-HCl (pH 7.4) (AmericanBio #AB14043; Sigma #T2413), 10mM NaCl (Sigma #71386), 3mM MgCl2 (Sigma #M1028), 1mM DTT (Sigma #43186) RNAse inhibitor (1000U/ml) (Roche #03335402001), 0.1% (v/v) TWEEN-20 (Bio-Rad #166-2404) 1% (m/v) BSA (Gemini Bio-Products #700-106P), nuclease-free water (AmericanBio #AB02123)) and resuspended gently 10x with 900 microliter set-up on P1000, without introduction of air. Left on ice, undisturbed, for 5 minutes. Filtered (pipetted directly) through 35 micrometer tube top cell strainer pre-wetted with “Buffer E”. 20 microliter sample

was taken for counts on hemocytometer. The suspension was centrifuged at 1000g, for 10 min at 4 ˚C on the same centrifuge and rotor. The supernatant was carefully and completely removed with pipettes in series: P1000, P200, P20 (last 10 microliters). Small volume of "Buffer F" was added immediately (1X Nuclei Buffer (from 20X) (10x Genomics #2000153/2000207), 1mM DTT (Sigma #43186), RNAse inhibitor (1000U/ml) (Roche #03335402001), Nuclease-Free Water (AmericanBio #AB02123)) based on expectation in first counts and pipetted gently 10x with P200 pipette tip, with set-up on P200 smaller than volume to not introduce air, then filtered (pipetted directly) through 35 micrometer tube top cell strainer pre-wetted with Buffer F into the new 1.5 ml polypropylene tube. The minimal sample (2-5 microliters) was diluted 10x with Buffer F for the second counts on hemocytometer and adjusted the nuclei concentration to 3.3 million per milliliter with "Buffer F". Nuclei were used in Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Kit (10x Genomics #PN-1000283) as soon as possible with recovery target of 10000 nuclei (used 5ul final nuclei solution) by following the user guide (10xGenomics CG000338 Rev E).

Single nucleus multiome: sequencing library preparation and sequencing

Library and sequencing were performed by following the user guide (10xGenomics CG000338 Rev E) on NovaSeq 6000 platform (Illumina).

Consensus gene annotation models across species

Since existing gene annotation structures in chimpanzees, macaques and marmosets are relatively incomplete as compared to those in humans, we sought to generate consensus annotation models through the whole genome alignment. As a first step, we downloaded the human (hg38) annotation

file in GTF format from GENCODE (v28), and transformed the names of unplaced scaffolds, unlocalized scaffolds and alternative haplotypes to match the names from UCSC Genome Browser. We next downloaded the chain files of the whole-genome alignments between humans and chimpanzees (panTro6), between humans and macaques (rheMac10), as well as between humans and marmosets (calJac3) from UCSC Genome Browser. Reciprocal liftOver was subsequently used (minMatch was set to 0.1 to account for the sequence dissimilarities among different species) to locate the putative orthologs of humans in the other three species by requiring that: 1) at least half of the exons from a given human gene were aligned to the other three species; 2) aligned exons were kept in the same order along the gene body; 3) exon length changes were < 50%. 4) intron length changes were < four folds. Lastly, the exons and introns were bridged to form a pre-mRNA model for each gene in each species.

We conducted additional analyses to evaluate the reliability of the consensus gene model (fig. S2): 1) We compared the average genome mappability scores calculated via GenMap (*61*) between species differentially expressed genes and background genes, and found a very similar distribution . 2) We analyzed the data based on publicly available gene annotations from NCBI RefSeq and found that the major findings in this study are robust to this alternative gene model.

Single-nucleus RNA-seq data processing, filtering and normalization

Based on the consensus gene models, we generated for each of the four species a customized reference genome which incorporated the associated gene annotation information to run the CellRanger (version 3.1.0) pipeline including the read alignment, barcode counting and unique molecular identifier (UMI) quantification, resulting in the filtered raw count matrices containing gene expression in all inferred cells for downstream quality control.

We exploited several steps to preprocess our data. First, we conducted an initial clustering for each individual of each species, and later removed low-quality clusters with low UMIs and/or mitochondrion content, and with no meaningful cluster identities. Second, the remaining cells were re-clustered, and doublet scores were calculated via the scrublet algorithm (*47*). Third, to avoid the false positives, a custom script was applied in order to spot biologically meaningful clusters which were falsely assigned as doublets by scrublet. Specifically, we calculated the AUROC scores, which summarized expression of predefined sets of cell type markers, to infer cell type identities by the R package AUCell (*48*). Cells assigned to more than two cell types were considered as putative doublets. Only the clusters predicted to have high doublet scores using both methods were regarded as doublets, and were excluded in the downstream analyses. Although current doublet-detection methods are sensitive to identify heterotypic doublets (i.e., doublets formed from different cell types), they are not sensitive enough to detect homotypic doublets (i.e., doublets formed from the same cell types). Homotypic doublets are expected to have much higher number of UMIs compared to other cells in the same cell type, a feature that can be harnessed to filter putative homotypic doublets. Fourth, to filter putative homotypic doublets or outliers, we excluded cells with extreme contents of UMIs (three-fold standard deviations away from the average UMI) for each major cell type. Lastly, to increase our power in identifying low-quality cells and outliers, cells from the same species were further integrated (see below), and low-quality cells and doublets were removed as mentioned above. Gene expression within each cell was then normalized using the *NormalizeData* function in the R package Seurat (scaling factor = 10000) (version 3.1.0) (*49, 62*).

Transcriptomic integration of the snRNA-seq data from the four primates

To embed all nuclei from the four primate species in the same low dimension space, and subsequently visualize them on the UMAP (Fig. 1B), we integrated the data using Seurat (*49*). Specifically, for each individual out of the 16 individuals from the four species, we identified the top 2,500 highly variable genes using the variance-stabilizing transformation (*FindVariableFeatures* function) implemented in the R package Seurat and selected the top 2,500 highly variable genes for integration using *SelectIntegrationFeatures* function. We subsequently identified the anchors across datasets using the function *FindIntegrationAnchors* (set *anchor.features* as the calculated integration highly variable genes) based on the top 30 dimensions from canonical correlation analysis (CCA), followed by hierarchical integration of normalized data using the *IntegrateData* function. Then we scaled the integrated data, performed principal components analysis (PCA) and selected significant principal components with elbow plots. These principal components were used for downstream UMAP visualization. Note that we extracted cells from each major subtype classes, clustered and annotated them independently, and lastly mapped the annotation back onto the global UMAP (see below).

To validate the correspondence of homologous subclasses between different species, we further integrated data using other independent methods: MNN and Harmony (*63, 97*). In the MNN integration, we calculated the highly variable genes as described above, followed by batch correction on the normalized data using the *fastMNN* function (k = 16) in the R *batchelor* package. The data was first merged within each species followed by merging across species. In the Harmony method, we used the same highly variable genes, scaled the normalized data within each individual followed by performing principal component analysis. The top 40 principal components were then passed to the *RunHarmony* function (lambda = 0.8) in the R *harmony* package. The resulted low

dimension components from both methods were passed to the *RunUMAP* function to generate UMAP layouts with "umap-learn" method and "correlation" metric.

Single-nucleus multiome data processing, filtering and normalization

Cellranger-arc 2.0.1 was used to build the reference genome with the same versions of the human genome (hg38) and gene annotation (GENCODE v28) as we used for the snRNA-seq data. This reference genome was harnessed for the following read alignment, barcode counting and UMI counting in Cellranger-arc. To call peaks in the ATAC assay, MACS2 (*51*) was applied for each sample with the parameters "-f BAM -g hs -q 0.05 --nolambda --nomodel" for the aligned ATAC-seq data from Cellranger-arc results. Peaks from each sample were then combined and merged if the peak occurred in at least 3 samples to form a master peak set.

The same quality control procedures as detailed above were applied to process the RNA assay. In the ATAC assay, we further removed some nuclei showing low quality (peak region fragments < 100, percentage of reads in peaks < 10, nucleosome signal > 4 or transcription start sites [TSS] enrichment < 2).

Clustering and annotation of cell subtypes

Nuclei across the four species were clustered into four major cell classes: excitatory neurons (*SLC17A7*+), inhibitory neurons (*GAD1/2*+), glia cells and non-neural cell, based on the expression of key marker genes. We *in silico* isolated the cells from each of the categories (glia cells and non-neural cells were combined into the non-neuronal cells) based on well-established markers (Fig. 1B), and then performed normalization, highly variable gene identification, data scaling, PCA and UMAP visualization, followed by clustering of the cells into subtypes using the

FindClusters function with the neighbor number chosen as 25. The identity of each subtype was assigned as detailed below:

1) For the excitatory neurons, we parcellated them into 10 subclasses representing neurons at different layers and having distinct projection properties, based on their hierarchical relationship. The layer distribution of each subclass was first determined by two ways: i) the expression of canonical layer markers, including the upper-layer excitatory neuron marker *CUX2* and *PCDH8*, L4 marker *RORB*, deep-layer markers including *BCL11B*, *FEZF2*, *HTR2C*, *SYT6*, *OPRK1*, and *CTGF*; ii) the alignment with the published human MTG snRNA-seq data which has layer distribution for each cell (*16*). Through these two methods, we assigned each of the excitatory neuron subclasses to the corresponding layers. We also validated the layer distributions by comparing our data with the human dlPFC spatial transcriptomic data as detailed below ("Transcriptomic comparisons with human dlPFC spatial transcriptomic data" section). Although the two L6 intratelencephalic subclasses in this study also mapped to some L5 cells from human MTG (fig. S11B), they were annotated as L6 neurons because they generally showed marker enrichment at L6 of the human dlPFC spatial transcriptome data (fig. S11D) and their transcriptomical counterparts in the motor cortex of humans, marmosets and mice were annotated as L6 intratelencephalic neurons (*22*). Next, the projection properties of each subclass were demonstrated by the alignment with the published scRNA-seq data from mouse where Retro-seq was used to measure the projections of excitatory neurons (*17*). This added projection information to each subclass (IT, intratelencephalic; ET, extratelencephalic neuron; NP, near projecting; CT, corticothalamic projection and L6B). Each subtype within a given subclass was finally labeled with the predicted laminar information of the subclass along with two to three conserved subtype

markers. These 40 excitatory neuron subtypes were organized into 10 subclasses representing layer and axon projection-specific excitatory neuron types: L2-3 IT, L3-5 IT-1, L3-5 IT-2, L3-5 IT-3, L6 IT-1, L6 IT-2, L5 ET, L5-6 NP, L6CT, L6B.

2) For the inhibitory neurons, we classified them into different subclasses based on the expression of classical inhibitory neuron markers including *SST*, *PVALB*, *LAMP5* and *VIP*, as well as markers highly correlated with their developmental origins (MGE: *LHX6* and *SOX6*; CGE: *ADARB2, NR2F2* and *PROX1)* (*16, 17, 19, 20, 31, 64*). This included 9 subclasses: *SST NPY*, *SST HGF*, *SST*, *PVALB*, *PVALB* chandelier cells (*PVALB* ChC), *LAMP5 LHX6*, *LAMP5 RELN*, *VIP*, *ADARB2 KCNG1*. The *SST NPY* subclass included only one subtype marked by *NPY* and *CHODL* expression and matched to mouse *Sst Chodl* subtype, representing putative long-range projection neurons (*17*). The *SST HGF* class included two subtypes (InN *SST HGF GABRQ* and InN *LHX6 HGF STON2*) marked by conserved *HGF* expression across species, but only exhibited *TH* expression in humans, macaques and marmosets. Other *SST* subtypes were considered as the *SST* subclass and were further separated by the expression of canonical *SST* subtype markers including *CALB1*, *CBLN4*, *RELN* (*16, 17, 19*). Some of them were transcriptomically aligned to mouse Martinotti and non-Martinotti cells (figs. S13C and S14C) (*16, 17, 19, 20, 31*). The *PVALB* ChC subclass consisted of two subtypes marked by *UNC5B* (*65*) and aligned with mouse chandelier cells. The rest of the *PVALB*+ inhibitory neuron subtypes were assigned to the *PVALB* subclass, largely considered as the *PVALB* basket cell subtypes. The *LAMP5 LHX6* subclass comprised of two subtypes (InN *LAMP5 LHX6 PROX1* and InN *LAMP5 LHX6 TAC1*) that were considered as MGE-derived inhibitory neurons based on previous studies (*31*) and mapped to Ivy cells in mouse hippocampus (*12, 66*). The rest of the *LAMP5*+ inhibitory neurons were assigned to the *LAMP5*

RELN subclass and largely mapped to neurogliaform cells. The *VIP* subclass was marked by evident *VIP* expression. Other inhibitory neuron subtypes were largely aligned to *Sncg*+ inhibitory neurons in mouse (fig. S13C), but exhibited very low expression of *SNCG*. We classified them as the *ADARB2 KCNG1* group based on their combinatory expression of *ADARB2* and *KCNG1* distinguishing them from other inhibitory neuron subtypes. Although we noticed some *LAMP5*+ or *ADARB2*+ layer 1 cells expressing *SST*, we did not classify them as *SST* inhibitory neurons, as they exhibited lower *SST* expression compared to other *SST* subtypes and they transcriptomically co-clustered with *LAMP5*+ or *ADARB2*+ inhibitory neurons. Each inhibitory neuron subtype was finally named with "InN"–the abbreviation of inhibitory neurons–followed by a major inhibitory neuron marker (e.g., *SST*, *PVALB*) and one or two conserved subtype markers.

3) For glia cells, we first clustered them into astrocytes (*AQP4*+), oligodendrocyte precursor cells (OPCs, *PDGFRA*+), oligodendrocytes (*MOG*+), and microglia (*P2RY12*+). Each subclass was further subdivided into different subtypes based on known and novel marker gene expression. In detail, the astrocytes were divided into interlaminar astrocytes (*GFAP*+*FABP7*+; iAstro) (*67*), protoplasmic astrocytes (*AQP4*+*SLC1A2*+; pAstro), and fibrous astrocytes (*GFAP*+*AQP1*+; fAstro) (*68*). Their putative laminar distributions were further validated by transcriptomical integration with the human MTG and motor cortex snRNA-seq data (*16, 22*). In humans and chimpanzees, we found an additional astrocyte subtype expressing *OSMR* and *TPST1*, indicating its identity of reactive astrocytes (rAstro) (*69*). OPCs and oligodendrocytes collectively formed the oligodendrocyte lineage. We classified them into OPCs (*PDGFRA*+*PCDH15*+) and committed OPCs (COPs, *GPR17*+*SOX4*+) (*70*), as well as newly formed (*MOG*+*CDH7*+; nOligo), early (*MOG*+*FRY*+; eOligo), middle (*MOG*+*OPALIN*+; mOligo) and late oligodendrocytes

(*MOG*+*GSN*+; lOligo) (*21*). Regarding microglia, although they represented an overall homogeneous cell population (*P2RY12*+*APBB1IP*+), we found one human-specific subtype with exclusive expression of *CCL3* and *CCL4* (*25*), as well as one Hominini-specific subtype with high expression of *GLDN* and *MYO1E*.

4) Non-neural cells were categorized into refined subtypes based on the expression of key marker genes. Immune cell subtypes were composed of macrophages (*F13A1*+*COLEC12*+), myeloid cells (*LSP1*+*LYZ*+), B cells (*EBF1*+*IGKC*+) and T cells (*SKAP1*+*CD247*+), all of which were characterized by well-established markers. For the endothelial cells, we subdivided them as arterial (aEndo), capillary (cEndo) and venous (vEndo), with shared expression of *CLDN5* but distinct expression of *DKK2*, *SLC7A5* and *IL1R1,* respectively (*23*). We meanwhile identified a subtype (*HBA1*+*HBB*+*KLF1*+*TFRC*+*)* with characteristics of red blood lineage cells, likely representing nucleated erythroblasts. Initial, pre-isolation, presence of nucleus in these cells was confirmed by high expression of genes encoding nuclear non-coding RNA (*MALAT1*) (*71*) and beta-2-microglobulin (*B2M*). A similar subtype was also reported in a recent adult human cerebral vasculature dataset (*72*). The mural subtypes were identified based on the expression of key marker genes as pericytes (PC) (*P2RY14*+*GRM8*+) and smooth muscle cells (SMCs), the latter of which were *ACTA2*+ and further classified as arterial (aSMC, *CNN1*+), arteriolar (aaSMC, *CYP1B1*+) and venous (vSMC, *CRISPLD2*+) (*24*). We also identified three subtypes as vascular leptomeningeal cells (VLMCs; *COL1A1*+*COL1A2*+) with high expression of *ABCA8*, *SLC13A3* and *SLC4A4*, respectively (*73*).

In analyzing the sn-multiome data, we used the human snRNA-seq data as the reference followed by Seurat label transfer (*49*) to annotate cell identity.

Hierarchical dendrogram showing subtype organizations

The hierarchical structure of all subtypes from our data was constructed in a dendrogram to mirror their taxonomy (Fig. 1B). We used the following strategy: 1) we derived the transcriptome profile of each subtype by averaging gene expression across all cells within the given subtype from the four species. 2) Within each of the three major classes (glia cells and non-neural cells were merged to one class), we extracted the subtype markers as detailed in "Detection of cell type marker genes" and calculated the inter-subtype transcriptomic similarity based on the expression correlation of these markers. 3) We calculated transcriptomic distance between each pair of subtypes by subtracting the correlation coefficients from 1 and built an hierarchical dendrogram using the 'ward.D2' algorithm. 4) The three dendrograms from the three major classes were merged to a single dendrogram using the union of subtype markers and the "merge_dendrogram" function in R.

Evaluation of subtype separability

To evaluate the reliability of subtype detection, we measured the subtype separation after randomly removing one donor in a given species. Specifically, we randomly removed one donor in a given species and performed data integration using the remaining three donors. The identity of each cell was masked followed by predicting its identity using a "neighbor-voting" strategy. Here, the k-nearest neighbors (k = 15) were calculated for each cell using the integrated principal component analysis dimensions, and the most abundant subtype label among these neighbors was considered as the predicted identity of the given cell. For each subtype, we compared the predicted subtype labels versus the reference subtype labels and used the area under curve (AUC) to assess the

reliability of subtype separation. These processes were repeated four times so that each of the four donors in a given species was removed in one calculation and the distributions of the AUC scores were visualized in figure S3A.

Evaluation of the relationship between subtype separability and data quality

To have an overview of how the cell type detection could be affected by the sequencing depth and the number of cells, we evaluated the subtype separability in the down-sampled data. We used the human dlPFC data here for demonstration. The data was subsampled by cells or genes (fig. S1C) and integrated across individuals as described in "Transcriptomic integration of cells from the four primates". We then predicted the identity of each cell in the down-sampled data following the strategy mention above "Evaluation of subtype separability". For each subtype, we compared the predicted subtype labels versus the reference subtype labels and used the area under curve (AUC) to assess reliability of the subtype separation.

Evaluation of species-specificity of species-specific subtypes

We applied the following steps to evaluate the species-specific presence of the five detected species-specific subtypes. For a given species-specific subtype, we calculated its markers separately in each of the species having this subtype, following the methods described in section "Detection of cell type marker genes". We then calculated the enrichment of the marker sets across all species using AUCell package (*48*), which uses a rank-based method and is robust to sequencing depth differences. We then visualized the enrichment score distribution on UMAP layouts in each species. Homologous cells of the species-specific subtypes are expected to be marked by high enrichment scores and be clustered together on UMAP.

Subtype abundance comparisons across species

Cell type composition comparison analysis can be limited by at least two factors: 1) low number of biological replicates; 2) dependency of frequency measures between subtypes (i.e., one subtype increases in abundance leads to the decrease of frequencies of other subtypes). To reduce such bias, we used the scCODA package (*50*) to carry out comparisons of cell subtype abundance between each pair of species. Because the abundance of some cell subclasses (e.g., oligodendrocytes and astrocytes) is sensitive to dissection bias, we performed the subtype comparisons within the following cell groups: excitatory neurons, inhibitory neurons, astrocytes, oligodendrocyte lineage, immune cells, and other non-neural cells. The false discovery rate was set at 0.2 in the analysis. In visualizing the results by networks (Fig. 2B), we set the nodes as subtypes with the sizes representing average subtype proportions in the given cell class (here glia and non-neural cells were merged to one class) across species.

Calculation of transcriptomic heterogeneity

We conducted two analytic approaches to measure the inter-cell transcriptomic heterogeneity in homologous subclasses across species. In the first method, we calculated the transcriptomic entropy in the actual data (total entropy) as well as the permutated data (unstructured entropy) and subtracted the unstructured entropy from the total entropy to define the transcriptomic heterogeneity. Specifically, we followed these procedures (fig. S9A): 1) Randomly down-sample the cells within a subclass to have a consistent number of cells across species. 2) Identify the highly variable genes (HVGs) in each species and use the union of HVGs (unHVGs) for the downstream analysis. 3) Calculate one-dimension UMAP in each species using the down-sampled data. 4) Split

the cells into m bins with equal bin widths along the UMAP axis and calculated average expression of the unHVGs in each bin. 5) Calculate Shannon entropy for each gene using the average expression obtained in step 4. 6) Permutate the gene expression 30 times and calculate the average entropy for each gene following the steps 3-5. 7) Subtract the entropy obtained in the permutated data (step 6) from the actual data (step 5) and summarize all entropy differences greater than 0.5. 8) Repeat the down-sampling 100 times and plot the distribution of entropy differences in violin plots. In this analysis, to avoid batch effects driving cell separations on the UMAP layout, the gene expression of the unHVGs were batch corrected using Seurat data integration. However, the entropy calculation still utilized the uncorrected expression values.

Here, we used the average expression along the UMAP bins to define entropy because the raw snRNA-seq data is much noisier. The Shannon entropy was calculated using the following equation:

$$H(X_j) = -\sum_{i=1}^{n} P(x_{ij})\ logP(x_{ij})$$

Here, *j* represents a gene and $H(X_j)$ denotes its entropy. For each gene *j*, its expression was parcellated into n bins based on expression levels and $P(x_{ij})$ indicates the probability of gene *j* having expression in bin *i*. In order to have comparable entropy measures across species, the same expression intervals were used to generate bins for the same gene in all the four species.

In the second approach, we measured the cell separation as the average Euclidean distances in the principal component analysis (PCA) dimensions. We followed the similar steps described above in calculating entropy differences: 1) and 2) are the same. 3) Generate 100 bootstrap replicates of the highly variable genes (80% of the genes) to avoid the influence of certain unstructured genes with high variance. 4) Perform PCA analysis across the bootstrap replicates and calculate average Euclidian distances in the first 10 PCA dimensions. 5) Repeat the down-

sampling 100 times and plot the distribution of average Euclidian distances in violin plots. Similarly, we also batch corrected the gene expression of the unHVGs in each species to avoid individual differences driving the PCA separations.

Global transcriptomic divergence across cell subtypes and species

We inspected the transcriptomic reorganization of each homologous subtype by assessing its transcriptome divergence across species. Specifically, we obtained the representative transcriptional profile of each subtype in each species by calculating the average expression across all cells from the given subtype and species for the union of top 1500 highly variable genes from all the four cell classes. For each subtype, the transcriptome divergence between each pair of species was then defined as Pearson correlation coefficient subtracted from 1 based on the transcriptional profiles of this subtype between the two species. To investigate the cross-subtype dynamics of the interspecies divergences, for each pairwise species we scaled the divergences across subtypes in each species pair to transform them into corresponding z-scores.

To ensure the robustness of the evolutionary comparisons, we examined the reliability of the observations from three perspectives. First, we selected a different set of genes for calculating the transcriptome divergence, that is, the top 2500 most highly variable genes across all cells. Second, we evaluated the effects of subtype size, namely, the number of cells within each type, by downsampling each subtype to 100 cells. A few subtypes with less 100 cells were not down-sampled and the original average expression values were used. Third, we defined the transcriptome divergence as the percent of differentially expressed genes between two given species instead of the Pearson's distance. For each of the three approaches, with other procedures being constant, we replicated the evolutionary comparisons as mentioned above.

Detection of cell type marker genes

Across the entire analyses of this study, we defined subtype marker genes by comparing gene expression levels between the given subtype and the remaining subtypes for each of the four primate species. We used the Seurat *FindMarkers* function to perform the Wilcoxon Rank Sum test, and genes with expression in more than 25% of cells in the given cell type and with the false discovery rate less than 0.01 were considered as the marker genes. In some analyses under specific circumstances, we incorporated additional requirements to adjust the criteria by further requiring a threshold of expression fold changes, expression levels, and expression percentages within other subtypes.

Detection of differentially expressed genes and species-specific genes

Across the entire analyses of this study, we detected the differentially expressed genes between pairwise species using the Wilcoxon Rank Sum test implemented in the Seurat function *FindMarkers* on the basis of the normalized data. We classified a gene as differentially expressed between two species for a certain subtype if it was expressed in more than 25% of cells of either species along with the false discovery rate less than 0.01. With regard to specialized analyses in certain subtypes, we adjusted the criteria in a context-dependent manner. For instance, we intersected the differentially expressed gene list with the marker gene list as defined in 'Detection of cell type marker genes' to identify genes with subtype-restricted differential expression.

We further searched for the species-specific changes including enriched and depleted ones by combining differentially expressed genes from all pairwise species comparisons. Genes specifically enriched/depleted in a given species for a given subtype were defined as the genes

with expression levels significantly higher/lower in this species as compared to all the other three species. Since we changed the criteria for defining differential expression in some settings, the selection of species-specific genes will be altered accordingly.

Cell type alignments within and across datasets

For the subtype alignments across different datasets, we assigned each cell in our data to a matched subtype from a published dataset via the Seurat function *TransferData* following the identification of their transferring anchors by *FindTransferAnchors* function. After all cells were properly assigned, the Sankey diagrams (river plots) were constructed using the R package *ggsankey*. For the subtype alignment with public data involving a small/medium number of subtypes, we examined the alignment patterns by calculating the Pearson correlation coefficient between every pair of subtypes, and displayed them in a heat map.

Gene ontology enrichment analysis

Across the entire analyses of this study, we uncovered the gene ontology (GO) enrichment by the Bioconductor package 'topGO' (http://bioconductor.org/packages/release/bioc/html/topGO.html) using the Fisher's exact test. GO terms under biological processes, molecular functions and cellular components were considered as significant if the enrichment p-value is < 0.05. To ensure the stability of the enriched terms, we employed another complementary approach. Specifically, we checked the enriched terms by uploading the test gene list and the background gene list to the Metascape web portal with the default parameters (*74*).

Transcriptomic comparisons with human dlPFC spatial transcriptomic data

In addition to the transcriptomic integration with human MTG data (*16*) to infer the layer distribution of dlPFC excitatory neurons, we applied two independent methods to evaluate the inferred laminar distribution of excitatory neuron subclasses: 1) assessing the expression of certain representative excitatory neuron subclass markers (*PCDH8*, *RORB*, *POU3F1*, *HTR2C*, *OPRK1*, *SYT6*, *CTGF*) in a human dlPFC spatial transcriptomic dataset (*18*). 2) assessing the enrichment of subclass marker gene sets across the cortical layers of the same spatial transcriptomic data. We extracted the top 30 markers of each subclass ranked by fold changes of average expression as described in the "Detection of cell type marker genes" section. We then scaled their expression in the spatial transcriptomic data and averaged the scaled expression for each data point. The distributions of the scaled expression were visualized in box plots.

Selection of highly subtype- and species-specific changes in excitatory neurons

Species-enriched or -depleted changes for each excitatory neuron subclass were identified as in 'Detection of differentially expressed genes and species-specific genes'. To restrict these genes to a subset which contained genes exclusively expressed in a given subclass of a certain species, we further required the genes to be differentially expressed between this subclass and any other subclasses in this species. For both within-species and cross-species comparisons, we used the *FindMarkers* function requiring a minimum expression ratio of 0.25, maximum background expression ratio of 0.15, fold changes of expression ratio $\geq$ 2.5, and fold changes of average expression $\geq$ 3.5.

Imputation of laminar distribution of dlPFC cells

In order to understand the putative laminar organization of dlPFC inhibitory neurons and how it relates to subtype identities, we imputed their laminar scores based on the published human middle temporal gyrus (MTG) data (*16*), where cells dissected at each cortical layer were captured for sequencing. Specifically, the union of the top 2000 highly variable genes from dlPFC and MTG data were selected to compute the anchors between the two datasets using the *FindTransferAnchors* function with canonical correlation algorithm in Seurat. Following that, we transferred the laminar labels from MTG data to dlPFC data via the *TransferData* function.

Enrichment of *LAMP5 LHX6* subtype markers in MGE- and CGE- derived inhibitory neurons

We extracted the markers from each subtype (InN *LAMP5 LHX6 PROX1* and InN *LAMP5 LHX6 TAC1*) of *LAMP5 LHX6* subclass using the *FindMarkers* function in Seurat, with a minimum expression ratio of 0.2, *logfc.threshold* greater than 0.2 and expression ratio fold changes no less than 1.5. We then calculated the average expression of these markers across inhibitory neuron subtypes and scaled the expression to the range of -3 to 3, that could avoid bias caused by potential outliers. The distribution of the scaled expression values in each species were later visualized according to inhibitory neuron developmental origins. We next asked if any of these markers were unique to MGE- or CGE-derived inhibitory neurons, similar to genes like *LHX6* and *PROX1*. We compared their fold changes between MGE and CGE inhibitory neuron subtypes and filtered with a threshold of 1.5. Only genes that are shown up in at least two species were kept and the top eight markers ranked by average fold changes across species were visualized in a dot plot.

Gene family enrichment analysis

Inspired by a previous study that inhibitory neuron identities are shaped by the divergent expression of different elements from the same sets of gene families (*65*), we evaluated how the gene families diverged across the four species. We downloaded the gene family information from HGNC database (https://www.genenames.org/data/genegroup/#!/), which defined gene families based on protein sequence or function similarity. To access how these gene families were differentially used between different inhibitory neuron subtypes in the same species as well as homologous inhibitory neuron subtypes across different species, we performed the gene family enrichment analysis via hypergeometric test (or one-sided Fisher's exact test). All the expressed genes were considered as the gene pool, and each set of cluster marker genes (fig. S14D) or species-enriched genes (fig. S14I) was regarded as a sampling from the gene pool. We tested for each gene family whether there were more genes than expected present in a given sampling pool using the R function *phyper*. Multiple testing adjustment of *p* values were carried using *p.adjust* function in R with "fdr" method. FDR cutoff was set at 0.1 to highlight significant enriched gene families.

Hexagon visualization of gene expression

UMAP layout visualization of gene expression is a convenient way to understand gene expression patterns. However, as cell number increase, some cells will be buried underneath, which obscuring result interpretation. Even adding transparent colors cannot solve the problem when the cell numbers are so enormous, like this dataset. Therefore, we applied hexagon visualization of gene expression overlaid on the UMAP layout using *schex* R package, which parcellates cells into thousands of bins and shows the summarized expression (i.e., mean expression) for each bin. We adopted this visualization for multiple analyses (e.g., Fig. 4D). which led to more accuracy.

Nissl staining

Adult human, chimpanzee, macaque, and marmoset dlPFC sections were mounted, dehydrated in progressive concentrations of ethanol followed by Histoclear, rehydrated in progressive dilutions of ethanol, stained with 0.1% cresyl violet acetate for 15 minutes, dehydrated in progressive concentrations of ethanol followed by Histoclear and coverslipped with Permount mounting solution (Electron Microscopy Services #17986-01). Nissl-stained sections were imaged with a Nikon Eclipse Ti2-E microscope under brightfield.

TH, SST and VIP immunohistochemistry

Antigen unmasking was performed with the Retriever 2100 (Electron Microscopy Services) using buffer A (EMS #62706-10). Sections were rinsed in PBS (2 x 5 min) and incubated in RNAscope Hydrogen Peroxide solution (ACD, Bio-Techne #322335) for 10 minutes at room temperature to quench endogenous peroxidase activity. Sections were washed in PBS (3 x 5 min) and incubated in blocking solution containing 5% (v/v) normal donkey serum (Jackson ImmunoResearch Laboratories) and 0.3% (v/v) Triton X-100 in PBS for 30 min at room temperature. Primary antibodies – TH (Abcam ab112, 1:500), SST (Santa Cruz Biotechnology sc-74556, 1:100), and VIP (Abcam ab8556, 1:2000) were diluted in blocking solution and incubated with tissue sections for 48 h at 4°C. Sections were washed with PBST (1X PBS + 0.3% Triton X-100) prior to being incubated with the appropriate fluorophore-conjugated or horseradish peroxidase (HRP)-conjugated secondary antibodies (Jackson ImmunoResearch Labs) for 30 min at room temperature. All secondary antibodies were raised in donkey and diluted at 1:250 in blocking solution. For HRP-conjugated secondary antibodies, sections were rinsed in PBST (3 x 5 min) and incubated

with TSA Plus Fluorescein (Akoya Biosciences #NEL741001KT) diluted 1:200 in 1X Plus Amplification Diluent. Sections were washed with PBST (3 x 5 min), treated with Autofluorescence Eliminator Reagent (Millipore #2160) according to manufacturer instructions, and coverslipped with Vectashield Plus Antifade Mounting Medium (Vector Laboratories #H-1000).

In situ hybridization (ISH)

For multiplex fluorescent ISH, the RNAscope Multiplex Fluorescent Reagent Kit v2 Assay was used (ACD, Bio-Techne, #323110). The manufacturer's pretreatment protocol for fixed frozen tissue was carried out with modifications (ACD, Bio-Techne, document no. 323100-USM). Antigen unmasking was performed with the Retriever 2100 (Electron Microscopy Services) using buffer A (EMS #62706-10). Sections were rinsed in PBS (2 x 5 min) and incubated in RNAscope Hydrogen Peroxide solution (ACD, Bio-Techne #322335) for 10 minutes at room temperature. Sections were washed with 0.1X PBS (2 x 5 min) and dried overnight. RNAscope multiplex fluorescent v2 detection protocol was performed with modifications (ACD, Bio-Techne, document no. 323100-USM). All washes were performed with PBST (1X PBS + 0.3% Triton X-100). Tyramide signal amplification was performed with TSA Plus Fluorescein (Akoya Biosciences #NEL741001KT) or TSA Plus Cy3 (Akoya Biosciences #NEL744001KT) diluted 1:200 in RNAscope Multiplex TSA Buffer. Following the RNAscope detection protocol, sections were treated with Autofluorescence Eliminator Reagent (Millipore #2160) and mounted as previously described. The following probes were used in our experiments: Hs-SST (ACD, Bio-Techne #310591), Hs-RELN (ACD, Bio-Techne #413051-C2), Hs-DDC (ACD, Bio-Techne #441631), Hs-TH (ACD, Bio-Techne #441651-C2), and Hs-FOXP2 (ACD, Bio-Techne #407261-C2).

For multiplex fluorescent ISH combined with immunofluorescence, the RNAscope Multiplex Fluorescent v2 Assay combined with Immunofluorescence - Integrated Co-Detection Workflow protocol was performed with modifications (ACD, Bio-Techne, document no. MK 51-150 TN, Appendix D). Antigen unmasking was performed with the Retriever 2100 (Electron Microscopy Services) using 1X RNAscope Co-Detection Target Retrieval Buffer (ACD, Bio-Techne #323180). Sections were rinsed in PBS (2 x 5 min), incubated in RNAscope Hydrogen Peroxide solution (ACD, Bio-Techne #322335) for 10 minutes at room temperature, and washed with PBS (2 x 5 min). Primary antibodies - TH (Immunostar #22941, 1:250), TH (Abcam ab112, 1:500), SST (Santa Cruz Biotechnology sc-74556, 1:100), DDC (Abcam ab3905, 1:250), or IBA1 (FUJIFILM Wako #019-1947, 1:500) were diluted in RNAscope Co-detection Antibody Diluent (ACD, Bio-Techne #323180) and incubated for 48 hours at 4° C. RNAscope multiplex fluorescent v2 Co-detection protocol was performed with previously described modifications (ACD, Bio-Techne, document no. 323100-USM). Following the RNAscope detection protocol, sections were incubated with fluorophore-conjugated or HRP-conjugated secondary antibodies diluted 1:250 in RNAscope Codetection Antibody Diluent (ACD, Bio-Techne #323180) for 30 min at room temperature. For HRP-conjugated secondary antibodies, sections were rinsed in PBS (3 x 5 min) and incubated with TSA Plus Cy3 diluted 1:200 in RNAscope Multiplex TSA Buffer. All sections were rinsed in PBST (3 x 5 min), treated with Autofluorescence Eliminator Reagent (Millipore #2160) and mounted as previously described.

Species-specific subtype RNA *in situ* hybridization

For species-specific subtype validation experiments, human and chimpanzee dlPFC multiplex ISH was performed as described above with the following probes: Hs-LAMP5 (ACD, Bio-Techne

#487691), Hs-CD8A (ACD, Biotchne #560391-C2) Hs-CUX2 (ACD, Bio-Techne #425581), Hs-PRLR (ACD, Bio-Techne #542551-C2), Hs-AQP4 (ACD, Bio-Techne #482441-C2), Hs-CHI3L1 (ACD, Bio-Techne #408121). Marmoset dlPFC multiplex ISH was performed as described above with the following modifications: to maintain tissue integrity, marmoset sections were fixed for an additional 48 hours in 10% neutral buffered formalin, subsequently washed with PBS (5 x 30min), and mounted. Antigen retrieval and protease incubation steps were replaced with a 30-minute RNAscope Custom Pretreatment reagent (ACD, Bio-Techne, #30040) incubation at 40°C. The following probes were utilized for marmoset dlPFC multiplex ISH: Cj-LAMP5 (ACD, Bio-Techne #532731-C2), Cj-AQP4 (ACD, Bio-Techne #500451-C3), Cj-CUX2 (ACD, Bio-Techne #554631-C2), Cj-PRLR (ACD, Bio-Techne #1188911-C1), Cj-CD8A, (ACD, Bio-Techne #1188921-C1), Cj-CHI3L1 (ACD, Bio-Techne #1188931-C1).

Imaging and Quantification of Tyrosine Hydroxylase (TH+) and Somatostatin (SST+) neurons

Tyrosine hydroxylase immunolabeled (TH+) cells were imaged from n=3 sections of human, macaque, marmoset, and mouse PFC. Z-stack images were acquired at 60x objective magnification using a Nikon Eclipse Ti2-E or Nikon A1 confocal microscope and the number of TH+/*SST* mRNA+ and TH+/SST+ cells counted manually.

Imaging and Quantification of TH+ and Vasoactive Intestinal Protein (VIP+) neurons

TH+ cells were imaged from n=3 sections of marmoset dlPFC. Z-stack images at 60x magnification were acquired using a Nikon Eclipse Ti2-E or Nikon A1 confocal microscope and the number of TH+/VIP+ cells counted manually.

Imaging and Quantification of species-specific subtype RNA *in situ* hybridization

For each probe pair (*AQP4*/*CHI3L1*, *CUX2*/*PRLR*, *LAMP5*/*CD8A*), three to five images were captured at 20x magnification from n=3 human, chimpanzee, and marmoset dlPFC sections on a Nikon A1 confocal microscope and the number of positive *AQP4*+, *CUX2*+, or *LAMP5*+ cells counted manually. From these positive cells, co-expression of the respective probe: *CHI3L1, PRLR,* or *CD8A* was manually counted as positive or negative and the percentage of cells expressing each marker calculated.

SST protein and mRNA sequence comparisons across species

We estimated the evolutionary age of somatostatin (SST) genes by phylostratigraphy, which employs protein sequence similarity algorithms (BLASTP, HMMER) to query the largest possible reference database to find the most evolutionarily distant species in which a sufficiently similar protein sequence is present (*75-78*). For any protein of at least 40 amino acids, the minimal evolutionary age (or the phylostratum), is indicated as the index of the NCBI taxonomic node at which the query species and subject species diverged. The evolutionary age of each taxonomic NCBI node is evaluated with TimeTree (*79*). Thus, a human protein also present in bacteria is ancient; a human protein also present in sea squirts is restricted to the Chordata phylum; a human protein also present in platypus is restricted to the Mammalia class; a human protein present also in Coquerel's sifaka is restricted to the Primate order; a human protein present only in humans is restricted to humans. We aligned protein sequences with ClustalW (*80*). We examined human SST sequence constraints in the gnomAD database (*81*) and correlated them to variant pathogenicity as a function of amino acid sequence (*82*). We aligned SST mRNA sequences via NCBI BLAST (*83*).

Evolutionary divergences in the oligodendrocyte lineage

In each of the subtypes along the oligodendrocyte lineage (excluding the newly formed oligodendrocytes because of its low abundance), we assessed their transcriptome divergence between each pair of species as the Pearson's distance on the basis of the expression of their highly variable genes (Fig. 5C, top). In parallel, we calculated the expression fold change of each gene between each pair of species, and ordered them in a descending order in each subtype (Fig. 5C, bottom), aiming to reflect the evolutionary dynamics along the oligodendrocyte lineage. The subtype-restricted genes along the oligodendrocyte lineage were identified by differential expression analyses as described in 'Detection of differentially expressed genes and species-specific genes'.

Interlaminar astrocyte heterogeneity and heterogeneity-related genes

We probed into the interlaminar astrocyte (iAstro: Astro *GFAP FABP7*) subtype heterogeneity mainly in humans as this species held more cells from this subtype as compared to other species. We re-clustered all human astrocytes using the same clustering procedure as in 'Clustering and annotation of subtypes', and detected three subclusters in the iAstro subtype. Next, we used the analysis of variance (ANOVA) to identify genes differentially expressed across the three subclusters (Bonferroni-corrected $p < 0.001$). Further, based on gene-gene expression similarities, these genes were organized into two modules by the visual inspection. Eigengene of each module was defined as the first principal component (PC1). In module two, the top 10 genes mostly correlated with the eigen gene were selected through the calculated Pearson correlation coefficients, as shown in the bottom of Fig. 5D.

FOXP2 immunohistochemistry

For the immunofluorescence staining, brain tissue samples were fixed in 4% paraformaldehyde at 4°C. Tissue sections were mounted and dried overnight. Antigen retrieval was performed using R-Buffer AG pH 6.0 (Electron Microscopy Sciences #62707-10) and Electron Microscopy Sciences Retriever according to manufacturer's instructions. Sections were washed in PBS (3 x 30 min) and incubated in blocking solution containing 5% (v/v) normal donkey serum (Jackson ImmunoResearch Laboratories), 1% (w/v) bovine serum albumin, and 0.4% (v/v) Triton X-100 in PBS overnight at 4°C. Primary antibodies (IBA1 Novos Biological #NB100-1028, FOXP2 ABCAM #ab16046, NeuN Millipore #MAB377) were diluted (IBA1 1:1,000, FOXP2 1:10,000, NeuN 1:2,000) in blocking solution and incubated with tissues sections overnight at 4°C. Sections were washed with PBS (3 x 30 min) before being incubated with the appropriate fluorescent secondary antibodies (Jackson ImmunoResearch Labs) overnight at 4°C. Secondary antibodies were raised in donkey and diluted at 1:250 in blocking solution. Sections were subsequently washed in PBS and stained for DAPI (1:10,000) for 10 minutes at room temperature. Sections were washed in PBS (3 x 5 min), dehydrated in 70% ethanol, treated with the Autofluorescence Eliminator Reagent (Millipore #2160) for 5-10 minutes, washed in 70% ethanol, washed again in PBS and cover-slipped with Vectashield (Vector laboratories, #H-1000). Sections were digitized with a Zeiss LSM 510 Meta Confocal Microscope and images were assembled in Zeiss Zen, ImageJ, Adobe Photoshop, and Adobe Illustrator.

For the immunohistochemistry of FOXP2 in primates and non-primate mammals, brain tissue samples were fixed in 4% paraformaldehyde for approximately 2 days at 4°C. Tissue sections were incubated in 1% hydrogen peroxide/PBS to quench endogenous peroxidase activity. Sections were washed in PBS (3 x 15 min) and incubated in blocking solution containing 5% (v/v)

normal donkey serum (Jackson ImmunoResearch Laboratories), 1% (w/v) bovine serum albumin, and 0.4% (v/v) Triton X-100 in PBS for 1 h at room temperature. Primary antibody against FOXP2 (Abcam #ab58599) was diluted (1:50,000) in blocking solution and incubated with tissues sections overnight at 4°C. Sections were washed with PBS (3 x 15 min) prior to being incubated with the appropriate biotinylated secondary antibodies (Jackson ImmunoResearch Labs) for 1.5 h at room temperature. Secondary antibodies were raised in donkey and diluted at 1:250 in blocking solution. Sections were subsequently washed in PBS and incubated with avidin-biotin-peroxidase complex (Vectastain ABC Elite kit; Vector Laboratories) for 1 h at room temperature. Finally, sections were washed in PBS (3 x 15 min) and signals were developed using a DAB peroxidase substrate kit according to the manufacturer's protocol (Vector Laboratories). Following washes in PBS, sections were mounted on Superfrost Plus charged slides, dried, dehydrated, and cover-slipped with Permount (Fisher Scientific). Sections were digitized using the Aperio CS2 slide scanner (Leica). Digitized images were assembled in Adobe Photoshop and Adobe Illustrator. We have performed Nissl staining in the section adjacent to the one stained against FOXP2 to delineate cortical layers. We then quantified the number of FOXP2-immunopositive nuclei in each cortical layer. Whenever possible, we also analyzed FOXP2 immunostaining in non-cortical brain regions such as striatum and used them as positive controls for the presence of FOXP2-positive nuclei.

Construction of mammalian phylogenetic tree

The comprehensive immunohistochemistry characterization of FOXP2 expression across 51 mammal cortices (table S3) prompted us to investigate the connections between the FOXP2 expression pattern and the phylogenetic relationship of these mammals. To construct the phylogenetic tree of these mammals, we inquired the *Latin* names of these species in

http://www.timetree.org/ and generated a network file ".nwk", which was then read by *read.tree* function in R *ape* package and visualized by R ggplot2 package. The expression of FOXP2 defined as the proportions of FOXP2+ nuclei was shown in a dot plot, with the species ordered according to the phylogenetic tree (Fig. 6D).

Detection of cell type-specific differentially accessible peaks

The ATAC-seq count matrices from the sn-multiome data were used for the differential peak accessibility test. Because one human donor HSB8050 (table S1) has very limited deep layer excitatory neurons (including the ones expressing *FOXP2*), which is probably attributed to dissection bias, only the rest four samples were included in this analysis and utilized to reconstructed a master peak set by merging peaks present in at least two samples. The logistic regression model implemented in Seurat was applied to find accessible peaks enriched in each cell subclass, by comparing each subclass to all other subclasses. Only peaks proximal to *FOXP2* (chromosome - 7, range – 113937174 ~ 114840607, hg38 coordinates) were tested. A minimum expression ratio of 0.025 was selected and the Bonferroni-adjusted p value threshold was set at 0.01. A threshold of minimum expression ratio fold change of 2 was selected for low abundance peaks (expression ratio < 0.1) to avoid potential noise, while the threshold was adjusted to 1.5 for peaks with high abundance (≥0.1).

FOXP2 gene regulatory network (GRN) inference via CellOracle

The ATAC-seq count matrices from the sn-multiome data were leveraged to construct a base GRN. Similarly, the sample HSB8050 (table S1) was not included because of the low numbers of deep layer neurons, which can lead to loss of peaks when constructing the master peak set. Monocle3

(*84*) was used to preprocess the data using LSI as the normalization method. Peaks and peak-to-peak co-accessibility were obtained by Cicero (*54*). Peaks overlapping a transcription start site (TSS) were annotated to the corresponding gene and only those peaks with a co-accessibility greater than 0.8 were retained. We added 11 regions with suggestive increased microglia chromatin accessibility (*85*). Regions were scanned for motifs using CellOracle's scan function (*55*), which uses the gimmemotifs motif scanner (fpr = 0.02, default motif database = gimme.vertebrate.v5.0 using binding and inferred motifs, and cumulative binding score cutoff =8) to generate an annotated peak-motif binary matrix which is the base GRN in CellOracle.

We processed the RNA-seq count matrices from the sn-multiome data using scanpy (*53*) to exclude non-expressed genes (total counts < 1) and normalize the expression per cell. We estimated gene dispersion using scanpy's filter_genes_dispersion function (with flavor='cell_ranger') and retained the top 3000 variable genes. The expression was log transformed and scaled before running a PCA. We used CellOracle's balancedKNN implementation (k = 756 nearest neighbors, b_sight = 756*8, b_maxl = 756*4) to compute neighbors in the data.

The computed base GRN was fit to each cell type using the CellOracle protocol: first, cell-type specific links were retrieved by fitting the GRN to the cell-type specific expression matrix using a bagging ridge regression model (bagging_number=20, alpha=10). The links in the resulting networks were filtered by p-value (< 0.001) and from those, only the *FOXP2*-related links were retained.

We additionally parcellated each *FOXP2*-related link into three models based on its sign (plus and minus sign of the coefficient) and direction: upstream genes positively regulating *FOXP2* (model I), downstream targets positively (model II) or negatively (model III) regulated by *FOXP2*.

Because we constructed networks in *FOXP2*-expressing cells, we did not consider the model where *FOXP2* is negatively regulated by upstream genes. For regulation model I and II, we further required that the expression ratio of a given gene in a subclass is larger than 0.1 or twice than its average expression ratios (requiring this value > 0.1) across all cell subclasses. While the first criterion is to only retain the most robustly expressed genes, the second criterion is to recover certain genes with low expression but high expression specificity. Similarly, we required genes to have expression ratios smaller than 0.1 or lower than most other cell subclasses for the regulation model III. We also computed a specificity score for each gene in each subclass defined as the fold changes of the expression in the given subclass versus the mean expression across all subclasses. A pseudo-value of 0.1 was added in both denominator and numerator to avoid 0 as denominator. In each cell type, the top genes in model I and II ranked by descending specificity scores and the top genes in model III ranked by ascending specificity scores were visualized by networks (fig. S22, A and C).

Regulon activity inference via pySCENIC

We used pySCENIC (*56*) to infer the activity of regulons in our snRNA-Seq dataset. A regulon is defined as the set of a transcription factor (TF) and its target genes. Using scanpy (*53*), we excluded from our dataset genes detected in less than 3 cells, and only kept cells expressing between 200 and 8,000 genes. The raw count matrix was analyzed with pySCENiC CLI's GRNBoost2 implementation, resulting in an adjacency table, linking TFs and their targets based on correlations between expression of genes across cells. Then, we used pySCENIC CLI's ctx to refine the targets in the regulons. We leveraged publicly available motif and TF databases from cistarget (https://resources.aertslab.org/cistarget/) to determine which target genes have an enrichment for

a corresponding motif of their TF. This allowed us to exclude indirect targets from the regulons based on the absence of cis-regulatory elements.

The activity of the regulons in each cell was scored using pySCENIC's AUCell (*48*): genes in each cell were ranked based on their expression level. Since the scoring method is ranking-based, AUCell is independent of the gene expression units and the normalization procedure. This rank was used to construct a recovery curve of all target genes in the regulon. For this, only the top ranking 5% genes were used, following pySCENIC's suggestions. The AUC of the regulon is used as the regulon activity. We also computed Regulon Specificity Scores from the activity matrix to determine the top active regulons in each of the cell types in our data via pySCENIC's regulon_specificity_scores function.

Mouse in utero electroporation experiments

All experiments using animals were carried out in accordance with a protocol approved by Yale University's Committee on Animal Research and NIH guidelines. Timed-pregnant CD-1 mice were used for in utero electroporation. Using day of vaginal plug as embryonic day 0.5, dams were anesthetized at E14.5, the uterine horn was exposed, and either empty plasmid or human/mouse FOXP2 cDNA and Gfp cDNA were electroporated into the ventricular wall (*86, 87*). At postnatal day 7, electroporated mice were sacrificed, their brains extracted into cooled aCSF, and then sectioned sagittally at 300μmin cooled aCSF. Sections were then microdissected, pooling for each mouse its neocortical layers 5-6 (infragranular, GFP-negative cells), L4 (granular, GFP-positive cells), and layers 2-3 (infragranular, GFP-negative cells), similar to previous techniques (*88*). RNA was isolated using Qiagen RNeasy Plus Mini Kit, with an on-column DNase treatment. RNA quantity and quality was measured using a Nanodrop Spectrophotometer and TapeStation 2200

(Agilent Technologies). Libraries for sequencing were generated with Illumina TruSeq. Mouse samples used were summarized in table S6.

Human FOXP2 cDNA (NM_014491) was purchased from Origene (Cat. # SC314508) and mouse Foxp2 cDNA (BC062926) was purchased from OpenBioSystems (Item MMM1013-9497440), and both were subcloned into pCAGEN. The pCAGEN-GFP construct has been described previously (*86, 87*). The resulting vectors were sequenced at Yale's Keck DNA Sequencing Lab to confirm codon usage. The pCAGEN-GFP construct has been described previously.

Analysis of mouse in utero electroporation RNA-seq data

Read alignment was performed via STAR (*57*) using the mouse mm10 genome and GENCODE v19 annotation, followed by HTSeq (*58*) for read counting. DESeq2 (*59*) was implemented for differential gene expression between FOXP2- and GFP-electroporated mouse cortices, retaining only genes with fold changes no smaller than 1.5 and false discovery rate no greater than 0.05. The differentially expressed genes were organized in table S7.

Co-expression analysis of brain disorder genes

For the genes genetically associated with neurological and neuropsychiatric diseases, we took advantage of previously curated gene lists for autism spectrum disorder (ASD) , attention deficit hyperactivity disorder (ADHD), developmental delay disorder (DDD), epilepsy (EP), schizophrenia (SCZ), major depressive disorder (MDD), multiple sclerosis (MS), bipolar disorder (BD), Alzheimer's disease (AD), Parkinson's disease (PD), frontotemporal dementia and amyotrophic lateral sclerosis (FTD-ALS) and stroke (*10, 89-93*). Because ASD encompassed only

13 genes, we extracted ASD-related genes from SFARI (https://gene.sfari.org/) under the syndromic category and with available gene scores no bigger than 2 (*94*).

We first evaluated the overall expression conservation for different disease gene lists using the average coefficients of variation (CV). For a given gene, we calculated its CV in a given subtype based on its expression ratios in this subtype across the four species. To reduce noises caused by low expression, for each gene we only included the subtypes that express the gene in more than 20% of the cells in at least one species. The CVs across the retained subtypes were averaged as the mean CV for the given gene. The distributions of mean CVs of multiple genes were visualized in density plots.

We next performed a gene clustering analysis to gain a more detailed view of the expression patterns of these disease risk genes. Prior to clustering, we classified genes as conserved or divergent category based on their expression conservation across species. Genes without any specie-specific expression were considered to be in the conserved category and other genes were assigned to the divergent category. We then identified gene co-expression modules in each category separately based on the following procedures. Specifically, we calculated the Pearson correlation coefficients between these genes based on their average expression across subclasses in all species, followed by defining the gene-gene distances as Pearson correlation coefficients subtracted from 1. We then performed hierarchical clustering using the *hclust* function in R with *ward.D2* algorithm. To define a high resolution of gene clusters, we parcellated the genes to different clusters using the *cutreeDynamic* function from R WGCNA package (*95*), with a minimum of 10 genes per module. Because the expression patterns were more heterogeneous for divergent genes, we set the module detection sensitivity at 3 for divergent genes and 2 for conserved genes. The resulted module eigengenes (table S8) were cut to 50 bins with equal bin

width, and the eigengene values were replaced by the bin indices. For each module, we compared the fold changes of the bin indices of each subclass across species and use a cutoff of 1.4 to determine species enrichment. We then used weighted pie plots to summarize the gene expression patterns across subclasses and species.

We took the advantage of the human transcription factor database (*96*) and intersected them with the gene co-expression modules, which highlighted several genes encoding key transcription factors displaying species-specific expression.

Identifying brain disorder risk genes showing subclass- and species-specific expression

To identify neurological and neuropsychiatric disease risk genes with subclass- and species-specific expression, we adopted a similar strategy described in 'Detection of differentially expressed genes and species-specific genes'. For both within-species and cross-species comparisons, we utilized the *FindMarkers* function to identify significant changes requiring a minimum expression ratio of 0.25, fold changes of expression ratio ≥ 1.5, log fold changes of average expression ≥ 0.35, false discovery rate ≤ 0.01. In order to be classified as species-specific genes, we required genes to have expression levels higher/lower in this species as compared to all three other species.

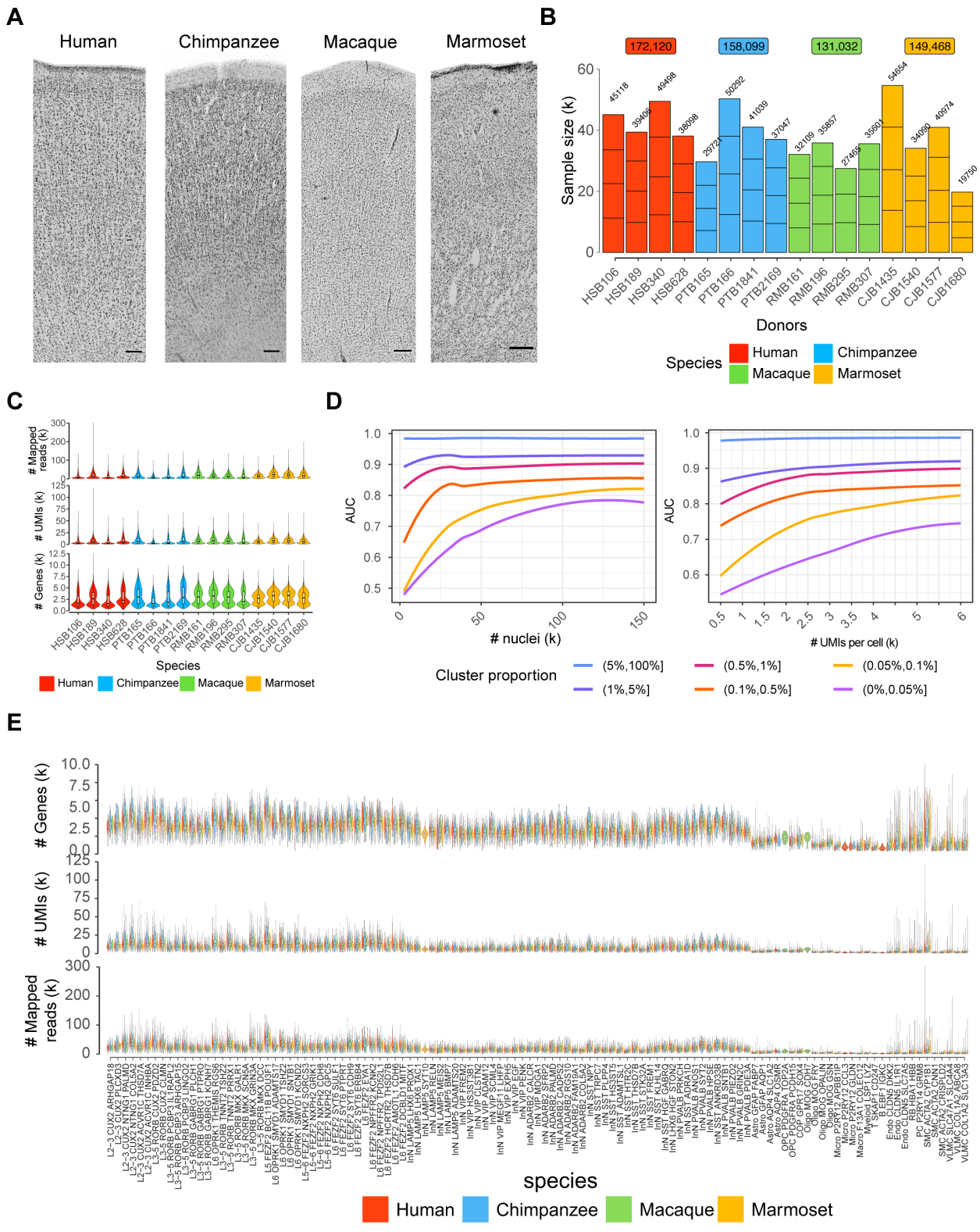

A
Human
Chimpanzee
Macaque
Marmoset
B
172,120
158,099
131,032
149,468
Sample size (k)
Donors
Species
Human
Chimpanzee
Macaque
Marmoset
C
Mapped reads (k)
UMIs (k)
Genes (k)
Species
D
AUC
nuclei (k)
UMIs per cell (k)
Cluster proportion
(5%,100%]
(1%,5%]
(0.5%,1%]
(0.1%,0.5%]
(0.05%,0.1%]
(0%,0.05%]
E
Genes (k)
UMIs (k)
Mapped reads (k)
species
Human
Chimpanzee
Macaque
Marmoset

Fig. S1. Overview of snRNA-seq data quality. (A) Nissl staining showing tissue cytoarchitecture and homology of the dlPFC in the four analyzed species. Scale bar: 200μm. **(B)** Bar plots showing the number of nuclei in each individual, with bars colored by species. Each segment on a bar represents a technical replicate of the given individual. Our data encompassed a balanced number of nuclei sampled in each species, including 172,120 nuclei in humans, 158,099 nuclei in chimpanzees, 131,032 nuclei in macaques and 149,468 nuclei in marmosets. **(C)** Violin plot showing the distribution of the number of uniquely mapped reads, unique molecular identifiers (UMIs), and detected genes across individuals. (**D**) Line plots showing our ability to detect rare subtypes. The human dlPFC data was subsampled to show how subsampling of cells (left) or UMIs (right) could affect the ability to differentiate subtypes (measured by AUC scores; materials and methods). Lines are colored based on subtype abundance. (**E**) Violin plots show the distribution of uniquely mapped reads, UMIs, and detected genes across subtypes and species, with box plots displaying the median, 25^{th} quantiles and 75^{th} quantiles. An average of 2,781 genes were detected per nucleus across all the four species.

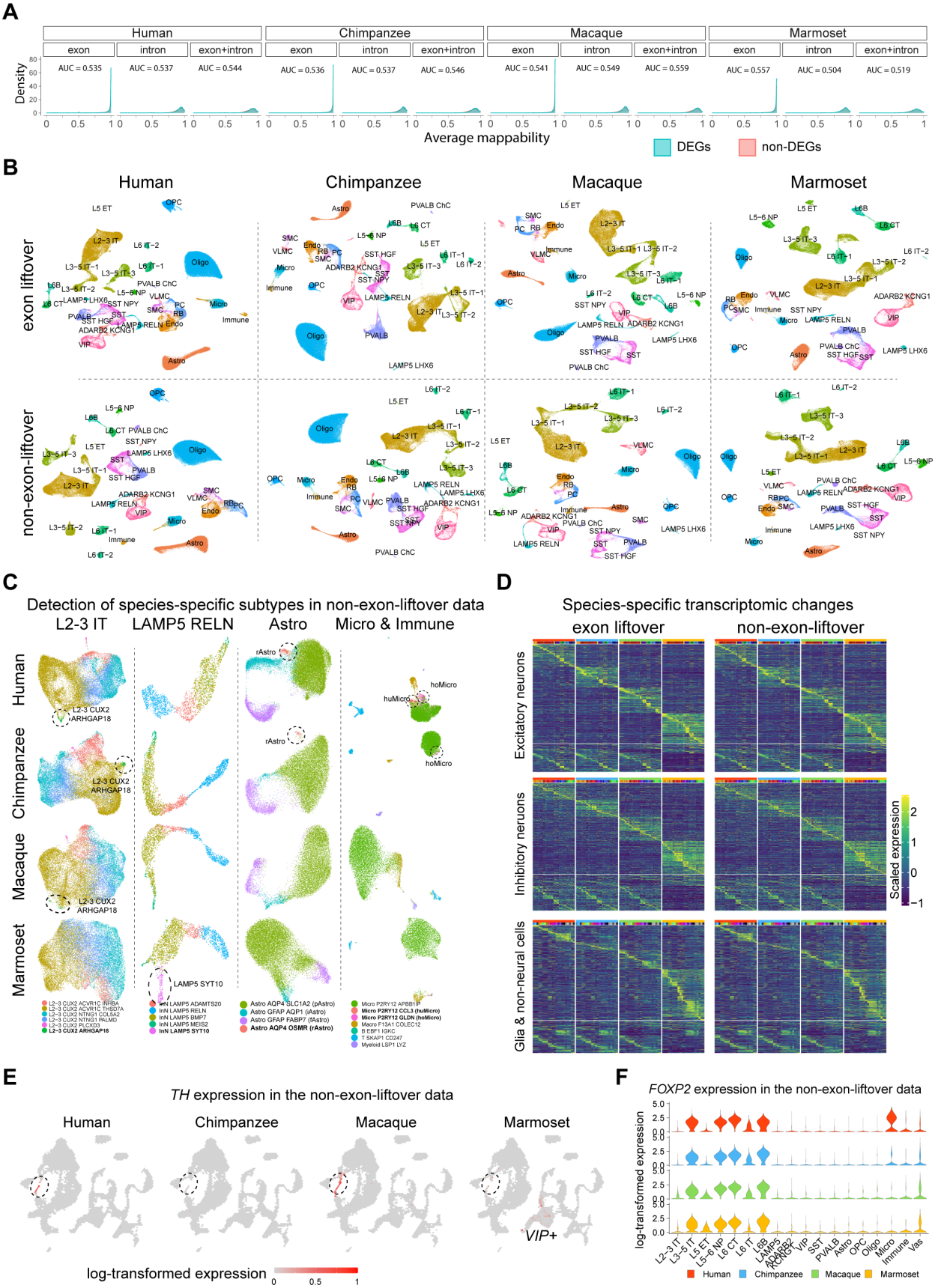

A
Human
Chimpanzee
Macaque
Marmoset
exon
intron
exon+intron
Density
AUC = 0.535
AUC = 0.537
AUC = 0.544
AUC = 0.536
AUC = 0.537
AUC = 0.546
AUC = 0.541
AUC = 0.549
AUC = 0.559
AUC = 0.557
AUC = 0.504
AUC = 0.519
Average mappability
DEGs
non-DEGs
B
Human
Chimpanzee
Macaque
Marmoset
exon liftover
non-exon-liftover
C
Detection of species-specific subtypes in non-exon-liftover data
L2-3 IT
LAMP5 RELN
Astro
Micro & Immune
Human
Chimpanzee
Macaque
Marmoset
L2-3 CUX2 ARHGAP18
LAMP5 SYT10
rAstro
hoMicro
huMicro
D
Species-specific transcriptomic changes
exon liftover
non-exon-liftover
Excitatory neurons
Inhibitory neurons
Glia & non-neural cells
Scaled expression
E
TH expression in the non-exon-liftover data
Human
Chimpanzee
Macaque
Marmoset
VIP+
log-transformed expression
F
FOXP2 expression in the non-exon-liftover data
log-transformed expression
Human
Chimpanzee
Macaque
Marmoset

Fig. S2. Robustness of major findings to different gene annotation models. (A) Species differentially expressed genes (DEGs) shows a similar distribution of mappability scores compared to background genes (non-DEGs). AUC, area under the ROC curve. **(B)** Subclass structure generated using the consensus gene model (exon-liftover) is consistent with that using the publicly available gene annotation model (non-exon-liftover). **(C)** Species-specific subtypes can be robustly recovered using the non-exon-liftover model. **(D)** Species-specific transcriptomic changes detected in the exon-liftover model are still robust to the non-exon-liftover model. (**E-F**) *TH* **(E)** and *FOXP2* **(F)** expression patterns remain unchanged in the non-exon-liftover model.

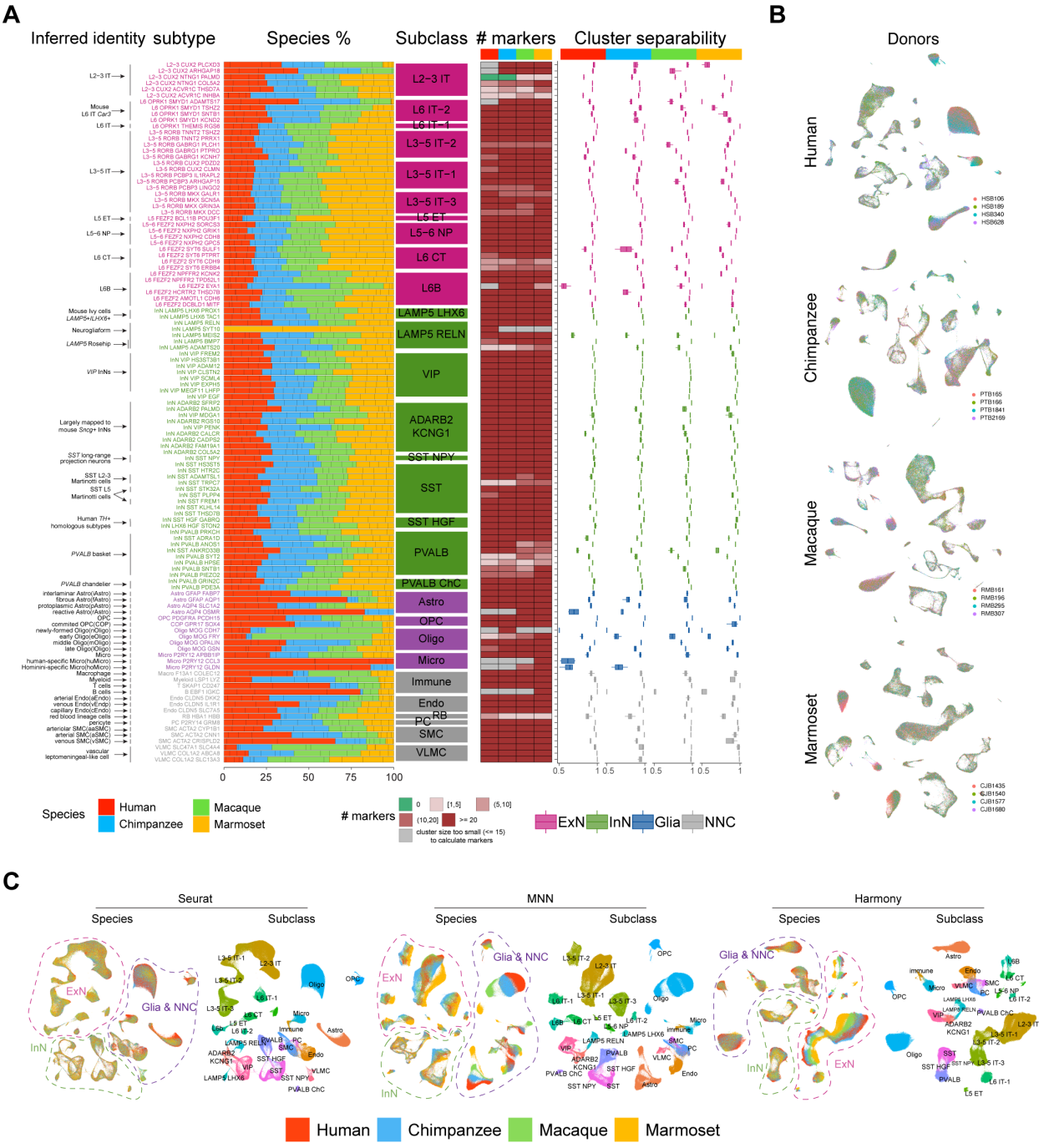


Fig. S3. Robustness of subtype clustering and integration. (A) From left to right: predicted cell identity of subtypes; subtype names; bar plot showing the relative donor proportion (bars are colored by species and segmented by donors); cell subclass names; heatmap showing the number of markers detected for each subtype; AUC scores measuring cluster separability in the down-sampled data (see Methods for details). (**B**) UMAP layout visualizing donor information suggests

there is little donor bias driving the cluster separation. **(C)** Integration of cells from the four species using Seurat (left) (*49*), MNN (middle)(*63*) and Harmony (right) (*97*). Cells were colored by species or subclass. ExN, excitatory neurons; InN, inhibitory neurons; NNC, non-neural cells; IT, intratelencephalic excitatory neurons; ET, extratelencephalic excitatory neurons; CT, corticothalamic projection excitatory neurons; NP, near-projecting excitatory neurons; *PVALB* ChC: *PVALB* chandelier cells; Astro, astrocytes; Oligo, oligodendrocytes; OPC: oligodendrocyte precursor cells; Micro: microglia; Macro, macrophages; T, T cells; B, B cells; Endo: endothelial cells; RB, red blood lineage cells; PC, pericytes; SMC, smooth muscle cells; VLMC, vascular leptomeningeal cells.

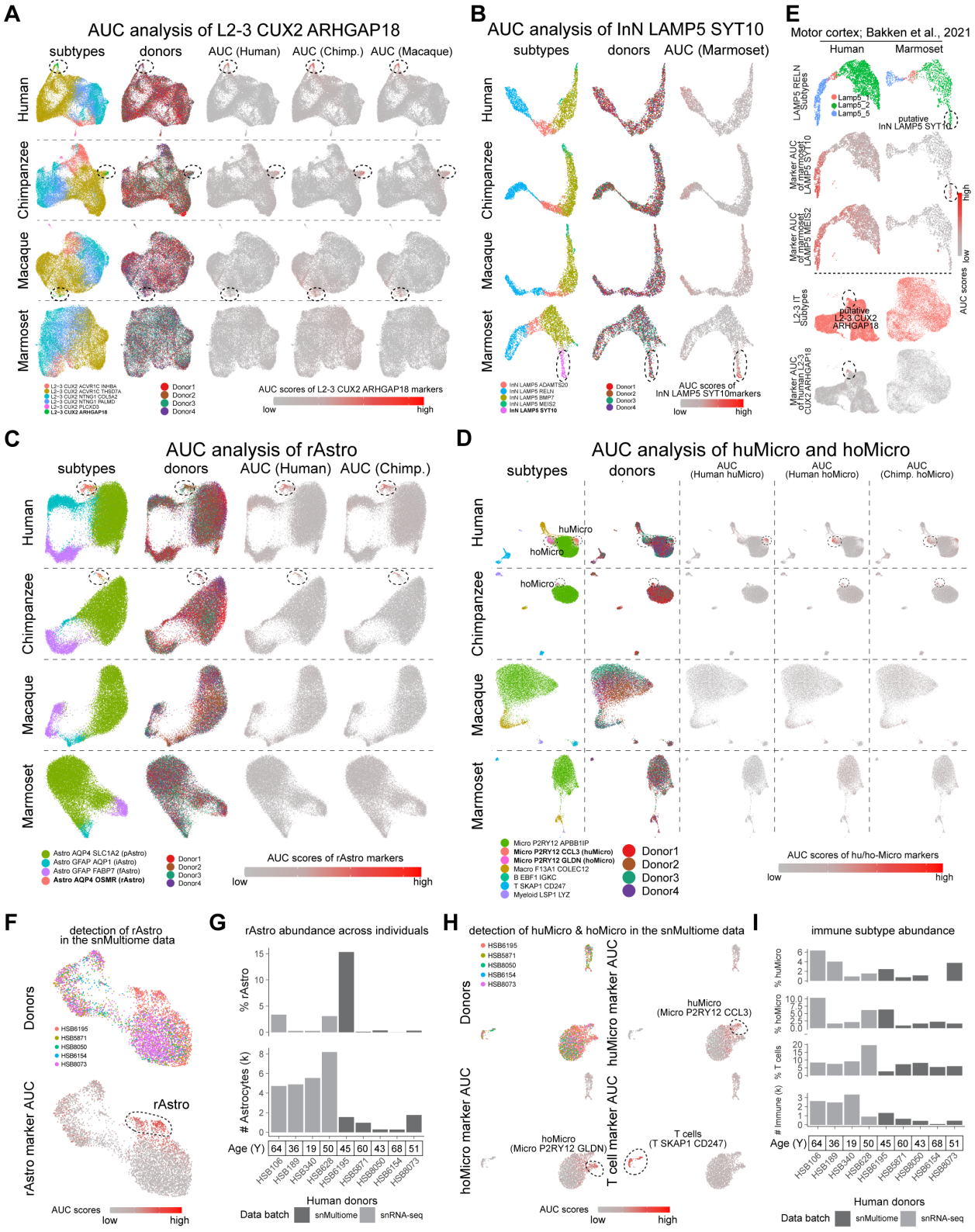


Fig. S4. Evaluating species-specificity of the detected species-specific subtypes. (A-D) AUC scores illustrating the enrichment of the markers for each species-specific subtype: L2-3 *CUX2*

ARHGAP18 (A), InN *LAMP5 SYT10* (B), Astro *AQP4 OSMR* (rAstro, C), Micro *P2RY12 CCL3* (huMicro) and Micro *P2RY12 GLDN* (hoMicro) (D). Species-specific subtypes were highlighted by dashed circles. (**E**) Similar cells marked by the markers of InN *LAMP5 SYT10* (top three rows) and L2-3 *CUX2 ARHGAP18* (bottom two rows; highlighted by dashed lines) were also detected in motor cortex (*22*) and the species-specificity still stands. (**F**) The Astro *AQP4 OSMR* subtype is also detected in the sn-multiome data (see fig. S21 for more details), as demonstrated by the AUC scores measuring cell type marker enrichment. (**G**) Bar plots showing the abundance of the Astro *AQP4 OSMR* subtype across the human donors. In the snRNA-seq data batch, where we have more power to measure subtype proportions, rAstro abundance tend to increase with age, which is consistent with previous reports (*98*). (**H**) Both huMicro and hoMicro are detected in the sn-multiome data. (**I**) The abundance of huMicro and hoMicro are variable across the analyzed human donors, while the T cells abundance are relatively more consistent between donors.

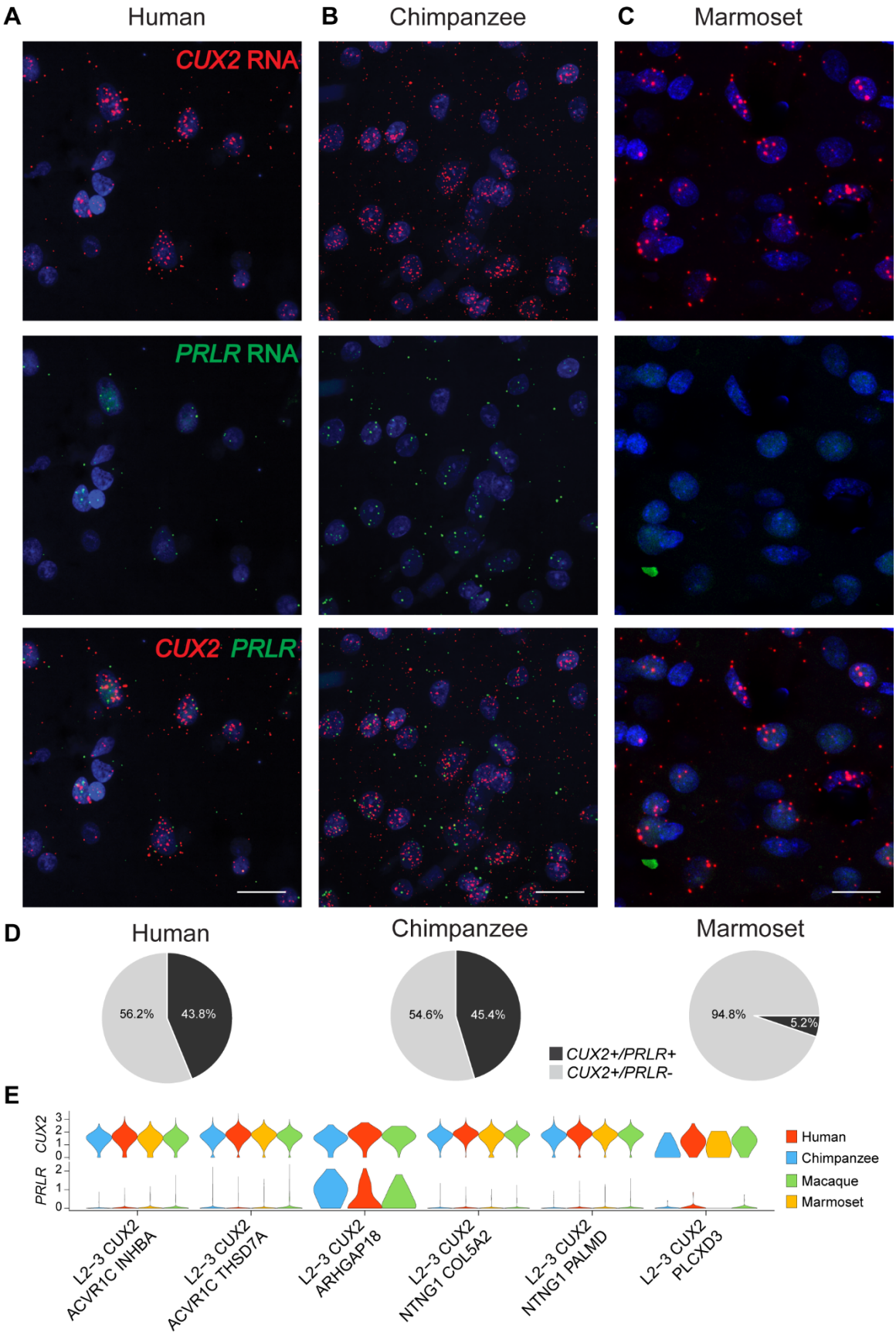

A
Human
B
Chimpanzee
C
Marmoset
CUX2 RNA
PRLR RNA
CUX2 PRLR
D
Human
Chimpanzee
Marmoset
56.2%
43.8%
54.6%
45.4%
94.8%
5.2%
CUX2+/PRLR+
CUX2+/PRLR-
E
CUX2
PRLR
Human
Chimpanzee
Macaque
Marmoset
L2–3 CUX2 ACVR1C INHBA
L2–3 CUX2 ACVR1C THSD7A
L2–3 CUX2 ARHGAP18
L2–3 CUX2 NTNG1 COL5A2
L2–3 CUX2 NTNG1 PALMD
L2–3 CUX2 PLCXD3

Fig. S5. RNA *in situ* hybridization validating the species-specificity of L2-3 *CUX2 ARHGAP18* subtype. (A) RNA *in situ* hybridization probing two representative markers (*CUX2* and *PRLR*) of the Catarrhini-specific subtype L2-3 *CUX2 ARHGAP18* in humans (A), chimpanzees (B) and marmosets. (C). Scale bar: 20μm in each species. (**D**) Quantification of cells expressing combinatory marker genes. n = 3. (**E**) Expression patterns of *CUX2* and *PRLR* across the L2-3 intratelencephalic subtypes in the snRNA-seq data.

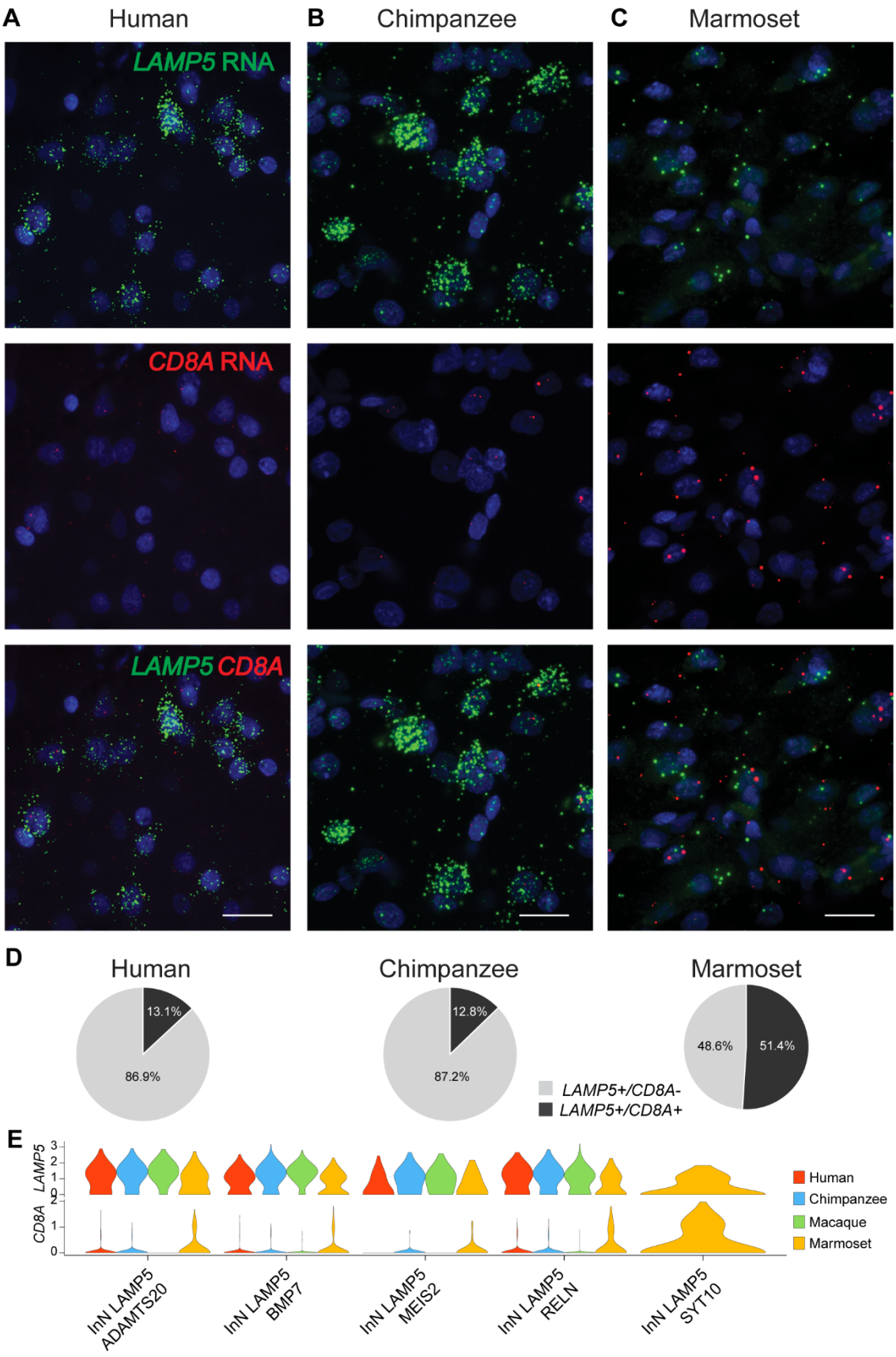

A
Human
B
Chimpanzee
C
Marmoset
LAMP5 RNA
CD8A RNA
LAMP5 CD8A
D
Human
Chimpanzee
Marmoset
13.1%
86.9%
12.8%
87.2%
48.6%
51.4%
LAMP5+/CD8A-
LAMP5+/CD8A+
E
LAMP5
CD8A
Human
Chimpanzee
Macaque
Marmoset
InN LAMP5 ADAMTS20
InN LAMP5 BMP7
InN LAMP5 MEIS2
InN LAMP5 RELN
InN LAMP5 SYT10

Fig. S6. RNA *in situ* hybridization validating the species-specificity of InN *LAMP5 SYT10* subtype. (A) RNA *in situ* hybridization probing two representative markers (*LAMP5* and *CD8A*) of the marmoset-specific subtype InN *LAMP5 SYT10* in humans (A), chimpanzees (B) and marmosets (C). Scale bar: 20μm in each species. (**D**) Quantification of cells expressing combinatory marker genes. n = 3. (**E**) Expression patterns of *LAMP5* and *CD8A* across the LAMP5 RELN subtypes in the snRNA-seq data.

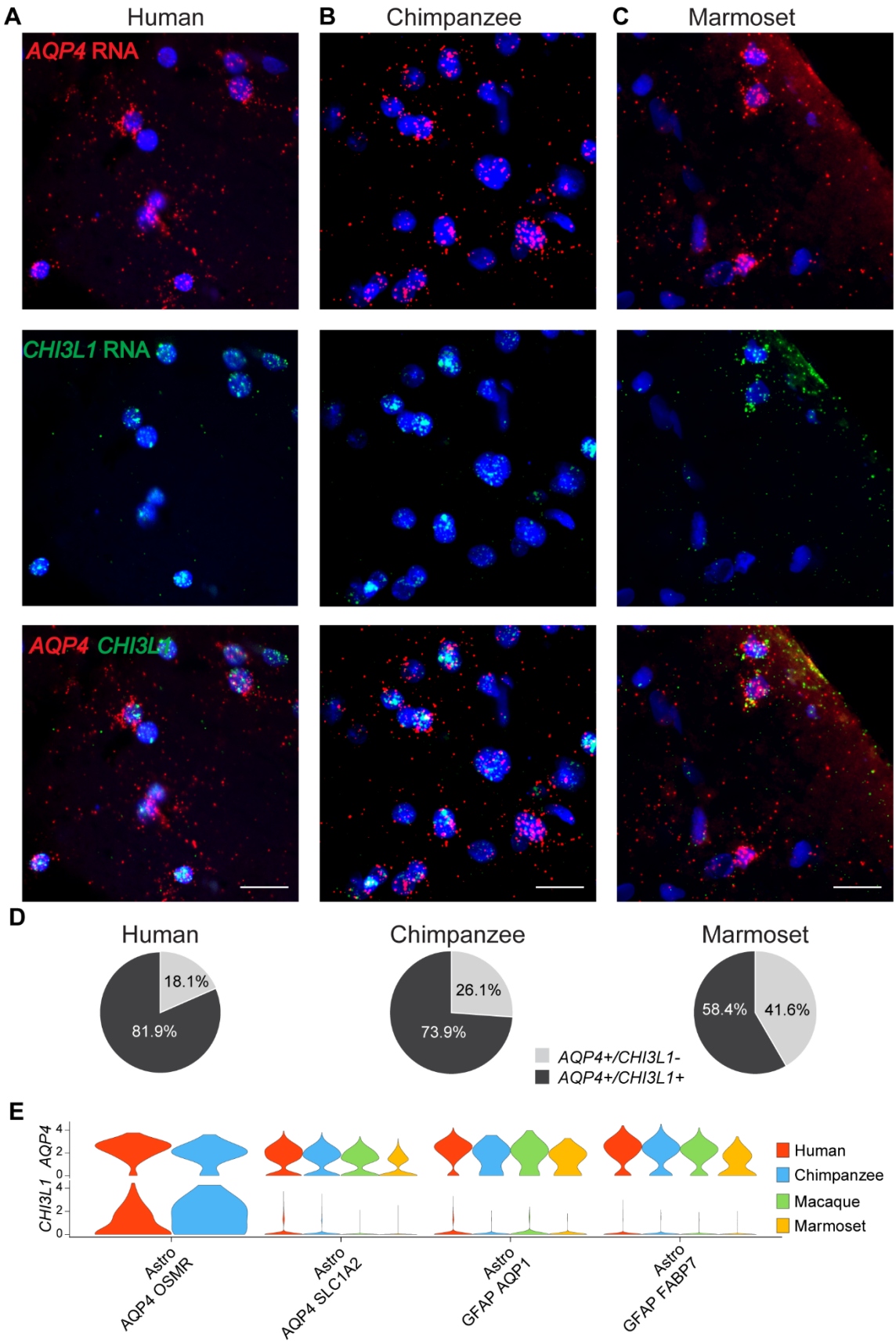

A
Human
B
Chimpanzee
C
Marmoset
AQP4 RNA
CHI3L1 RNA
AQP4 CHI3L1
D
Human
Chimpanzee
Marmoset
18.1%
81.9%
26.1%
73.9%
58.4%
41.6%
AQP4+/CHI3L1-
AQP4+/CHI3L1+
E
AQP4
CHI3L1
Human
Chimpanzee
Macaque
Marmoset
Astro AQP4 OSMR
Astro AQP4 SLC1A2
Astro GFAP AQP1
Astro GFAP FABP7

Fig. S7. RNA *in situ* hybridization validating the species-specificity of Astro *AQP4 OSMR* subtype. (A) RNA *in situ* probing two representative markers (*AQP4* and *CHI3L1*) of the Hominini-specific subtype Astro *AQP4 OSMR* in humans (A), chimpanzees (B) and marmosets (C). Scale bar: 20μm in each species. (**D**) Quantification of cells expressing combinatory marker genes. n = 3. (**E**) Expression patterns of *AQP4* and *CHI3L1* across the astrocyte subtypes in the snRNA-seq data.

A

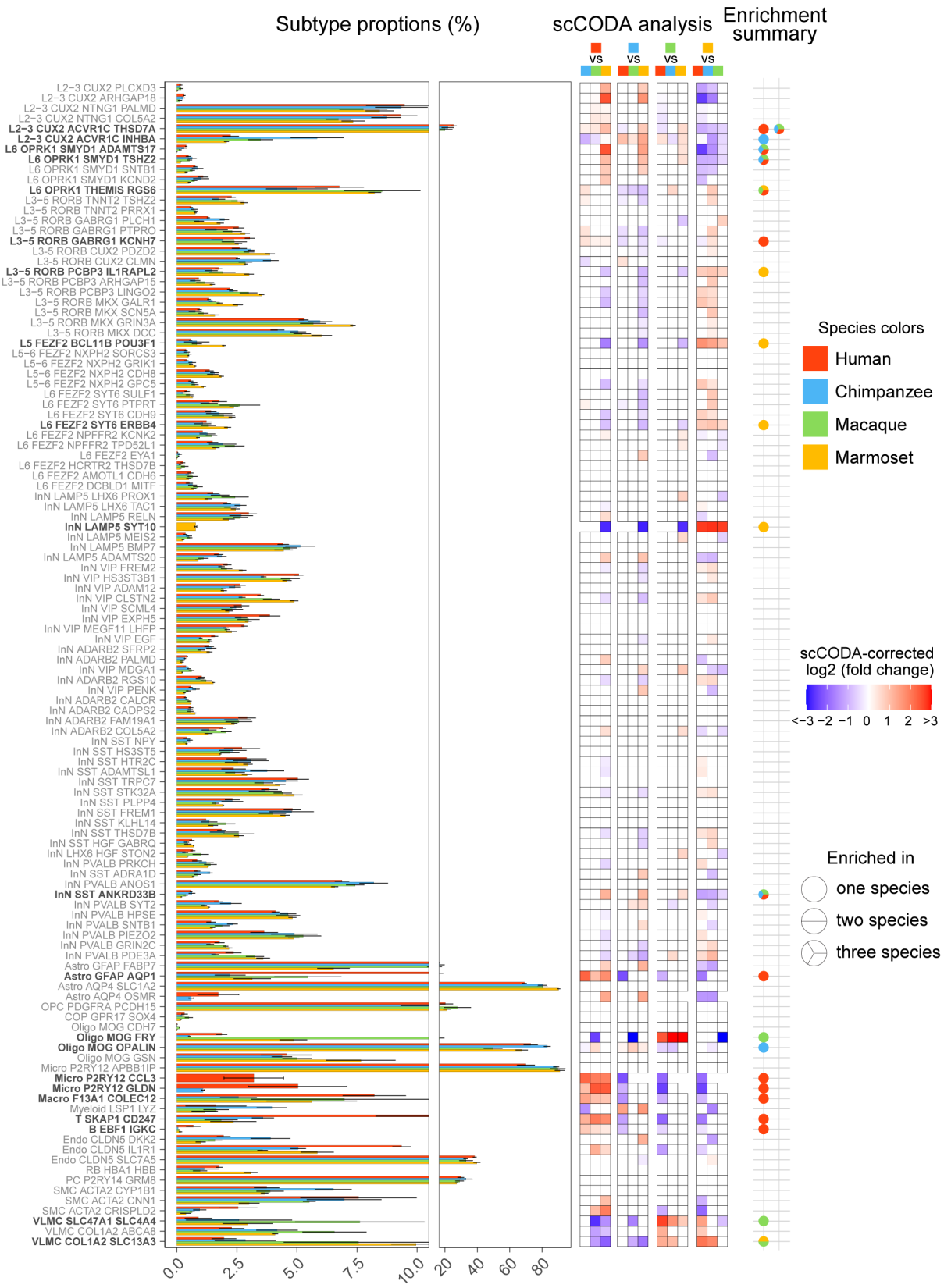

Subtype proptions (%)
scCODA analysis
Enrichment summary
vs
vs
vs
vs
L2−3 CUX2 PLCXD3
L2−3 CUX2 ARHGAP18
L2−3 CUX2 NTNG1 PALMD
L2−3 CUX2 NTNG1 COL5A2
L2−3 CUX2 ACVR1C THSD7A
L2−3 CUX2 ACVR1C INHBA
L6 OPRK1 SMYD1 ADAMTS17
L6 OPRK1 SMYD1 TSHZ2
L6 OPRK1 SMYD1 SNTB1
L6 OPRK1 SMYD1 KCND2
L6 OPRK1 THEMIS RGS6
L3−5 RORB TNNT2 TSHZ2
L3−5 RORB TNNT2 PRRX1
L3−5 RORB GABRG1 PLCH1
L3−5 RORB GABRG1 PTPRO
L3−5 RORB GABRG1 KCNH7
L3-5 RORB CUX2 PDZD2
L3-5 RORB CUX2 CLMN
L3−5 RORB PCBP3 IL1RAPL2
L3−5 RORB PCBP3 ARHGAP15
L3−5 RORB PCBP3 LINGO2
L3−5 RORB MKX GALR1
L3−5 RORB MKX SCN5A
L3−5 RORB MKX GRIN3A
L3−5 RORB MKX DCC
L5 FEZF2 BCL11B POU3F1
L5−6 FEZF2 NXPH2 SORCS3
L5−6 FEZF2 NXPH2 GRIK1
L5−6 FEZF2 NXPH2 CDH8
L5−6 FEZF2 NXPH2 GPC5
L6 FEZF2 SYT6 SULF1
L6 FEZF2 SYT6 PTPRT
L6 FEZF2 SYT6 CDH9
L6 FEZF2 SYT6 ERBB4
L6 FEZF2 NPFFR2 KCNK2
L6 FEZF2 NPFFR2 TPD52L1
L6 FEZF2 EYA1
L6 FEZF2 HCRTR2 THSD7B
L6 FEZF2 AMOTL1 CDH6
L6 FEZF2 DCBLD1 MITF
InN LAMP5 LHX6 PROX1
InN LAMP5 LHX6 TAC1
InN LAMP5 RELN
InN LAMP5 SYT10
InN LAMP5 MEIS2
InN LAMP5 BMP7
InN LAMP5 ADAMTS20
InN VIP FREM2
InN VIP HS3ST3B1
InN VIP ADAM12
InN VIP CLSTN2
InN VIP SCML4
InN VIP EXPH5
InN VIP MEGF11 LHFP
InN VIP EGF
InN ADARB2 SFRP2
InN ADARB2 PALMD
InN VIP MDGA1
InN ADARB2 RGS10
InN VIP PENK
InN ADARB2 CALCR
InN ADARB2 CADPS2
InN ADARB2 FAM19A1
InN ADARB2 COL5A2
InN SST NPY
InN SST HS3ST5
InN SST HTR2C
InN SST ADAMTSL1
InN SST TRPC7
InN SST STK32A
InN SST PLPP4
InN SST FREM1
InN SST KLHL14
InN SST THSD7B
InN SST HGF GABRQ
InN LHX6 HGF STON2
InN PVALB PRKCH
InN SST ADRA1D
InN PVALB ANOS1
InN SST ANKRD33B
InN PVALB SYT2
InN PVALB HPSE
InN PVALB SNTB1
InN PVALB PIEZO2
InN PVALB GRIN2C
InN PVALB PDE3A
Astro GFAP FABP7
Astro GFAP AQP1
Astro AQP4 SLC1A2
Astro AQP4 OSMR
OPC PDGFRA PCDH15
COP GPR17 SOX4
Oligo MOG CDH7
Oligo MOG FRY
Oligo MOG OPALIN
Oligo MOG GSN
Micro P2RY12 APBB1IP
Micro P2RY12 CCL3
Micro P2RY12 GLDN
Macro F13A1 COLEC12
Myeloid LSP1 LYZ
T SKAP1 CD247
B EBF1 IGKC
Endo CLDN5 DKK2
Endo CLDN5 IL1R1
Endo CLDN5 SLC7A5
RB HBA1 HBB
PC P2RY14 GRM8
SMC ACTA2 CYP1B1
SMC ACTA2 CNN1
SMC ACTA2 CRISPLD2
VLMC SLC47A1 SLC4A4
VLMC COL1A2 ABCA8
VLMC COL1A2 SLC13A3
0.0
2.5
5.0
7.5
10.0
20
40
60
80
Species colors
Human
Chimpanzee
Macaque
Marmoset
scCODA-corrected log2 (fold change)
<−3 −2 −1 0 1 2 >3
Enriched in
one species
two species
three species

Fig. S8. Changes in subtype abundance across the four analyzed primates. (A) Left: bar plots illustrating the subtype proportions in each major cell group. Here to accommodate dissection bias, the proportions were calculated among the following six cell groups: excitatory neurons, inhibitory neurons, astrocytes, oligodendrocyte lineage cells, immune cells, and other non-neural cells. Middle: heatmap showing the scCODA-corrected fold changes of subtype abundance between each pair of species. Non-zero values represent significant changes. Right: pie charts summarizing the patterns of species enrichment of subtype abundance. Note that the subtype "L2-3 *CUX2 ACVR1C THSD7A*" is enriched in human compared to other non-human primates and also enriched in Catarrhini compared to marmosets, therefore being labeled by two pie charts.

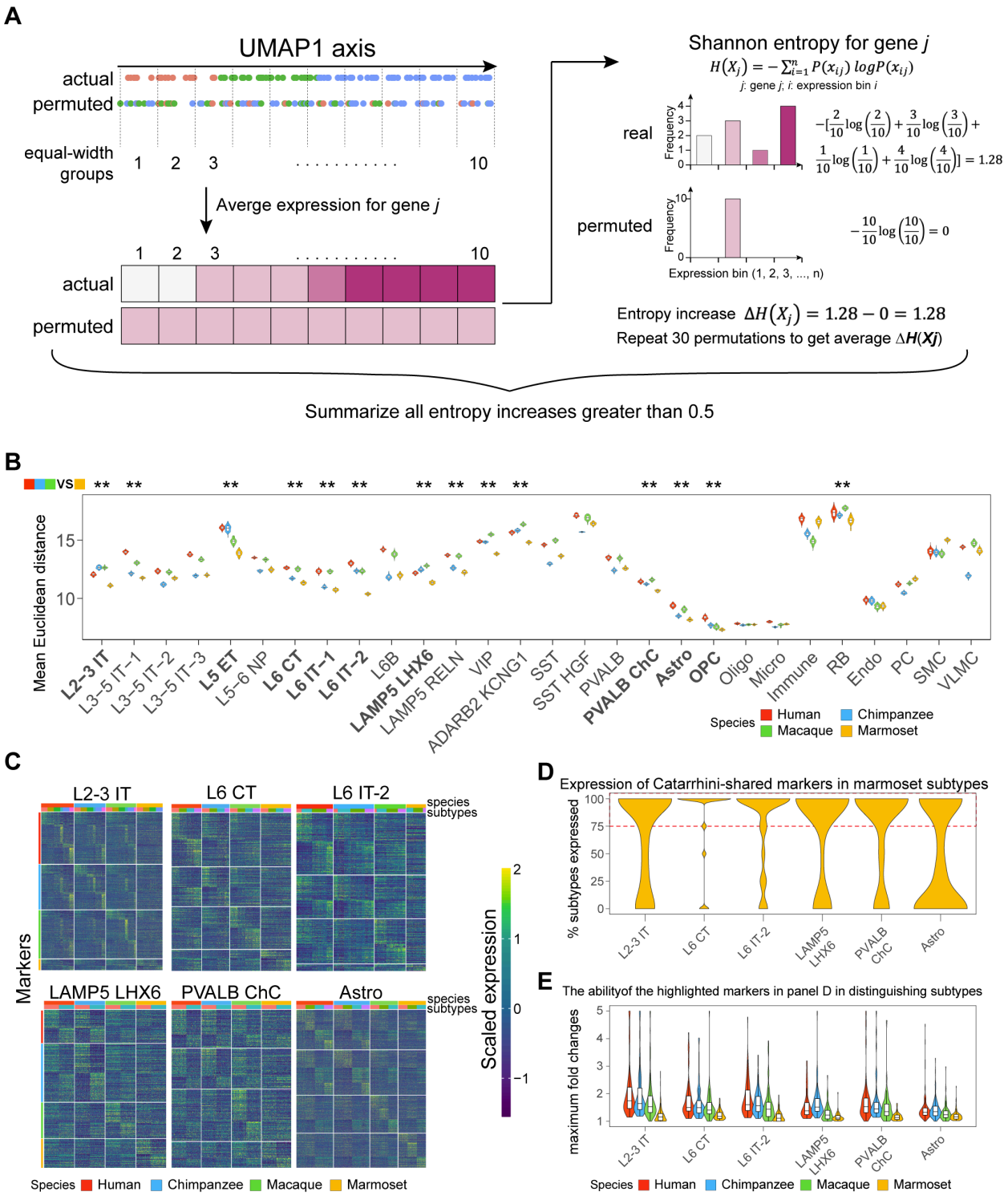


Fig. S9. Species differences in inter-cell transcriptomic heterogeneity in the homologous subclasses. (A) Diagram illustrating how the subclass heterogeneity scores were calculated (see materials and methods for details). **(B)** We first subsampled each subclass to have the same number

of cells across species. Within each down-sampled dataset, we generated bootstrap replicates of the highly variable genes and performed the principal component analysis on the bootstrap replicates, which can avoid the influence of certain unstructured genes with high variance. We calculated the mean Euclidean distances across the bootstrap replicates for each given down-sampled dataset and visualized the distribution of the mean Euclidean distances across 100 down-sampling replicates. Significance was tested via Wilcoxon Rank Sum test followed by Bonferroni correction (**: adjusted $p < 0.01$). Subclasses tested with significant entropy increases in Catarrhini (Fig. 2C) are highlighted in bold text. (**C**) Expression of subtype marker genes in the subclasses showing Catarrhini-specific entropy increases (Fig. 2C). Certain subclasses without any subtypes defined were not shown here. Overall, marmoset shows lower number of markers distinguishing the subtypes. (**D-E**) We extracted Catarrhini-conserved (≥ 2 species) subtype markers in each subclass and evaluated their expression patterns in the marmoset subtypes hierarchically related to the same subclass. Many markers distinguishing Catarrhini subtypes are also broadly expressed in marmoset, but show much lower variation compared to Catarrhini, as shown by their maximum fold changes between subtypes (E).

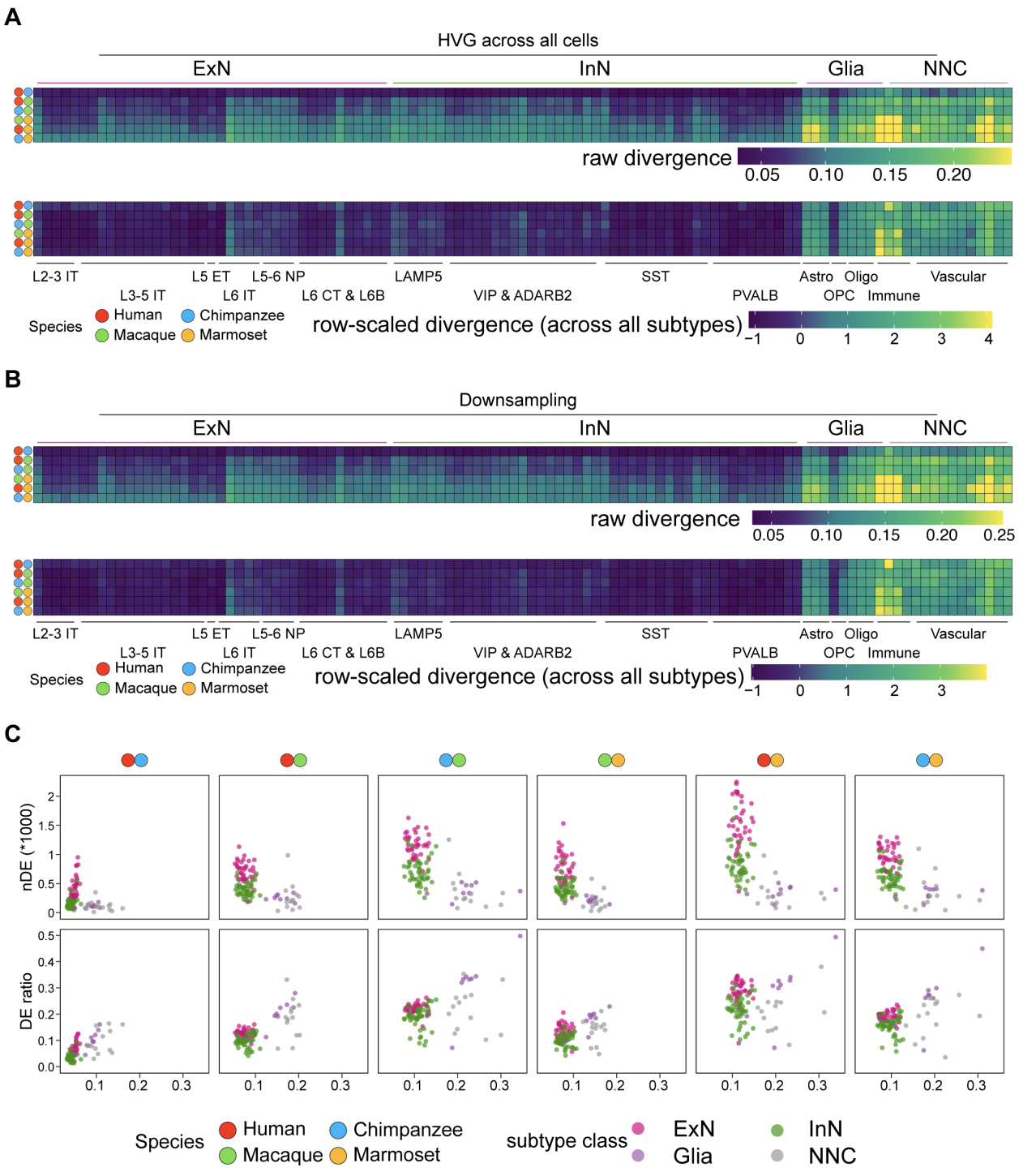

A
HVG across all cells
ExN
InN
Glia
NNC
raw divergence
0.05 0.10 0.15 0.20
L2-3 IT
L3-5 IT
L5 ET
L6 IT
L5-6 NP
L6 CT & L6B
LAMP5
VIP & ADARB2
SST
PVALB
Astro
OPC
Oligo
Immune
Vascular
Species
Human
Chimpanzee
Macaque
Marmoset
row-scaled divergence (across all subtypes)
−1 0 1 2 3 4
B
Downsampling
ExN
InN
Glia
NNC
raw divergence
0.05 0.10 0.15 0.20 0.25
L2-3 IT
L3-5 IT
L5 ET
L6 IT
L5-6 NP
L6 CT & L6B
LAMP5
VIP & ADARB2
SST
PVALB
Astro
OPC
Oligo
Immune
Vascular
Species
Human
Chimpanzee
Macaque
Marmoset
row-scaled divergence (across all subtypes)
−1 0 1 2 3
C
nDE (*1000)
DE ratio
Species
Human
Chimpanzee
Macaque
Marmoset
subtype class
ExN
InN
Glia
NNC

Fig. S10. Robustness of transcriptomic divergence among homologous subtypes to different strategies. (**A-B**) Transcriptomic divergence across species pairs and subtypes using two different strategies: using highly variable genes (HVGs) summarized from all cells (A) and down-sampling the cell number in each subtype (B). **(C)** Correlation between transcriptomic divergence (x-axis), and the number of differentially expressed genes (nDE, top y-axis) or the proportion of differentially expressed genes (DE ratio, bottom y-axis) among the expressed genes (expression ratio $\geq$ 0.3). Each dot represents the comparison of a subtype in a given species pair. ExN, excitatory neurons; InN, inhibitory neurons; NNC, non-neural cells.

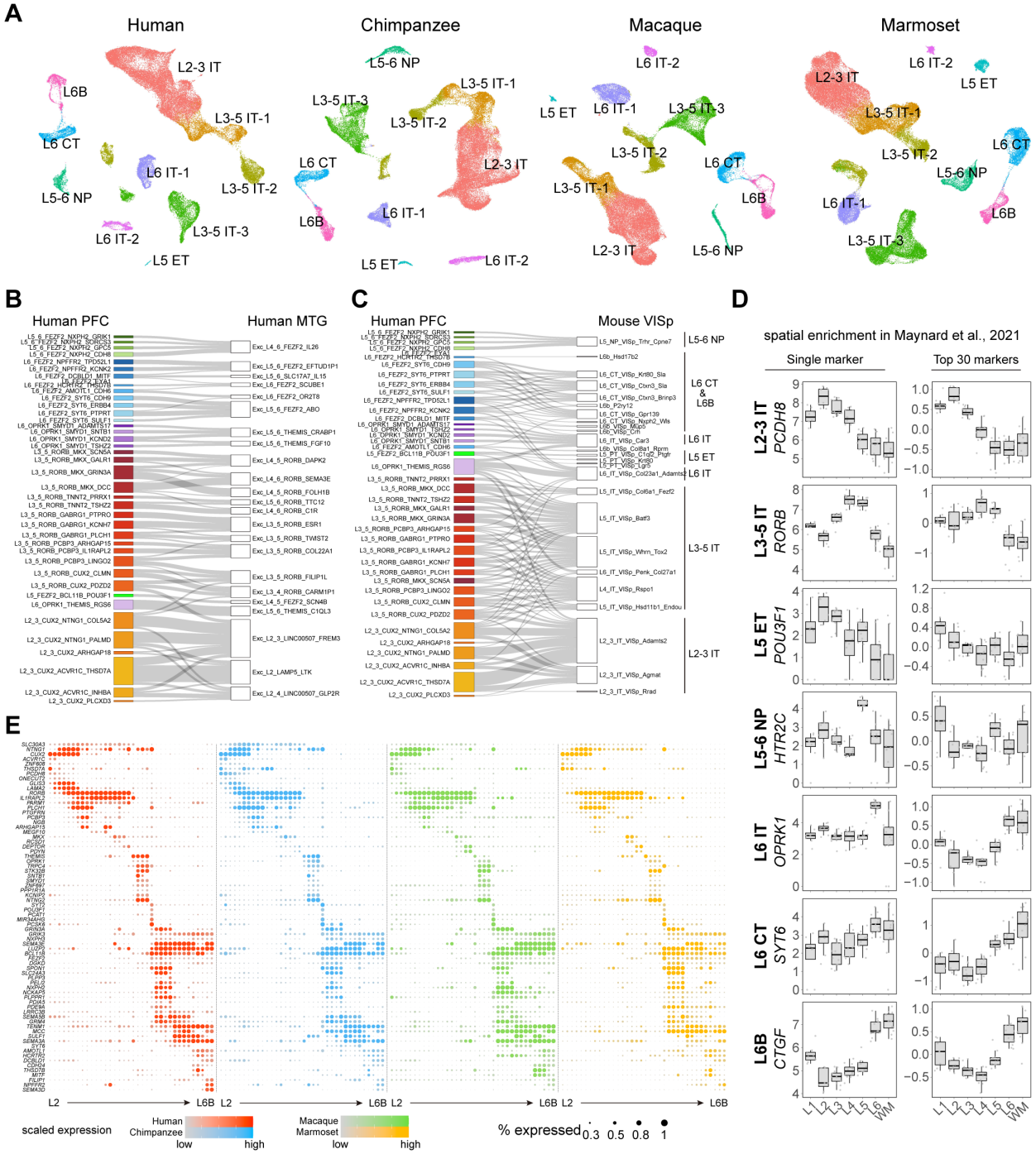

A
Human
Chimpanzee
Macaque
Marmoset
L2-3 IT
L6B
L3-5 IT-1
L6 CT
L6 IT-1
L3-5 IT-2
L5-6 NP
L3-5 IT-3
L6 IT-2
L5 ET
L5-6 NP
L3-5 IT-3
L3-5 IT-1
L3-5 IT-2
L6 CT
L2-3 IT
L6 IT-1
L6B
L5 ET
L6 IT-2
L5 ET
L6 IT-1
L3-5 IT-3
L3-5 IT-2
L6 CT
L3-5 IT-1
L6B
L2-3 IT
L5-6 NP
L6 IT-2
L2-3 IT
L5 ET
L3-5 IT-1
L6 CT
L3-5 IT-2
L5-6 NP
L6 IT-1
L6B
L3-5 IT-3
B
Human PFC
Human MTG
C
Human PFC
Mouse VISp
L5-6 NP
L6 CT & L6B
L6 IT
L5 ET
L6 IT
L3-5 IT
L2-3 IT
D
spatial enrichment in Maynard et al., 2021
Single marker
Top 30 markers
L2-3 IT PCDH8
L3-5 IT RORB
L5 ET POU3F1
L5-6 NP HTR2C
L6 IT OPRK1
L6 CT SYT6
L6B CTGF
L1 L2 L3 L4 L5 L6 WM
E
L2
L6B
scaled expression
Human
Chimpanzee
low
high
Macaque
Marmoset
low
high
% expressed
0.3 0.5 0.8 1

Fig. S11. Transcriptomic organization and conservation of glutamatergic excitatory neurons across the four primates. (A) Separate clustering of glutamatergic excitatory neuron subtypes across species. (**B**) Sankey diagram showing the excitatory neuron taxonomy correspondence between this data (left) and the human medial temporal gyrus (MTG, right) (*16*) based on transcriptomic integration. In each dataset, cells are organized according to their types. (**C**) As in (B), but between our data (left) and the mouse primary visual cortex (VISp, right) (*17*). (**D**) Left: expression of subclass markers across cortical layers in the human PFC spatial transcriptomic data (*18*). Right: average scaled expression of the top 30 subclass markers across cortical layers in the same human PFC spatial transcriptomic data (*18*). Regarding the representative L5 ET marker, we selected *POU3F1* as it has higher specificity than other genes including *FEZF2* and *BCL11B*. Some L5 ET markers are also expressed in other cortical layers (e.g., *POU3F1* exhibits low expression in L2-3 IT excitatory neurons). The L5 ET marker expression is enriched in both L2-3 and L5 in the spatial transcriptome data, likely attributed to the expression of L5 ET markers in L2-3IT excitatory neurons and the low abundance of ET excitatory neurons within L5 cells. (**E**) Dot plots showing the expression of marker genes across subtypes and species. Dots are colored by species, and their color gradients and sizes represent the scaled average expression and percentage of cells expressing a given gene, respectively.

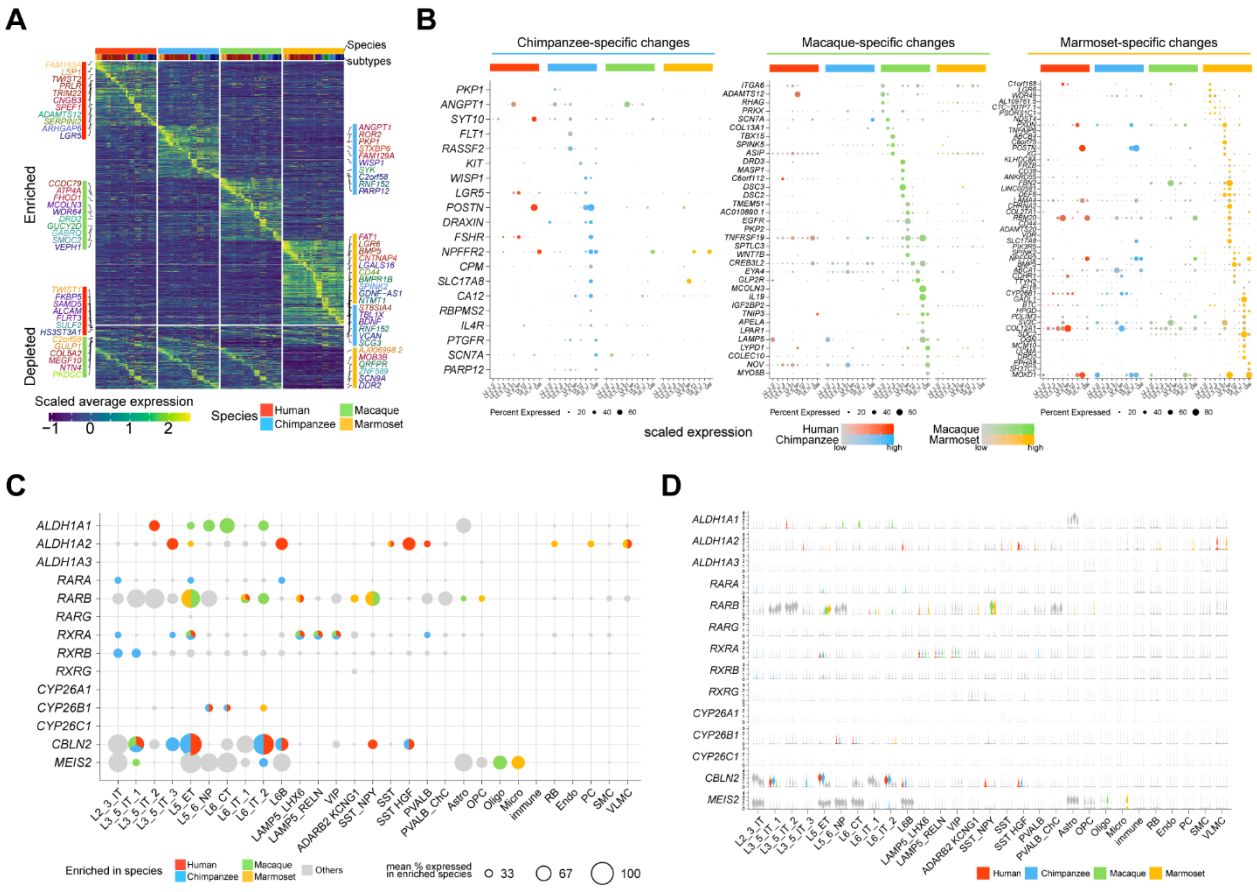


Fig. S12. Subtype- and species-specific transcriptomic differences in glutamatergic excitatory neuron subclasses. (A) Heat map showing standardized average expression across all the subtypes of the four species for genes enriched or depleted in specific subtypes of a given species. Species and subtype information are colored on the top, with subtype colors consistent with Figure 3A. Genes were colored according to the subtypes where they exhibit enrichment or depletion. **(B)** Dot plots displaying expression patterns for genes highly specific in each excitatory neuron subclass of chimpanzees (left), macaques (middle) and marmosets (right). **(C)** Weighted pie plots summarizing the subclass and species differences in the expression of key genes involved in retinoic acid signaling pathway. The colors denote the species in which the gene expression is enriched, with grey color indicating comparable expression across species. The size of the pies represents the average expression ratios in the enriched species. If the pies are colored grey, then

all species were included to compute the average expression ratios. (**D**) Violin plots showing the detailed expression patterns of retinoic acid signaling pathway genes. In each subclass, the four violin areas (from left to right) represent human (red), chimpanzee (blue), macaque (green) and marmoset (orange), respectively. If a gene is enriched in a subset of species, the corresponding violin areas will be colored according to the species where the gene is enriched in.

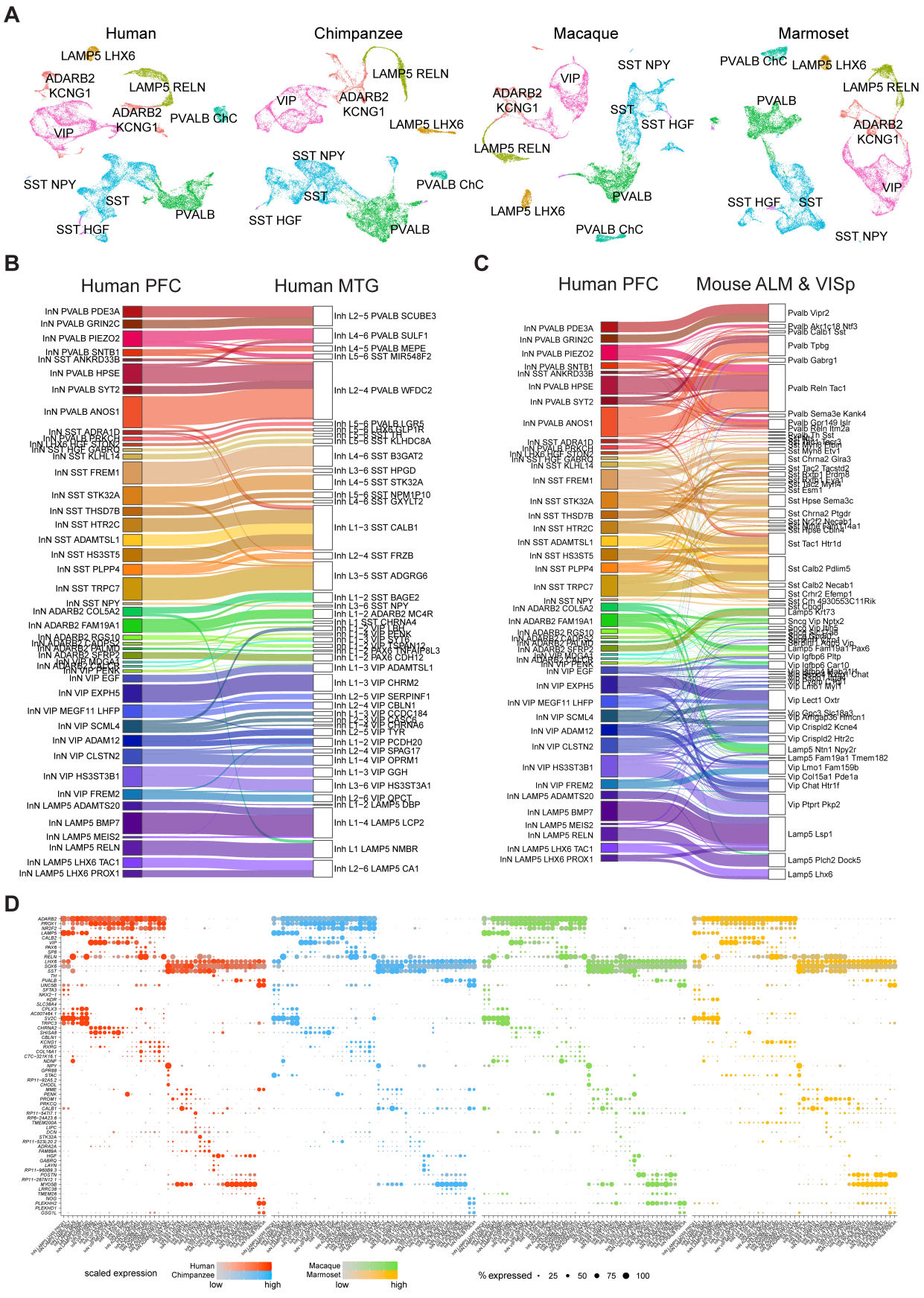

A
Human
LAMP5 LHX6
ADARB2
KCNG1
LAMP5 RELN
ADARB2
KCNG1
PVALB ChC
VIP
SST NPY
SST
PVALB
SST HGF
Chimpanzee
LAMP5 RELN
VIP
ADARB2
KCNG1
LAMP5 LHX6
SST NPY
PVALB ChC
SST
SST HGF
PVALB
Macaque
SST NPY
VIP
ADARB2
KCNG1
SST
SST HGF
LAMP5 RELN
PVALB
LAMP5 LHX6
PVALB ChC
Marmoset
PVALB ChC
LAMP5 LHX6
LAMP5 RELN
PVALB
ADARB2
KCNG1
VIP
SST HGF
SST
SST NPY
B
Human PFC
Human MTG
InN PVALB PDE3A
InN PVALB GRIN2C
InN PVALB PIEZO2
InN PVALB SNTB1
InN SST ANKRD33B
InN PVALB HPSE
InN PVALB SYT2
InN PVALB ANOS1
InN SST ADRA1D
InN PVALB PRKCH
InN LHX6 HGF STON2
InN SST HGF GABRQ
InN SST KLHL14
InN SST FREM1
InN SST STK32A
InN SST THSD7B
InN SST HTR2C
InN SST ADAMTSL1
InN SST HS3ST5
InN SST PLPP4
InN SST TRPC7
InN SST NPY
InN ADARB2 COL5A2
InN ADARB2 FAM19A1
InN ADARB2 RGS10
InN ADARB2 CADPS2
InN ADARB2 PALMD
InN ADARB2 SERPI2
InN VIP MDGA1
InN ADARB2 PENK
InN VIP PENK
InN VIP EGF
InN VIP EXPH5
InN VIP MEGF11 LHFP
InN VIP SCML4
InN VIP ADAM12
InN VIP CLSTN2
InN VIP HS3ST3B1
InN VIP FREM2
InN LAMP5 ADAMTS20
InN LAMP5 BMP7
InN LAMP5 MEIS2
InN LAMP5 RELN
InN LAMP5 LHX6 TAC1
InN LAMP5 LHX6 PROX1
Inh L2–5 PVALB SCUBE3
Inh L4–6 PVALB SULF1
Inh L4–5 PVALB MEPE
Inh L5–6 SST MIR548F2
Inh L2–4 PVALB WFDC2
Inh L5–6 PVALB LGR5
Inh L5–6 SST KLHDC8A
Inh L4–6 SST B3GAT2
Inh L3–6 SST HPGD
Inh L4–5 SST STK32A
Inh L5–6 SST NPM1P10
Inh L4–6 SST GXYLT2
Inh L1–3 SST CALB1
Inh L2–4 SST FRZB
Inh L3–5 SST ADGRG6
Inh L1–2 SST BAGE2
Inh L3–6 SST NPY
Inh L1–2 ADARB2 MC4R
Inh L1–3 SST CHRNA4
Inh L1–4 VIP PENK
Inh L1–3 SST SYT6
Inh L1–2 VIP TSPAN12
Inh L1–2 PAX6 TNFAIP8L3
Inh L1–3 VIP ADAMTSL1
Inh L1–3 VIP CHRM2
Inh L2–5 VIP SERPINF1
Inh L2–4 VIP CBLN1
Inh L1–3 VIP CCDC184
Inh L2–3 VIP CASZ1
Inh L1–4 VIP CHRNA6
Inh L2–5 VIP TYR
Inh L1–2 VIP PCDH20
Inh L2–4 VIP SPAG17
Inh L1–4 VIP OPRM1
Inh L1–3 VIP GGH
Inh L3–6 VIP HS3ST3A1
Inh L2–6 VIP QPCT
Inh L1–2 LAMP5 DBP
Inh L1–4 LAMP5 LCP2
Inh L1 LAMP5 NMBR
Inh L2–6 LAMP5 CA1
C
Human PFC
Mouse ALM & VISp
InN PVALB PDE3A
InN PVALB GRIN2C
InN PVALB PIEZO2
InN PVALB SNTB1
InN SST ANKRD33B
InN PVALB HPSE
InN PVALB SYT2
InN PVALB ANOS1
InN SST ADRA1D
InN PVALB PRKCH
InN LHX6 HGF STON2
InN SST HGF GABRQ
InN SST KLHL14
InN SST FREM1
InN SST STK32A
InN SST THSD7B
InN SST HTR2C
InN SST ADAMTSL1
InN SST HS3ST5
InN SST PLPP4
InN SST TRPC7
InN SST NPY
InN ADARB2 COL5A2
InN ADARB2 FAM19A1
InN ADARB2 RGS10
InN ADARB2 CADPS2
InN ADARB2 PALMD
InN ADARB2 SERPI2
InN VIP MDGA1
InN ADARB2 PENK
InN VIP PENK
InN VIP EGF
InN VIP EXPH5
InN VIP MEGF11 LHFP
InN VIP SCML4
InN VIP ADAM12
InN VIP CLSTN2
InN VIP HS3ST3B1
InN VIP FREM2
InN LAMP5 ADAMTS20
InN LAMP5 BMP7
InN LAMP5 MEIS2
InN LAMP5 RELN
InN LAMP5 LHX6 TAC1
InN LAMP5 LHX6 PROX1
Pvalb Vipr2
Pvalb Akr1c18 Ntf3
Pvalb Calb1 Sst
Pvalb Tpbg
Pvalb Gabrg1
Pvalb Reln Tac1
Pvalb Sema3e Kank4
Pvalb Gpr149 Islr
Pvalb Reln Itm2a
Sst Myh8 Etv1
Sst Chrna2 Glra3
Sst Tac2 Tacstd2
Sst Esm1
Sst Hpse Sema3c
Sst Chrna2 Ptgdr
Sst Hpse Cbln4
Sst Tac1 Htr1d
Sst Calb2 Pdlim5
Sst Calb2 Necab1
Sst Chrna2 Glra3
Lamp5 Krt73
Vip Igfbp6 Pltp
Vip Lect1 Oxtr
Lamp5 Ntn1 Npy2r
Lamp5 Fam19a1 Tmem182
Vip Lmo1 Fam159b
Vip Col15a1 Pde1a
Vip Chat Htr1f
Vip Ptprt Pkp2
Lamp5 Lsp1
Lamp5 Plch2 Dock5
Lamp5 Lhx6
D
scaled expression
Human
Chimpanzee
low
high
Macaque
Marmoset
low
high
% expressed
25
50
75
100

Fig. S13. Transcriptomic organization and conservation of GABAergic inhibitory neuron subtypes. (A) Separate clustering of inhibitory neurons in each species with cells colored by subclasses. **(B)** Sankey diagram showing the correspondence of the inhibitory neuron subtypes defined in our data versus human medial temporal gyrus (MTG) (*16*). **(C)** As in B, but showing the comparison between human dlPFC and mouse primary visual cortex and anterior lateral motor cortex (VISp & ALM) (*17*). **(D)** Dot plots showing the expression patterns of canonical inhibitory neuron marker genes and species-conserved inhibitory neuron subtype marker genes. Dots are colored by species, with the color gradients representing the scaled average expression and sizes depicting the expression ratio of a certain gene, respectively.

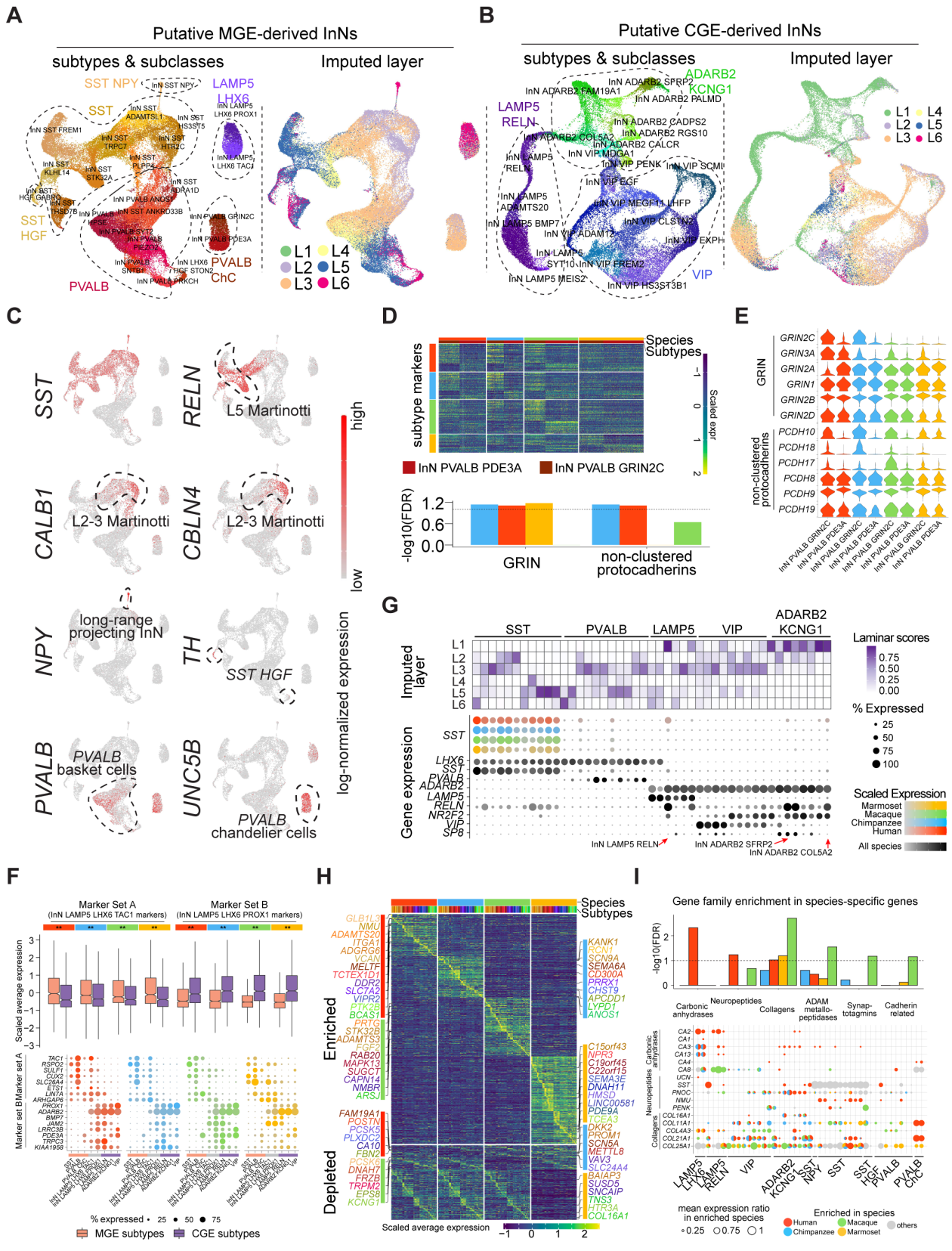
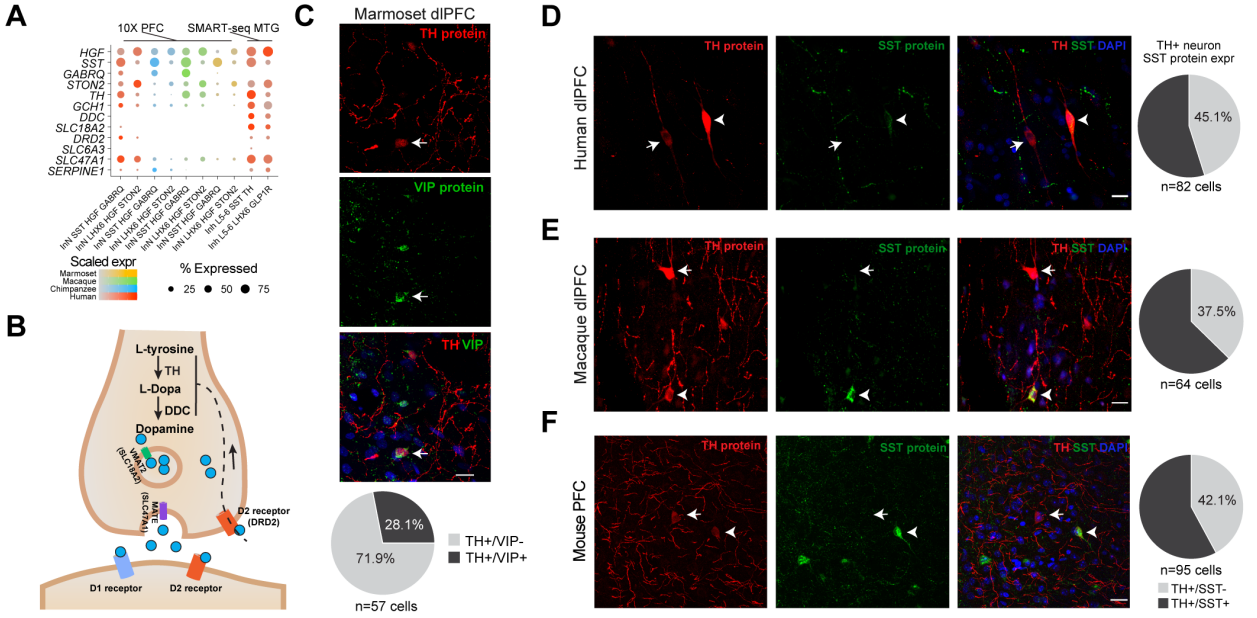
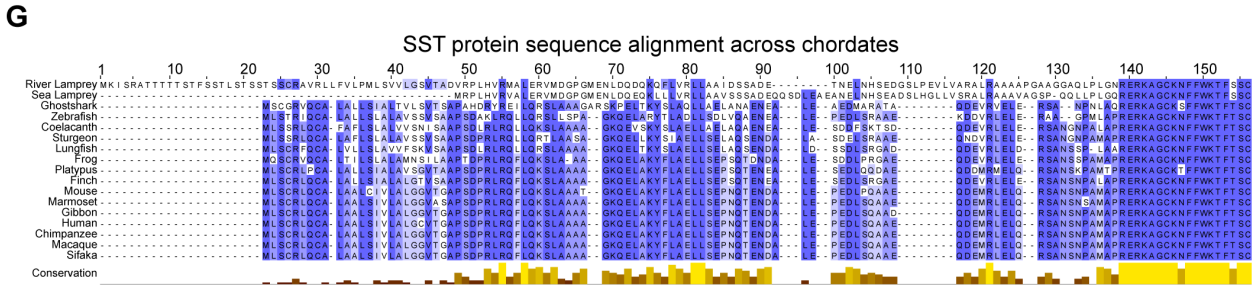
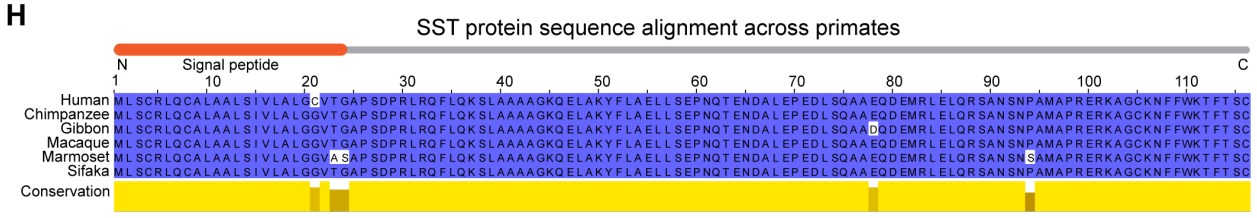

A
Putative MGE-derived InNs
subtypes & subclasses
Imputed layer
SST NPY
LAMP5 LHX6
SST
SST HGF
PVALB
PVALB ChC
L1 L2 L3 L4 L5 L6
B
Putative CGE-derived InNs
subtypes & subclasses
Imputed layer
ADARB2 KCNG1
LAMP5 RELN
VIP
C
SST
RELN
L5 Martinotti
CALB1
L2-3 Martinotti
CBLN4
L2-3 Martinotti
NPY
long-range projecting InN
TH
SST HGF
PVALB
PVALB basket cells
UNC5B
PVALB chandelier cells
high
low
log-normalized expression
D
Species
Subtypes
subtype markers
Scaled expr
InN PVALB PDE3A
InN PVALB GRIN2C
-log10(FDR)
GRIN
non-clustered protocadherins
E
GRIN
GRIN2C
GRIN3A
GRIN2A
GRIN1
GRIN2B
GRIN2D
non-clustered protocadherins
PCDH10
PCDH18
PCDH17
PCDH8
PCDH9
PCDH19
G
SST
PVALB
LAMP5
VIP
ADARB2 KCNG1
Imputed layer
Gene expression
Laminar scores
% Expressed
Scaled Expression
Marmoset
Macaque
Chimpanzee
Human
All species
InN LAMP5 RELN
InN ADARB2 SFRP2
InN ADARB2 COL5A2
F
Marker Set A
(InN LAMP5 LHX6 TAC1 markers)
Marker Set B
(InN LAMP5 LHX6 PROX1 markers)
Scaled average expression
Marker set B
Marker set A
% expressed
MGE subtypes
CGE subtypes
H
Species
Subtypes
Enriched
Depleted
Scaled average expression
I
Gene family enrichment in species-specific genes
-log10(FDR)
Carbonic anhydrases
Neuropeptides
Collagens
ADAM metallo-peptidases
Synap-totagmins
Cadherin related
mean expression ratio in enriched species
Enriched in species
Human
Macaque
Chimpanzee
Marmoset
others

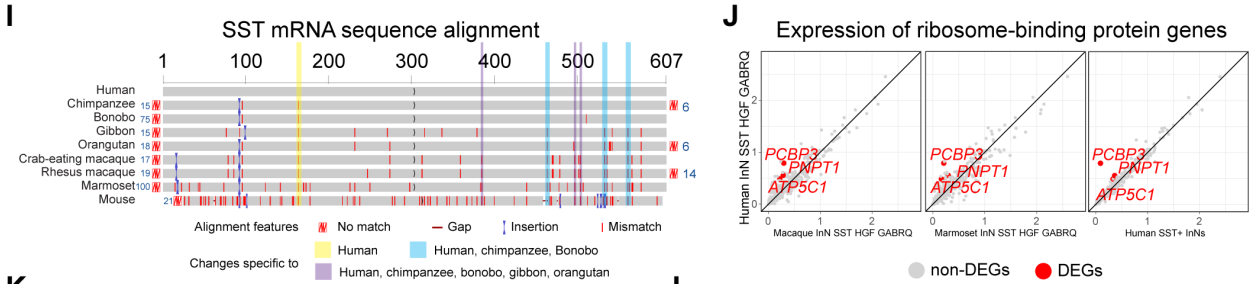
Fig. S14. Shared and divergent features of GABAergic inhibitory neuron subtypes. (A-B) UMAP showing putative MGE- (A) or CGE- (B) derived inhibitory neurons. **(C)** Expression of selected MGE inhibitory neuron markers in human MGE-derived inhibitory neurons. Inferred cell type identity is labeled in text. *RELN* is considered as a marker labeling majority of the SST Martinotti cells (*19, 31*). Based on transcriptomic integration (fig. S13C), subtypes inhibitory neuron *SST FREM1* and inhibitory neuron *SST STK32A* in our primate data were mapped to mouse L5 Martinotti cells (cluster *Sst Chrna2 Glra3* and *Sst Chrna2 Ptgdr*) (*17, 99*). Three other subtypes (InN *SST HS3ST5*, InN *SST ADAMTSL1* and InN *SST TRPC7*) marked by *CALB1* and *CBLN4* were mapped to mouse upper layer Martinotti cells (cluster *Sst Calb2 Necab1* and *Sst Calb2 Necab1*). **(D)** Top: expression of PVALB chandelier subtype markers. Bottom: enriched gene families in PVALB chandelier subtype markers. GRIN: glutamate ionotropic receptor NMDA Subunit 1. (**E**) Expression of selected GRIN and non-clustered cadherin genes that can distinguish the two *PVALB* chandelier subtypes. (**F**) Top: The markers of subtype InN *LAMP5 LHX6 TAC1* are enriched in MGE-derived inhibitory neuron subtypes while the markers of subtype InN *LAMP5 LHX6 PROX1* are enriched in CGE-derived inhibitory neuron subtypes. Bottom: Expression of selective markers from the two subtypes demonstrated their differential similarity to MGE- and CGE-derived inhibitory neurons. (**G**) *SST* expression across inhibitory neuron subtypes. There are 3 L1 inhibitory neuron subtypes (arrows) showing human-enriched *SST* expression. (**H**) Standardized average expression of species-enriched or -depleted genes across inhibitory neuron subtypes. Subtype colors are consistent with Figure 4A. Genes were colored according to the subtypes where they exhibit enrichment or depletion. (**I**) Top: Gene family enrichment in genes displaying species-specific expression in inhibitory neurons, with color denoting species. The dashed line represents the cutoff of false discovery rate (FDR) at 0.1. Bottom: Weighted pie plots

summarizing the expression patterns of the genes in the human-enriched gene families. The colors denote the species enrichment with grey color representing conserved expression. The sizes of pies represent the average expression ratios of genes in the enriched species (species-specific) or all species (conserved expression). Here, the carbonic anhydrases gene family was reported to regulate neuronal signaling through modulation of pH transients (*100*).


A
10X PFC
SMART-seq MTG
HGF
SST
GABRQ
STON2
TH
GCH1
DDC
SLC18A2
DRD2
SLC6A3
SLC47A1
SERPINE1
Scaled expr
% Expressed
B
L-tyrosine
TH
L-Dopa
DDC
Dopamine
VMAT2 (SLC18A2)
MATE (SLC47A1)
D2 receptor (DRD2)
D1 receptor
D2 receptor
C
Marmoset dlPFC
TH protein
VIP protein
TH VIP
28.1%
71.9%
TH+/VIP-
TH+/VIP+
n=57 cells
D
Human dlPFC
TH protein
SST protein
TH SST DAPI
TH+ neuron SST protein expr
45.1%
n=82 cells
E
Macaque dlPFC
37.5%
n=64 cells
F
Mouse PFC
42.1%
n=95 cells
TH+/SST-
TH+/SST+


G
SST protein sequence alignment across chordates
River Lamprey
Sea Lamprey
Ghostshark
Zebrafish
Coelacanth
Sturgeon
Lungfish
Frog
Platypus
Finch
Mouse
Marmoset
Gibbon
Human
Chimpanzee
Macaque
Sifaka
Conservation


H
SST protein sequence alignment across primates
N
Signal peptide
C
Human
Chimpanzee
Gibbon
Macaque
Marmoset
Sifaka
Conservation


I
SST mRNA sequence alignment
Human
Chimpanzee
Bonobo
Gibbon
Orangutan
Crab-eating macaque
Rhesus macaque
Marmoset
Mouse
Alignment features
No match
Gap
Insertion
Mismatch
Changes specific to
Human
Human, chimpanzee, Bonobo
Human, chimpanzee, bonobo, gibbon, orangutan
J
Expression of ribosome-binding protein genes
Human InN SST HGF GABRQ
Macaque InN SST HGF GABRQ
Marmoset InN SST HGF GABRQ
Human SST+ InNs
PCBP3
PNPT1
ATP5C1
non-DEGs
DEGs

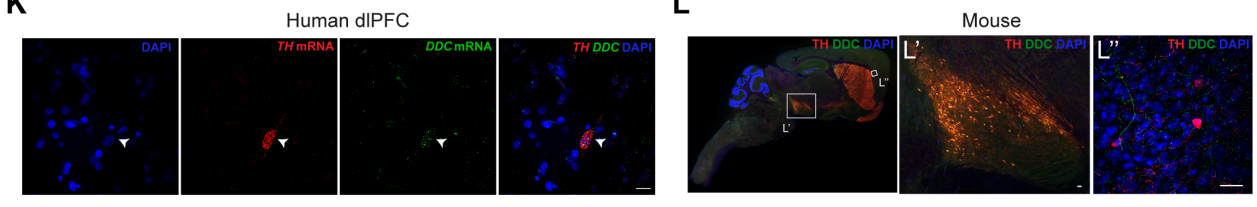

K
Human dlPFC
DAPI
TH mRNA
DDC mRNA
TH DDC DAPI
L
Mouse
TH DDC DAPI

Fig. S15. Multicolor immunofluorescence staining with antibodies against SST, TH and DDC. **(A)** Dot plot showing the expression of cluster markers and key components of dopamine synthesis and signaling in our dataset (columns 1-8), as well as the published human MTG dataset (columns 9-10) (*16*). Subtypes *Inh L5−6 SST TH* and *Inh L6 LHX6 GLP1R* are matched to subtypes InN *SST HGF GABRQ* and InN *LHX6 HGF STON2*, respectively. The depletion of *TH* expression in chimpanzees is correlated with the specific upregulation of *SERPINE1*, a gene encoding protein reported to suppress dopamine release (*101*). (**B**) Diagram showing the hypothesis of how each key component participate in the dopamine synthesis and signaling. **(C)** TH and VIP co-labeling in marmosets. n = 3. Scale bar: 20μm. **(D-E)** Immunofluorescence of TH and SST with tyramide signal amplification in human (D) and macaque (E) dlPFC tissues. Scale bars: 20μm for each species. (**F**) Immunofluorescence of TH and SST without tyramide signal amplification in mouse PFC tissues. Scale bar: 20μm. (**G-H**) SST protein sequence alignment across multiple Chordata (G) and primate species (H). SST encodes a Chordate-specific secreted protein with protein domain organization. The sequence changes of SST proteins in chordates are largely centered on the regulatory regions (N terminal and middle), but are depleted in the C terminal where the mature SST resides. Entering the primate lineage, SST protein sequences become highly conserved, with a human-specific amino acid change at position 21: glycine (Gly) to cysteine (Cys). Consistent with Cys depletion among human pathogenic variants (*82*), at residue Cys21 the human population gnomAD database (*81*) records one rare benign p.Cys21Tyr SST1 variant that is only tolerated as a heterozygous allele. (**I**) SST mRNA sequence alignment across multiple primate species and mouse. (**J**) Ribosome-binding protein genes (*102*) enriched in the human InN *SST HGF GABRQ* subtype compared to other SST interneurons as well as the homolog subtypes in macaques and marmosets. (**K**) RNA in situ hybridization of *TH* and *DDC* in human dlPFC tissues. Scale bar:

20μm. (**L**) Immunofluorescence of TH and DDC in mouse brain. The proteins are colocalized in substantia nigra/ventral tegmental area (L'), but not in mouse cortex (L''). Scale bar: 20μm.

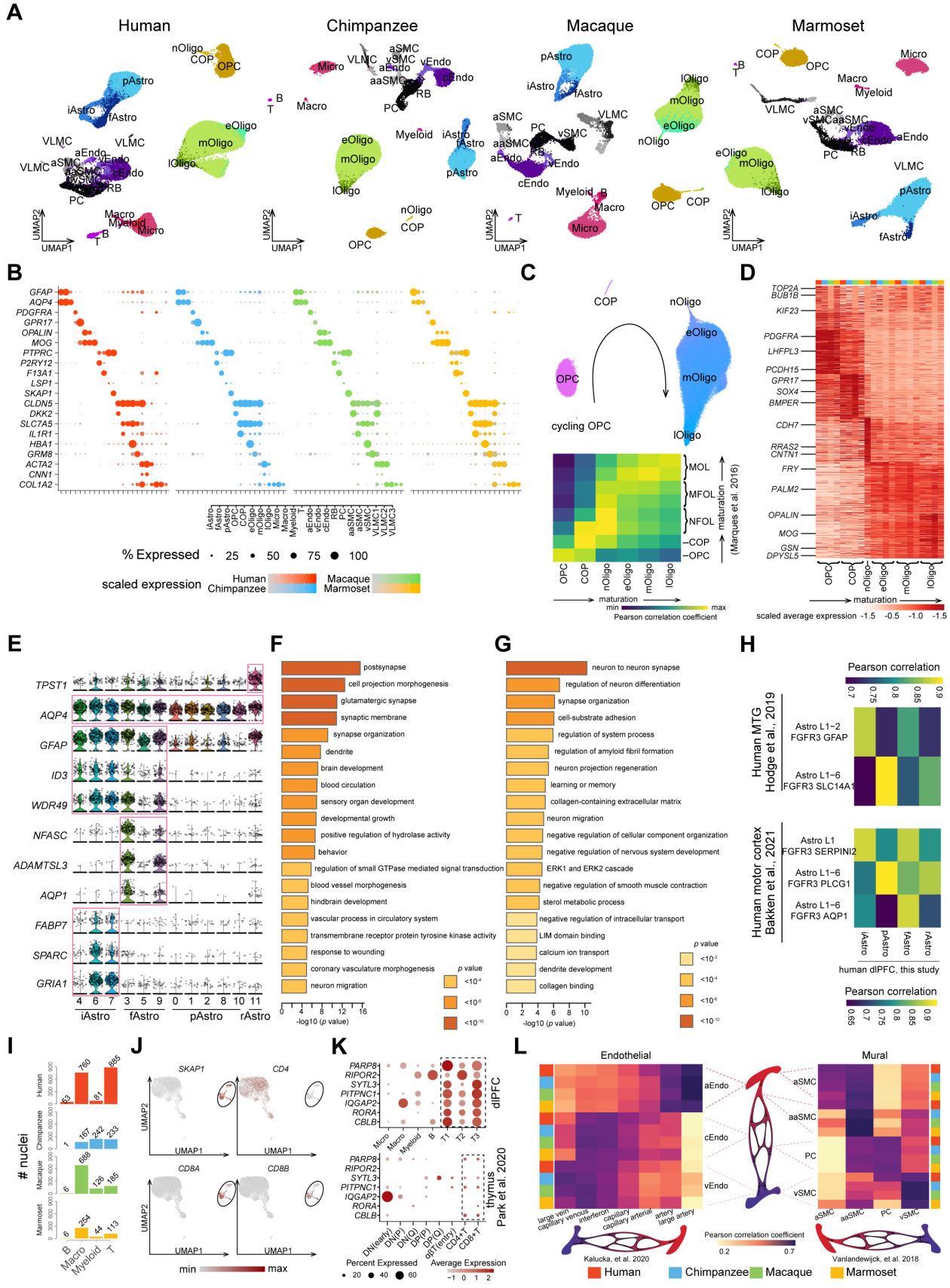

A
Human
Chimpanzee
Macaque
Marmoset
UMAP1
UMAP2
B
GFAP
AQP4
PDGFRA
GPR17
OPALIN
MOG
PTPRC
P2RY12
F13A1
LSP1
SKAP1
CLDN5
DKK2
SLC7A5
IL1R1
HBA1
GRM8
ACTA2
CNN1
COL1A2
% Expressed 25 50 75 100
scaled expression
Human
Chimpanzee
Macaque
Marmoset
C
cycling OPC
MOL
MFOL
NFOL
COP
OPC
maturation
(Marques et al. 2016)
Pearson correlation coefficient
D
TOP2A
BUB1B
KIF23
PDGFRA
LHFPL3
PCDH15
GPR17
SOX4
BMPER
CDH7
RRAS2
CNTN1
FRY
PALM2
OPALIN
MOG
GSN
DPYSL5
scaled average expression
E
TPST1
AQP4
GFAP
ID3
WDR49
NFASC
ADAMTSL3
AQP1
FABP7
SPARC
GRIA1
iAstro
fAstro
pAstro
rAstro
F
postsynapse
cell projection morphogenesis
glutamatergic synapse
synaptic membrane
synapse organization
dendrite
brain development
blood circulation
sensory organ development
developmental growth
positive regulation of hydrolase activity
behavior
regulation of small GTPase mediated signal transduction
blood vessel morphogenesis
hindbrain development
vascular process in circulatory system
transmembrane receptor protein tyrosine kinase activity
response to wounding
coronary vasculature morphogenesis
neuron migration
p value
-log10 (p value)
G
neuron to neuron synapse
regulation of neuron differentiation
synapse organization
cell-substrate adhesion
regulation of system process
regulation of amyloid fibril formation
neuron projection regeneration
learning or memory
collagen-containing extracellular matrix
neuron migration
negative regulation of cellular component organization
negative regulation of nervous system development
ERK1 and ERK2 cascade
negative regulation of smooth muscle contraction
sterol metabolic process
negative regulation of intracellular transport
LIM domain binding
calcium ion transport
dendrite development
collagen binding
p value
-log10 (p value)
H
Pearson correlation
Human MTG
Hodge et al., 2019
Astro L1–2
FGFR3 GFAP
Astro L1–6
FGFR3 SLC14A1
Human motor cortex
Bakken et al., 2021
Astro L1
FGFR3 SERPINI2
Astro L1–6
FGFR3 PLCG1
Astro L1–6
FGFR3 AQP1
human dlPFC, this study
I
nuclei
Human
Chimpanzee
Macaque
Marmoset
B
Macro
Myeloid
T
J
SKAP1
CD4
CD8A
CD8B
min
max
K
PARP8
RIPOR2
SYTL3
PITPNC1
IQGAP2
RORA
CBLB
dlPFC
thymus
Park et al. 2020
Percent Expressed
Average Expression
DN(early)
DN(P)
DN(Q)
DP(P)
DP(Q)
αβT(entry)
CD4-T
CD8-T
L
Endothelial
Mural
aEndo
cEndo
vEndo
aSMC
aaSMC
PC
vSMC
large vein
capillary venous
interferon
capillary
capillary arterial
artery
large artery
Kalucka, et al. 2020
Pearson correlation coefficient
Vanlandewijck, et al. 2018
Human
Chimpanzee
Macaque
Marmoset

Fig. S16. Taxonomy of glia cell and non-neural cell subtypes. (**A**) UMAP visualizations of glia cells and non-neural cells in humans, chimpanzees, macaques and marmosets. (**B**) Expression of shared subclass marker genes across homolog subtypes and species. Homolog subtypes with small cluster size (< 15 cells) were not included for visualization. (**C**) Top: UMAP showing subtypes in the oligodendrocyte differentiation and maturation trajectory (indicated by the black arrow). Bottom: Transcriptome similarities (Pearson correlation coefficients) between subtypes from this study (columns) and those from the mouse oligodendrocytes (rows) (*21*). (**D**) Scaled average expression of cell state-restricted genes (rows) in the subtypes along oligodendrocyte differentiation and maturation (columns). Species information is color-barcoded on the top. Because there are less than 10 nOligos in humans, chimpanzees and marmosets, only the macaque nOligos are included here for visualization. (**E**) Expression of astrocyte subtype markers across the 12 astrocyte subclusters with subtype labeled. (**F-G**) Bar plot exhibiting the gene ontology enrichment of module one (F) and module two (G) genes from Fig. 5D. (**H**) Transcriptomic comparisons between this data and the human MTG data (top) (*16*) as well as motor cortex data (top) (*22*). Astro L1-2 FG*FR3 GFAP* and Astro L1 *FGFR3 SERPINI2*: putative iAstro; Astro L1-6 *FGFR3 SLC14A1* and Astro L1-6 *FGFR3 PLCG1*: putative pAstro; Astro L1-6 *FGFR3*: putative fAstro. (**I**) The numbers of immune cell subtype in each of the four primates. (**J**) Human T cells (marked by circles) include predominant expression of *CD8A* and *CD8B* surface markers rather than *CD4*, identifying them as putative CD8+ T cells that may be more likely to infiltrate the brain for surveillance of dysfunctional cells (*103*). T cells are marked by the circles. (**K**) Expression of novel brain-resident T cell markers in our data (top) and in the thymus (bottom) (*104*). Genes expressed in more than 50% of cells in brain T cells and less than 10% of cells in thymic T cells were selected for visualization. (**L**) Middle and bottom: Illustration of the brain vascular

architecture from arteries, arterioles, capillaries to veins. Left and right: Transcriptome similarities (Pearson correlation coefficients) between subtypes from this study (rows) and those from mouse brains (columns) (*23, 24*).

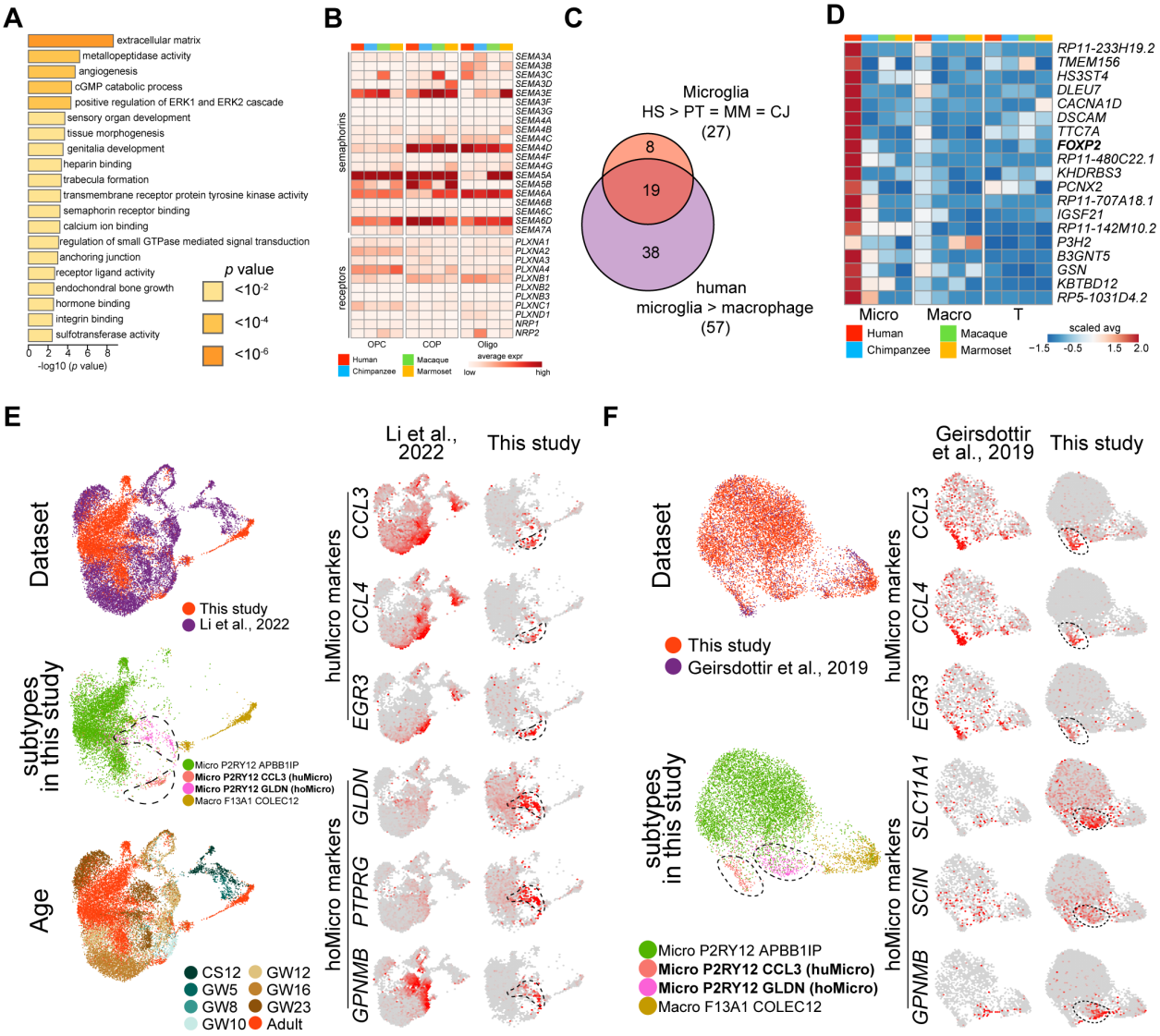


Fig. S17. Species divergence of glia cell and non-neural cell subtypes within the dlPFC across the investigated primates. (**A**) Bar plot exhibiting the gene ontology enrichment of the union of evolutionarily divergent genes along the oligodendrocyte differentiation and maturation processes. (**B**) Heat map displaying the average expression of genes encoding semaphorins (upper rows) and of genes encoding their receptors (lower rows) in OPCs, COPs and oligodendrocytes (Oligos) of the four primate species (columns). Species information is color-barcoded on the top. OPC: oligodendrocyte precursor cells; COP: committed OPCs; Oligo: oligodendrocytes; (**C**) Venn diagram showing the overlap between the 27 genes specifically enriched in microglia of humans

versus other species and the 57 genes significantly overrepresented in human microglia versus macrophages. HS: human; PT: chimpanzee; MM: macaque; CJ: marmoset. (**D**) Heat map of the resulting 19 genes from (C) showing their scaled average expression in microglia, macrophages and T cells of the four primate species. Species information is color-barcoded on the top. (**E**) Left: integration between this study and the fetal human microglia dataset (*105*). Right: hexagon visualization on the UMAP layout showing the shared markers of hoMicro and huMicro between the two datasets. (**F**) Same as in (E), but showing the integration between this study and another adult human microglia dataset (*25*). huMicro: human-specific microglia subtype (Micro *P2RY12 CCL3*); hoMicro: Hominini-specific microglia subtype (Micro *P2RY12 GLDN*).

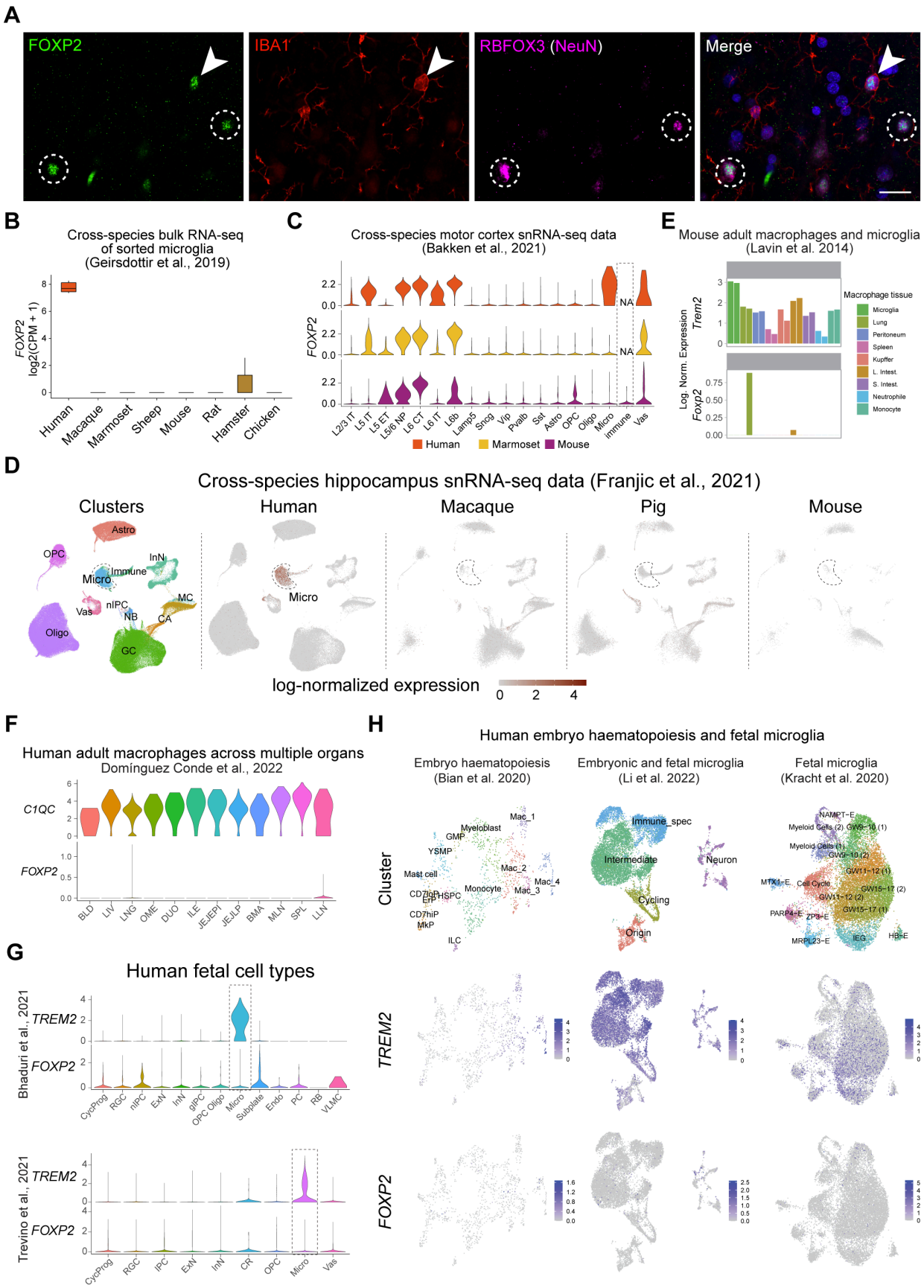

A
FOXP2
IBA1
RBFOX3 (NeuN)
Merge
B
Cross-species bulk RNA-seq of sorted microglia (Geirsdottir et al., 2019)
FOXP2 log2(CPM + 1)
Human
Macaque
Marmoset
Sheep
Mouse
Rat
Hamster
Chicken
C
Cross-species motor cortex snRNA-seq data (Bakken et al., 2021)
FOXP2
NA
L2/3 IT
L5 IT
L5 ET
L5/6 NP
L6 CT
L6 IT
L6b
Lamp5
Sncg
Vip
Pvalb
Sst
Astro
OPC
Oligo
Micro
Immune
Vas
Human
Marmoset
Mouse
E
Mouse adult macrophages and microglia (Lavin et al. 2014)
Trem2
Foxp2
Log. Norm. Expression
Macrophage tissue
Microglia
Lung
Peritoneum
Spleen
Kupffer
L. Intest.
S. Intest.
Neutrophile
Monocyte
D
Cross-species hippocampus snRNA-seq data (Franjic et al., 2021)
Clusters
Human
Macaque
Pig
Mouse
Astro
OPC
InN
Immune
Micro
MC
nIPC
Vas
NB
CA
Oligo
GC
log-normalized expression
0
2
4
F
Human adult macrophages across multiple organs
Domínguez Conde et al., 2022
C1QC
FOXP2
BLD
LIV
LNG
OME
DUO
ILE
JEJEPI
JEJLP
BMA
MLN
SPL
LLN
G
Human fetal cell types
Bhaduri et al., 2021
TREM2
FOXP2
CycProg
RGC
nIPC
ExN
InN
gIPC
OPC Oligo
Micro
Subplate
Endo
PC
RB
VLMC
Trevino et al., 2021
TREM2
FOXP2
CycProg
RGC
IPC
ExN
InN
CR
OPC
Micro
Vas
H
Human embryo haematopoiesis and fetal microglia
Embryo haematopoiesis (Bian et al. 2020)
Embryonic and fetal microglia (Li et al. 2022)
Fetal microglia (Kracht et al. 2020)
Cluster
Mac_1
Myeloblast
GMP
YSMP
Mast cell
Mac_2
Mac_4
Mac_3
CD7loP HSPC
Monocyte
CD7hiP
MkP
ILC
Immune_spec
Intermediate
Neuron
Cycling
Origin
NAMPT−E
Myeloid Cells (2) GW9−10 (1)
Myeloid Cells (1)
GW9−10 (2)
GW11−12 (1)
MTX1−E
Cell Cycle
GW15−17 (2)
GW11−12 (2)
GW15−17 (1)
PARP4−E
ZFP36−E
MRPL23−E
IEG
HB−E
TREM2
FOXP2

Fig. S18. Human-specific expression of *FOXP2* in microglia. (A) Triple immunofluorescent validation that FOXP2 is detected in IBA1-positive microglia and RBFOX3/NeuN-positive neurons in human dlPFC. Scale bar: 20μm. **(B-E)** Expression of *FOXP2* in RNA-seq data of sorted bulk microglia from eight species (B) (*25*), cross-species snRNA-seq data from motor cortex (C) (*22*), cross-species snRNA-seq data from hippocampus (D) (*41*), cross-tissue RNA-seq data of adult mouse macrophages and microglia (E) (*106*). These expression patterns confirm the human-specific *FOXP2* expression in microglia. (**F**) Expression of *FOXP2* across macrophages from multiple adult human organs (*107*) suggests *FOXP2* expression is specific to brain microglia. BLD: blood; LIV: liver; LNG: lung; OME: omentum; DUO: duodenum; ILE: ileum; JEJEPI: Jejunum epithelial fractions; JEJLP: Jejunum lamina propria; BMA: bone marrow; MLN: mesenteric lymph nodes; SPL: spleen; LLN: lung-draining lymph nodes. (**G-H**) *FOXP2* expression is not detected in human embryonic and fetal microglia (*38, 105, 108-110*).

Fig. S19. FOXP2 immunostaining in non-primate mammalian neocortex. (A) Immunostaining of FOXP2 in the cerebral cortex of non-primate mammals. Scale bar is 100μm for each species.

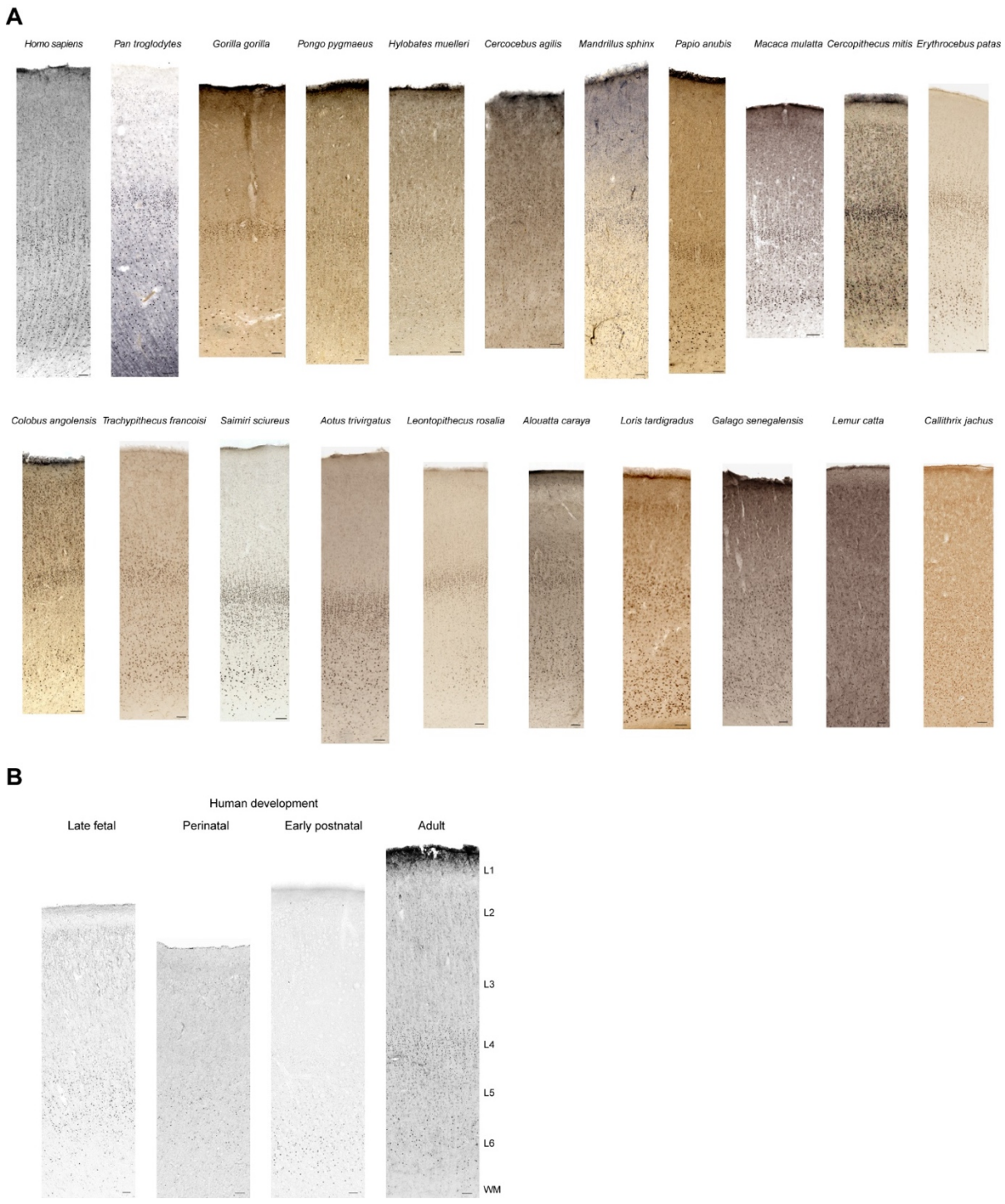


Fig. S20. FOXP2 immunostaining in primate neocortex. (A) Immunostaining of FOXP2 in primate cerebral cortex. Scale bar is 100μm for each species. **(B)** Immunostaining of FOXP2 in late fetal, perinatal, early postnatal and adult human dlPFC tissues. Scale bars: 100μm.

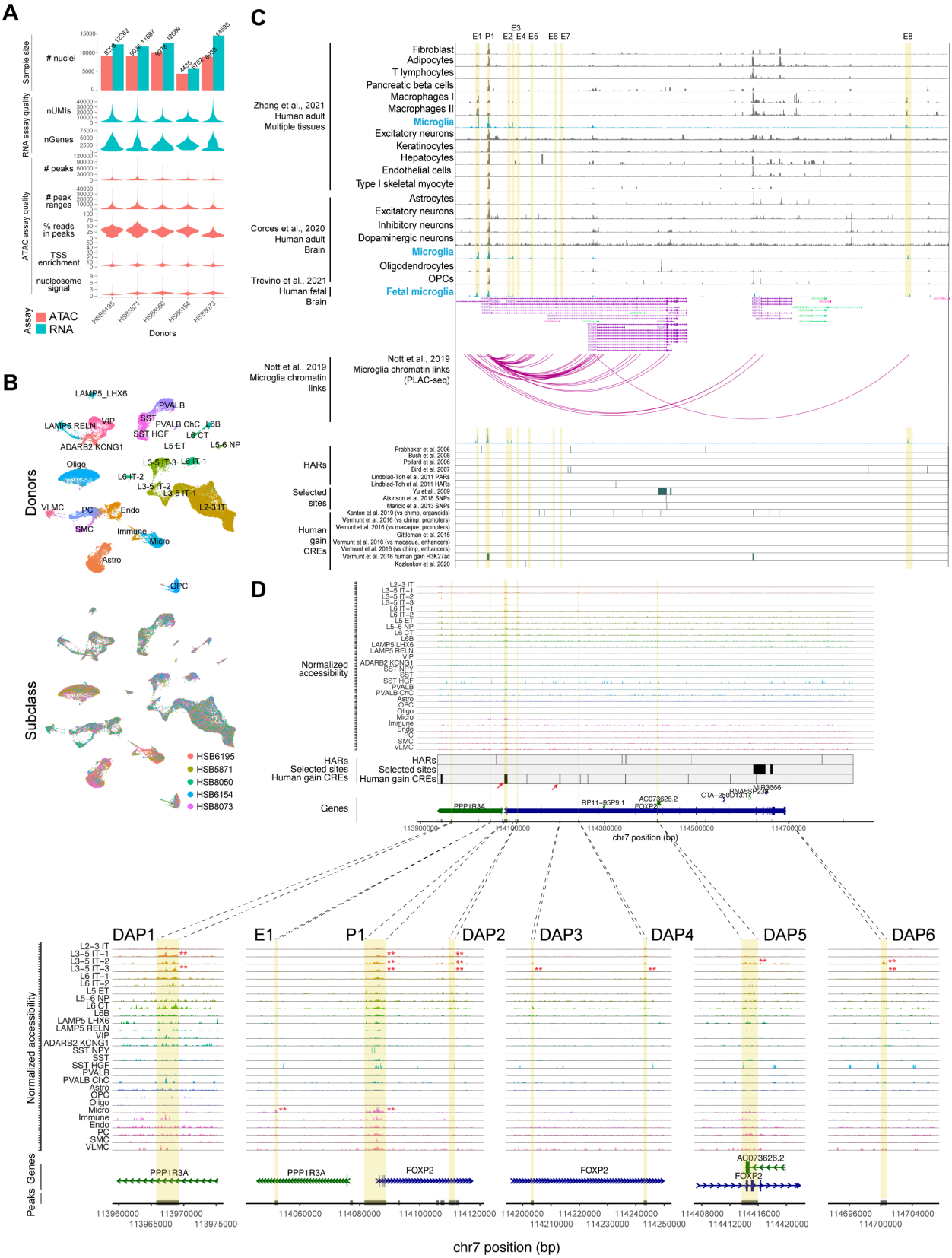

A
Sample size
RNA assay quality
ATAC assay quality
nuclei
nUMIs
nGenes
peaks
peak ranges
% reads in peaks
TSS enrichment
nucleosome signal
Assay
ATAC
RNA
Donors
HSB6195
HSB5871
HSB8050
HSB6154
HSB8073
B
Donors
Subclass
LAMP5_LHX6
PVALB
SST
VIP
LAMP5 RELN
PVALB ChC
L6B
SST HGF
L6 CT
ADARB2 KCNG1
L5 ET
L5-6 NP
Oligo
L6 IT-2
L6 IT-1
L3-5 IT-3
L3-5 IT-2
L3-5 IT-1
L2-3 IT
VLMC
PC
Endo
SMC
Immune
Micro
Astro
OPC
C
E1 P1
E3
E2 E4 E5
E6 E7
E8
Zhang et al., 2021
Human adult
Multiple tissues
Fibroblast
Adipocytes
T lymphocytes
Pancreatic beta cells
Macrophages I
Macrophages II
Microglia
Excitatory neurons
Keratinocytes
Hepatocytes
Endothelial cells
Type I skeletal myocyte
Corces et al., 2020
Human adult
Brain
Astrocytes
Excitatory neurons
Inhibitory neurons
Dopaminergic neurons
Microglia
Oligodendrocytes
OPCs
Trevino et al., 2021
Human fetal
Brain
Fetal microglia
Nott et al., 2019
Microglia chromatin
links
Nott et al., 2019
Microglia chromatin links
(PLAC-seq)
HARs
Selected sites
Human gain CREs
Prabhakar et al. 2006
Bush et al. 2008
Pollard et al. 2006
Bird et al. 2007
Lindblad-Toh et al. 2011 PARs
Lindblad-Toh et al. 2011 HARs
Yu et al., 2009
Atkinson et al. 2018 SNPs
Maricic et al. 2013 SNPs
Kanton et al. 2019 (vs chimp, organoids)
Vermunt et al. 2016 (vs chimp, promoters)
Vermunt et al. 2016 (vs macaque, promoters)
Gittleman et al. 2015
Vermunt et al. 2016 (vs macaque, enhancers)
Vermunt et al. 2016 (vs chimp, enhancers)
Vermunt et al. 2016 human gain H3K27ac
Kozlenkov et al. 2020
D
Normalized accessibility
HARs
Selected sites
Human gain CREs
Genes
PPP1R3A
RP11-95P9.1
AC073626.2
FOXP2
CTA-250D13.1
MIR3666
RN45SP253
chr7 position (bp)
DAP1
E1
P1
DAP2
DAP3
DAP4
DAP5
DAP6
Peaks
chr7 position (bp)

Fig. S21. Regulatory mechanisms underlying the cell type- and species-specific *FOXP2* expression. (A) Top: bar plots showing the number of cells passed quality control in the RNA assay and ATAC assay of sn-multiome data. Bottom: violin plots showing sample quality metrics. (**B**) UMAP visualizations showing subclass and donor information in the RNA assay. (**C**) Coverage plots showing the approximate coordinates of promoter (P1) and multiple cis-regulatory elements (E1-E8) around the *FOXP2* gene body. We identified one cis-regulatory element specific to adult and fetal human microglia across multiple datasets (*85, 109, 111*). This enhancer topologically interacts with the *FOXP2* gene body as shown by the Proximity ligation-assisted ChIP-Seq (PLAC-Seq) data (*43*). However, none of the cis-regulatory elements overlaps with any reported human accelerated regions (HARs), human-specific single nucleotide polymorphisms (SNPs) or human gain cis-regulatory elements (CREs) (*14, 112-122*). Only the promoter region is overlapped with one human gain CRE. (**D**) Differentially accessible peaks (DAP) across subtypes proximal to *FOXP2*. Subclasses showing enrichment of DAPs were indicated by stars (**: adjusted $p < 0.01$). We only detected DAPs in L3-5 IT and microglia subtypes. One DAP (E1) is overlapped with the microglia-specific enhancer reported in panel C and another (P1) is overlapped with the *FOXP2* promoter region. DAP3 enriched in L3-5 IT neurons is overlapped with one reported human gain cis-regulatory element (*120*).

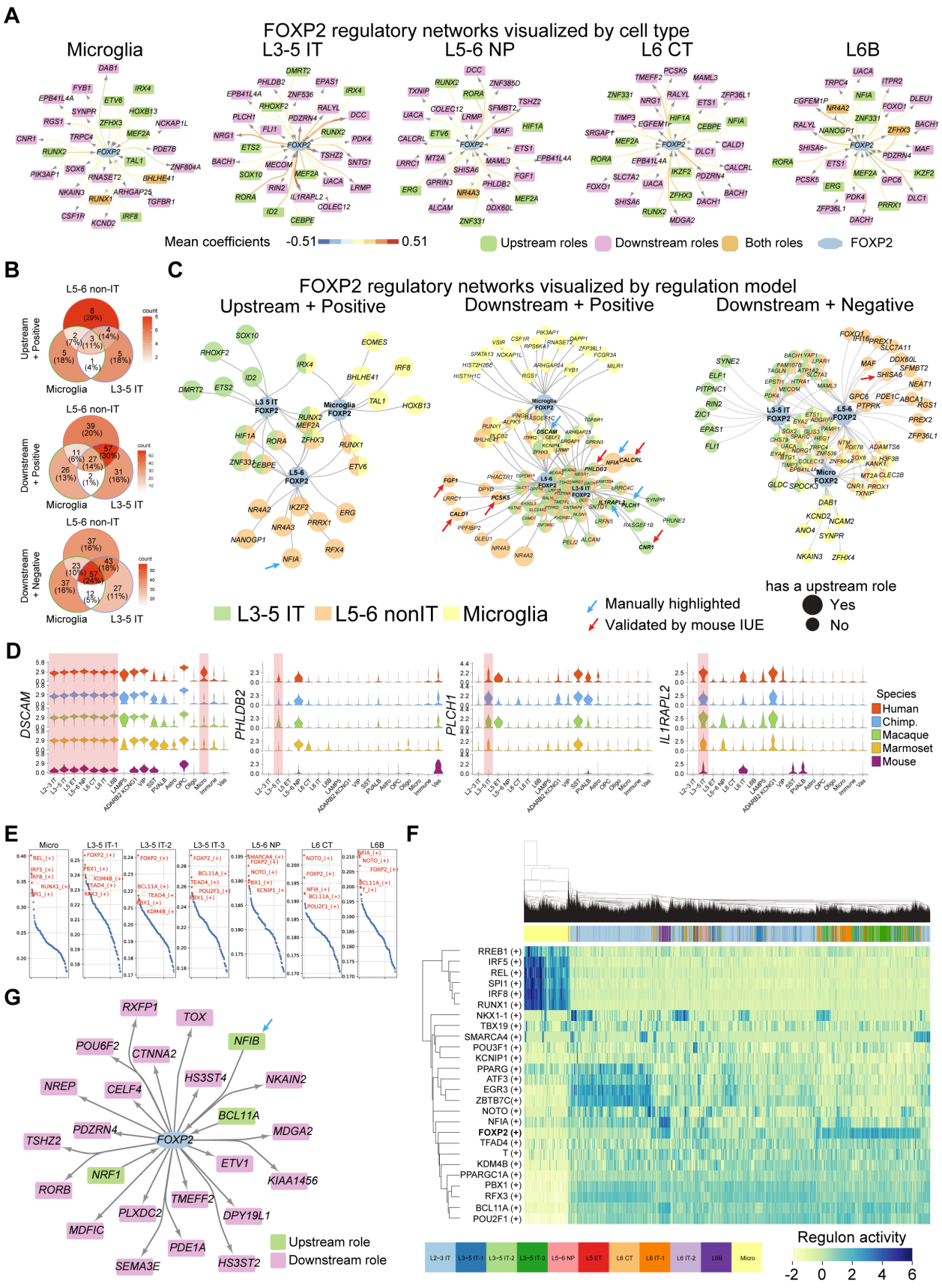

A
FOXP2 regulatory networks visualized by cell type
Microglia
L3-5 IT
L5-6 NP
L6 CT
L6B
Mean coefficients -0.51 0.51
Upstream roles
Downstream roles
Both roles
FOXP2
B
L5-6 non-IT
Upstream + Positive
Microglia
L3-5 IT
Downstream + Positive
Downstream + Negative
C
FOXP2 regulatory networks visualized by regulation model
Upstream + Positive
Downstream + Positive
Downstream + Negative
L3-5 IT
L5-6 nonIT
Microglia
Manually highlighted
Validated by mouse IUE
has a upstream role
Yes
No
D
DSCAM
PHLDB2
PLCH1
IL1RAPL2
Species
Human
Chimp.
Macaque
Marmoset
Mouse
E
Micro
L3-5 IT-1
L3-5 IT-2
L3-5 IT-3
L5-6 NP
L6 CT
L6B
F
RREB1 (+)
IRF5 (+)
REL (+)
SPI1 (+)
IRF8 (+)
RUNX1 (+)
NKX1-1 (+)
TBX19 (+)
SMARCA4 (+)
POU3F1 (+)
KCNIP1 (+)
PPARG (+)
ATF3 (+)
EGR3 (+)
ZBTB7C(+)
NOTO (+)
NFIA (+)
FOXP2 (+)
TFAD4 (+)
T (+)
KDM4B (+)
PPARGC1A (+)
PBX1 (+)
RFX3 (+)
BCL11A (+)
POU2F1 (+)
Regulon activity
-2 0 2 4 6
G
RXFP1
TOX
NFIB
POU6F2
CTNNA2
HS3ST4
NKAIN2
NREP
CELF4
BCL11A
TSHZ2
PDZRN4
FOXP2
MDGA2
ETV1
NRF1
KIAA1456
RORB
TMEFF2
PLXDC2
DPY19L1
MDFIC
PDE1A
HS3ST2
SEMA3E
Upstream role
Downstream role

Fig. S22. CellOracle- and pySCENIC-predicted *FOXP2* regulatory networks. (A) *FOXP2* regulatory networks predicted in multiple *FOXP2*-expressing subclasses in humans. We defined three models of regulation: 1) *FOXP2*-upstream regulators with positive regulation, 2) downstream targets positively regulated by *FOXP2*, 3) downstream targets negatively regulated by *FOXP2*. In each subclass, only top 10 genes of each regulatory model ranked by their expression specificity (descending order for positive regulation and ascending order for negative regulation) are visualized in the plot. Genes additionally validated by the mouse in utero electroporation experiments were also included for visualization. Positive regulation (red) and negative regulation (blue) are indicated by the edge colors. (**B**) Venn diagrams illustrating the overlap of FOXP2 regulators/targets across the three cell groups for each regulatory model. For simplicity, L5-6 NP, L6CT and L6B subclasses were merged to the "L5-6 non-IT" group. (**C**) FOXP2 regulatory networks visualized separately by regulatory models. Interesting genes (blue arrows) and the targets validated by the mouse in utero electroporation experiments (red arrows) are highlighted by arrows. Only the top 25 genes in each cell type and each regulation model were visualized for conciseness. Here *NFIA* is predicted as a downstream target of *FOXP2* in L5-6 non-IT group, which confirms previous studies (*44*). (**D**) Expression of selective *FOXP2* targets across subclasses and species. Here, *DSCAM* is predicted as a target positively regulated by *FOXP2* and it also shows human-specific expression in microglia. Likewise, three other genes are positively regulated by *FOXP2* in L3-5 IT and they also display enriched expression in L3-5 IT of the analyzed primates compared to mice. (**E-G**) Overview of the specificity scores (E), activity (F) and networks (G) of the pySCENIC-predicted regulons.

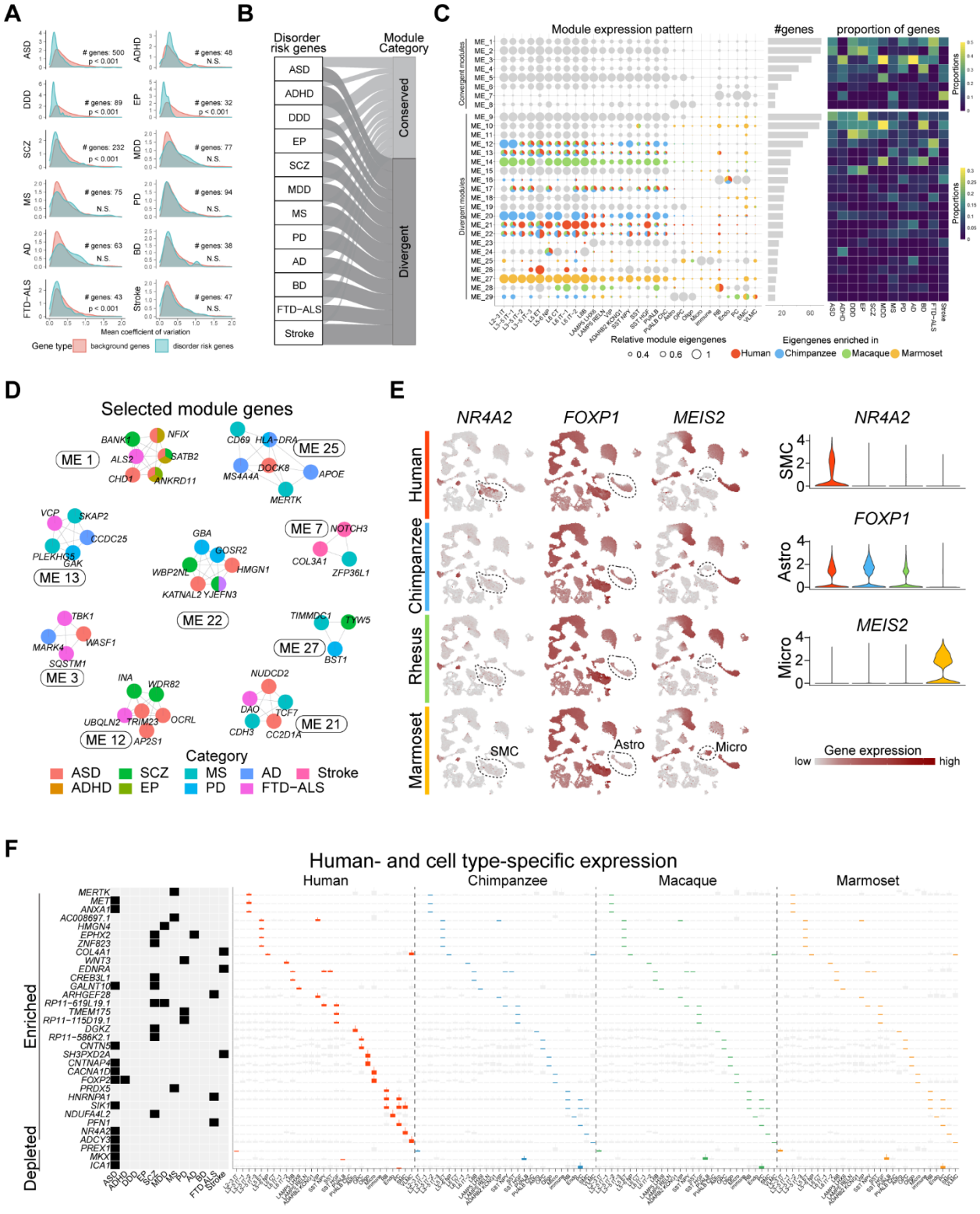

A
B
C
D
E
F
Mean coefficient of variation
Gene type
background genes
disorder risk genes
Disorder risk genes
Module Category
Conserved
Divergent
Module expression pattern
#genes
proportion of genes
Relative module eigengenes
Eigengenes enriched in
Human
Chimpanzee
Macaque
Marmoset
Selected module genes
Category
NR4A2
FOXP1
MEIS2
SMC
Astro
Micro
Gene expression
low
high
Human- and cell type-specific expression
Enriched
Depleted

Fig. S23. Species-conserved and -divergent expression of risk genes associated with twelve major brain disorders. (**A**) Distribution of mean coefficient of variation (CV) in disease risk genes and background genes. The significance was tested using Wilcoxon Rank Sum test. (**B**) Disease risk genes were assigned to either convergent and divergent modules based on their expression patterns (see materials and methods for details). The width of each ribbon represents the proportion of genes belonging to a specific category. 35% of the disease risk genes were assigned to the conserved expression modules. (**C**) Left: pie plots showing the eigengenes (PC1) of convergent and divergent modules summarized from the brain disorder genes. ME: module eigengenes. The colors of pies represent the species in which the module eigengenes are enriched. The grey colors denote comparable expression across species. Middle: gene module sizes. Right: proportions of the disease risk genes assigned to different modules. While this result shows the majority of the modules are enriched in neuron subclasses, it also shows preferential enrichment of stroke risk genes in multiple vascular subclasses (module 7 and module 16) and AD risk genes in immune cells (module 25), commensurate with previous genome-wide association or snRNA-seq studies (*11, 90, 123*). (**D**) Co-expression networks showing representative genes from the selected modules, with nodes representing genes and edges denoting Pearson correlation coefficients between nodes. For example, two PD risk genes (*GBA* and *GOSR2*, module 22) regulating the degradation and secretion of alpha-synuclein display Hominini-enriched expression in a subset of neuronal subclasses. Only intra-module connections are visualized for clarity. (**E**) Expression of *NR4A2* (left), *FOXP1* (middle) and *MEIS2* (right) across species overlaid onto the UMAP layouts. The subclasses showing significant evolutionary changes are marked in dashed circles. The violin plots in the bottom summarize the expression of the three genes in the corresponding subclass across species. (**F**) Expression of subclass-specific brain disorder genes

with either enriched or depleted expression specific to humans. Cell subclasses showing species-enrichment or -depletion of gene expression are color by species. Genes specific to other species are summarized in table S9. The MS risk genes *MERTK* has been reported to modulate microglia activities to regulate myelination (*124*).

Table S1. Human, chimpanzee, macaque and marmoset specimens used for snRNA-seq and sn-multiome. Metadata information of the human and nonhuman primate samples.

Table S2. Human, chimpanzee, and marmoset specimens used for Nissl staining, immunohistochemistry, and *in situ* hybridization. Metadata information of the human and nonhuman primate samples.

Table S3. Summary of FOXP2 immunohistochemistry across 51 mammals. This table summarizes the FOXP2 immunohistochemistry in the cerebral neocortex and other brain regions across 51 mammals and also detailed the samples used for this experiment.

Table S4. Peaks proximal to FOXP2 showing enrichment in FOXP2-expressing subclasses. By performing differential peak enrichment tests (Materials and Methods), we identified a list of peaks in the 1.5x zoom-out range of *FOXP2* displaying enrichment in a subset of *FOXP2*-expressing subclasses.

Table S5. CellOracle predicted FOXP2 regulatory networks. CellOracle predicted *FOXP2* regulatory networks

Table S6. Mouse samples used for the in utero electroporation experiment. Metadata information of the mouse samples analyzed in the in utero electroporation experiment.

Table S7. Differentially expressed genes between FOXP2- and GFP-electroporated mouse cortices. A summary of differentially expressed genes identified by comparing FOXP2- versus GFP-electroporated mouse brain cortices.

Table S8. List of conserved and divergent co-expression modules of brain disorder risk genes. We applied co-expression analysis to parcellate the brain disorder risk genes into different modules based on their expression pattern similarity. This table shows the module that each of these gene belongs to and its correlation to the summarized expression pattern of the given module, measured as the principal component 1 (or module eigengenes) of all gene expression data within that module.

Table S9. Subclass- and species-specific expression of brain disorder risk genes. A summary of brain disease genes showing subclass- and species-specific expression. Those genes showing human-specific expression were shown in figure S23D and were not included. A value of "0" in the table means the given gene is not enriched/depleted in the given subtype subclass while a value of "1" refers to enrichment or depletion.

Table S10. Full names of the gene acronyms mentioned in this study. Detailed descriptions of the gene symbol acronyms included in subtype names and mentioned in the main text.