Supporting Information

Kalish et al. 10.1073/pnas.1717871115

SI Experimental Procedures

Animals. Single-cell sequencing and FISH were performed on tissue from wild-type C57Bl6/J mice obtained from the Jackson Laboratory. Tg(Prkcd-glc-1/CFP,-Cre) ["Prkcd-Cre" (90)] mice were recovered from cryopreservation by the Jackson Laboratory and subsequently backcrossed to the C57Bl6/J background for at least 10 generations to achieve >97% strain purity (analysis by Charles River Laboratories). Prkcd-Cre mice were crossed with Gt(ROSA)26Sortm5(CAG-Sun1/sfGFP)Nat/J obtained from the Jackson Laboratory (29).

Generation of Single-Cell Suspensions. Mice were euthanized on postnatal days 5, 10, 16, and 21. Each time point includes data from eight mice total processed as four independent samples of two mice each. After isoflurane anesthetization, mice were decapitated and the brain was isolated. Mouse brains were dissected and 300-μm coronal sections were made on a Leica VT1000S vibratome. The dorsal LGNs were then microdissected in ice-cold PBS after visual identification using a Nikon SMZ-10A bright-field dissection microscope. LGNs from four animals were pooled to create one replicate per time point. LGN tissue was transferred to dissociation media: HBSS (Life Technologies), 10 mM Hepes (Sigma), 172 mg/L kynurenic acid (Sigma), 0.86 g/L $MgCl_2{\cdot}6H_2O$ (Sigma), and 6.3 g/L D-glucose (Sigma). This solution was saturated with 95% oxygen and 5% CO_2 and was pH-adjusted to 7.35 before use. Papain (20 U/mL; Worthington), Pronase from *Streptomyces griseus* (1 mg/mL; Sigma), proteinase XXIII from *Aspergillus melleus* (3 mg/mL; Sigma), and DNase (2 mg/mL; Worthington) were added to the dissociation media. Dissociation was carried out at 37 °C for 1 h. The samples were then triturated, filtered, and spun at 300 × *g* for 5 min. The pellet was resuspended in trypsin inhibitor solution (dissociation media with 1% BSA and 1% ovomucoid). Following gradient centrifugation (70 × *g* for 10 min), the cells were washed in dissociation media containing 0.04% BSA and resuspended in dissociation media containing 0.04% BSA and 15% OptiPrep (Sigma).

Single-Cell RNA Sequencing Using inDrops. Individual cells were captured and barcoded using the inDrops platform as previously described (30). Briefly, single-cell suspensions were fed into a microfluidic device that packaged the cells with barcoded hydrogel microspheres and reverse transcriptase/lysis reagents. After cell encapsulation, primers were photo-released by UV exposure. Two libraries of ∼3,000 cells each were collected for each sample. Indexed libraries were pooled and sequenced on a NextSeq 500 (Illumina).

Sequencing Data Processing. Transcripts were processed according to a previously published pipeline (30). Briefly, this pipeline was used to build a custom transcriptome from the Ensembl GRCm38 genome and GRCm38.84 annotation using Bowtie 1.1.1 (95), after filtering the annotation gtf file (gencode.v17.annotation.gtf filtered for feature_type: "gene"; gene_type: "protein_coding"; and gene_status: "KNOWN"). Read quality control and mapping against this transcriptome were performed. Unique molecular identifiers (UMIs) were used to link sequence reads back to individual captured molecules. All steps of the pipeline were run using default parameters unless explicitly stated.

Quality Control, Dimensionality Reduction, and Clustering of Cells. All cells were combined into a single dataset. All mitochondrially encoded genes were removed from the dataset. Cells with fewer than 500 or more than 15,000 UMI counts were excluded. Cells were then clustered using the Seurat R package (96). The data were log-normalized and scaled to 10,000 transcripts per cell. Variable genes were identified using the following parameters: x.low.cutoff: 0.0125; x.high.cutoff: 3; y.cutoff: 0.5. We limited the analysis to the top 30 principal components. Clustering resolution was set to 0.6. Clusters containing fewer than 100 cells were discarded. The expression of known marker genes was used to assign each cluster to one of the main cell types. Due to the developmental regulation of neuronal markers, a combination of *Snap25*, *Slc17a6*, and *Stmn2* was used to identify excitatory neurons. *Gad1*, *Olig1*, *Aqp4*, *Cldn5*, *Vtn*, *Cx3cr1*, and *Mrc1* were used to identify inhibitory neurons, oligodendrocytes, astrocytes, endothelial and smooth muscle cells, pericytes, microglia, and macrophages, respectively. Clusters with significant expression of two or more markers were removed, as they represented doublet clusters resulting from simultaneous capture of two or more cells in a single droplet.

Clusters from each cell type were combined and the Seurat-based clustering was repeated to characterize subtype diversity. Clustering resolution was set to 1.2 for cell subtype clustering to maximize discrimination of subtype diversity. The Seurat FindMarkers function was used to identify genetic markers of cellular subtypes. To confirm the expression of genes within specific subpopulations of cell types, we performed differential gene expression analysis between cell clusters using Monocle 2 (37). We compared each cell cluster with the entire cell-type population. The data were modeled and normalized using a negative binomial distribution. A false discovery rate (FDR) less than 0.05 and a log2-fold change in expression greater than 1 was used as the threshold of enrichment of a given gene within a cell cluster.

Identification of Developmentally Regulated Genes. To identify developmentally regulated genes in the LGN, we performed a differential gene expression analysis using Monocle 2 (37). The analysis was conducted on each cell type, comparing each consecutive pair of time points (P5 vs. P10, P10 vs. P16, P16 vs. P21). The data were modeled and normalized using a negative binomial distribution. Genes whose FDR was less than 0.05, and whose log2-fold change in expression was either greater than 1 or smaller than −1, were considered developmentally regulated. log2-fold change was calculated from the depth-normalized data after adding 0.1 to the expression of each gene.

To cluster genes in excitatory neurons based on their temporal gene expression pattern, transcript counts were weighted using zingeR to eliminate zero inflation (https://www.biorxiv.org/content/early/2017/06/30/157982). We then performed weighted linear regression and downweighted the highly zero-inflated cells from the estimation of the dispersion parameter. EdgeR (97) was used to identify genes that most strongly varied by time, and hierarchical clustering was used to group genes by expression pattern.

Coexpression Analysis. To identify genes with similar cell type-specific expression patterns, we calculated Pearson correlation coefficients for depth-normalized gene expression between genes of interest and all other expressed genes. We applied a Bonferroni correction for multiple comparisons and reported the adjusted *P* value. An adjusted *P* value <0.05 was considered statistically significant.

Gene Ontology Analysis. Enrichment of functionally related genes in the dataset was assessed by performing an unbiased gene

ontology (GO) analysis using the PANTHER classification system (98).

Single-Molecule Fluorescence in Situ Hybridization. Wild-type C57Bl6/J mice (Jackson Laboratory) were euthanized with isoflurane and their brains were rapidly dissected and embedded in OCT (Tissue-Tek) on dry ice; 20-μm-thick sections were sliced on a Leica CM1950 cryostat and stored at −80 °C until use. Single-molecule multiplexed fluorescence in situ hybridizations (FISHs) were performed using the RNAscope platform (Advanced Cell Diagnostics; ACD) according to the manufacturer's protocol for fresh-frozen sections. Commercial probes obtained from ACD detected the following genes: *Stmn2*, *Aldh1l1*, *Cx3cr1*, *Pecam*, *Olig1*, *Grin2b*, *Sparcl1*, *C1qa*, *Ndnf*, *Mbp*, *Efna5*, *Plxna1*, and *Prkcd*.

To quantify gradient expression of *EphrinA5* and *PlexinA1*, FISH was performed as above and images were obtained on an Olympus FluoView 1000 confocal laser scanning microscope. Fluorescence intensity was quantified using ImageJ software (https://imagej.nih.gov/ij/).

Axon Tracing. Choleratoxin subunit B conjugated to Alexa Fluor 647 or 555 was injected into the vitreous of either the left or right eye of Sun1/Prkcd-Cre positive or Sun1/Prkcd-Cre negative mice. Mice were killed 24 h after injection and transcardially perfused with 4% paraformaldehyde (Electron Microscopy Sciences). Fixed brains were sectioned at 50 μm thickness on a Leica VT1000S vibratome. Sections were stained with Dapi and mounted on glass microscopy slides (Fisher) for imaging.

Imaging. Fluorescent axon tracing and FISH images were taken on an Olympus FluoView 1000 confocal laser scanning microscope, equipped with the following objectives: 10× air 0.4 N.A., 20× air 0.75 N.A., 40× oil 1.3 N.A., and 60× oil 1.42 N.A.

Comparison with Allen LGN Data. To quantify the degree to which cell types were similar to the LGN scRNA-seq data from the Allen Institute for Brain Science, we applied the MetaNeighbor approach (https://www.biorxiv.org/content/early/2017/06/16/150524). Spearman correlation coefficients were calculated between all pairs of cells within a dataset. The correlations were ranked and standardized so that all values lay between 0 and 1. Only data from excitatory neurons, interneurons, and oligodendrocytes were included in the analysis. A neighbor voting algorithm was then used to predict the identity of the test set based on similarity to the training data. Our data and the Allen LGN data were alternately used as the test and training set. The ability to recover cells of the same type is reported as the mean area under the receiver operator characteristic curve. All cell types were compared across 106 gene sets from the GO Consortium goslim_generic collection (August 2015), which was filtered for terms appearing in the GO Consortium mouse annotations gene_association.mgi.gz and those containing 20 to 1,000 genes.

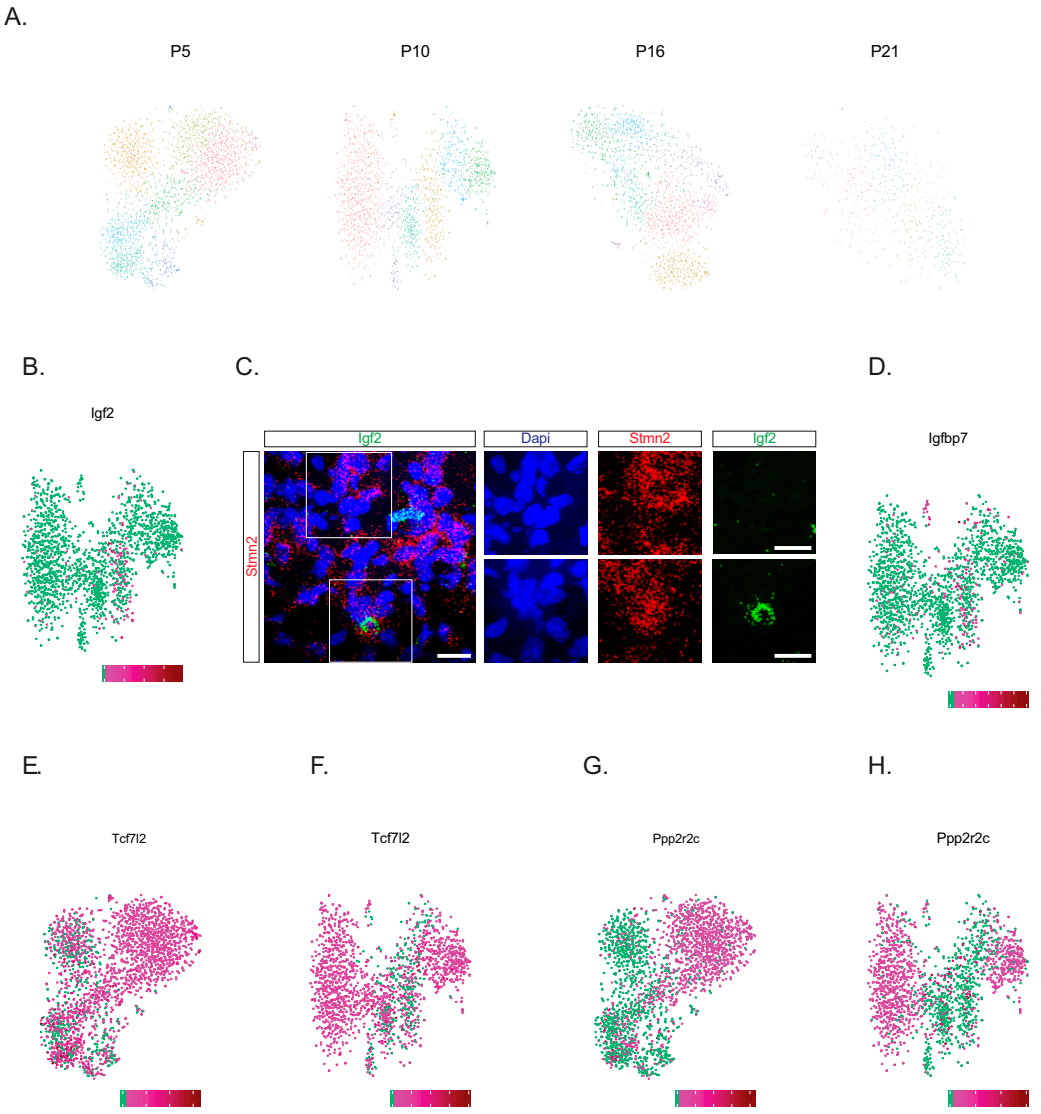

Fig. S1. Transcriptional heterogeneity of excitatory neurons. (*A*) t-SNEs of subclusters of excitatory neurons at each age. (*B*) *Igf2* expression across excitatory clusters at P10. (Scale, 0 to 120 transcripts per cell.) (*C*) Confocal images of FISH in P10 LGN sections probed for *Igf2* (green) and the excitatory neuron marker *Stmn2* (red). White boxes are magnified (*Right*). (Scale bars, 10 μm.) (*D*) *Igfbp7* expression across excitatory clusters at P10. (Scale, 0 to 46 transcripts per cell.) (*E*) *Tcf7l2* expression across excitatory clusters at P5. (Scale, 0 to 70 transcripts per cell.) (*F*) *Tcf7l2* expression across excitatory clusters at P10. (Scale, 0 to 110 transcripts per cell.) (*G*) Expression of *Ppp2r2c* across excitatory clusters at P5. (Scale, 0 to 44 transcripts per cell.) (*H*) Expression of *Ppp2r2c* across excitatory clusters at P10. (Scale, 0 to 30 transcripts per cell.)

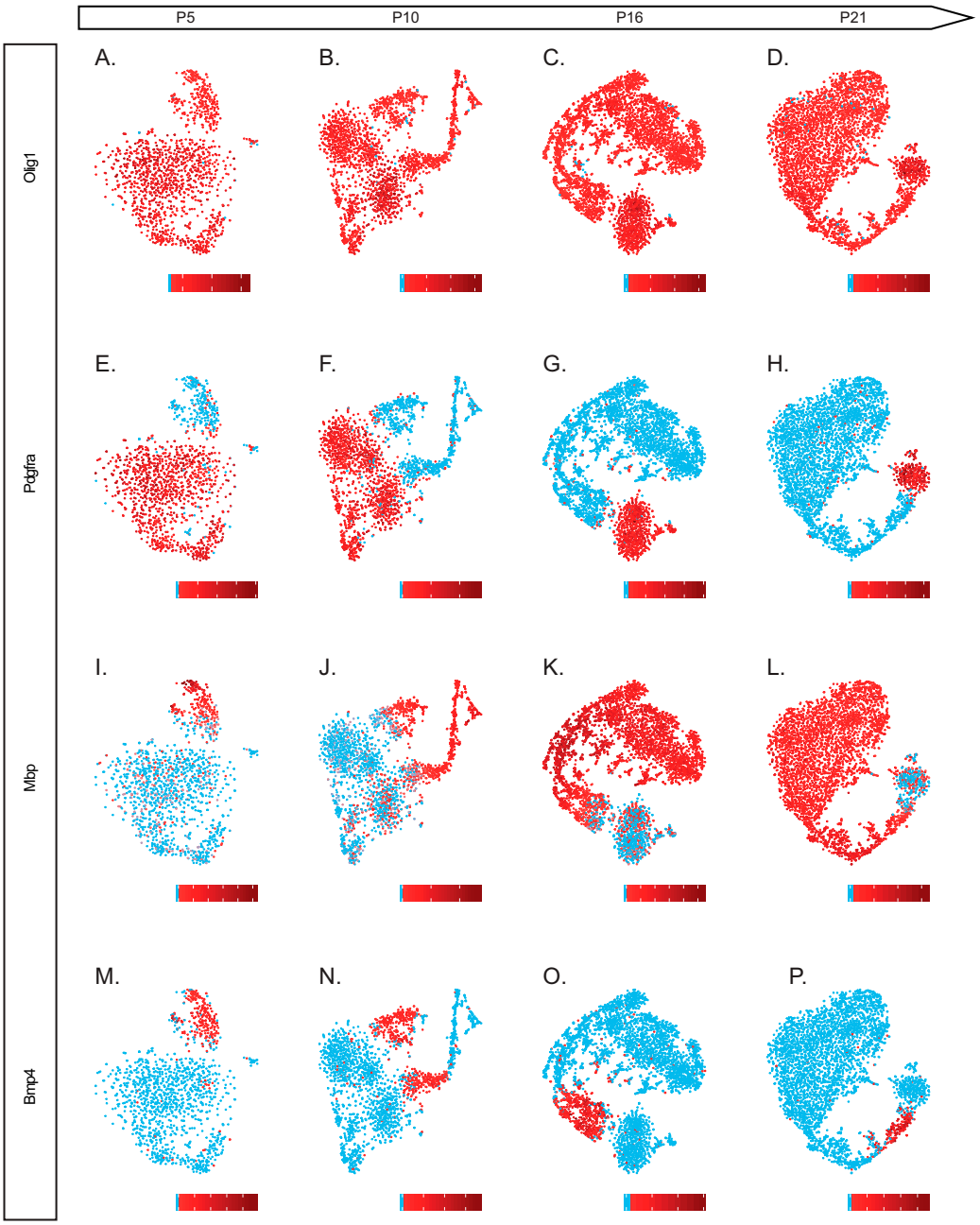

Fig. S2. Trajectory of oligodendrocyte maturation. (*A*) *Olig1* expression across oligodendrocyte clusters at P5. (Scale, 0 to 200 transcripts per cell.) (*B*) *Olig1* expression across oligodendrocyte clusters at P10. (Scale, 0 to 300 transcripts per cell.) (*C*) *Olig1* expression across oligodendrocyte clusters at P16. (Scale, 0 to 325 transcripts per cell.) (*D*) *Olig1* expression across oligodendrocyte clusters at P21. (Scale, 0 to 300 transcripts per cell.) (*E*) *Pdgfra* expression across oligodendrocyte clusters at P5. (Scale, 0 to 120 transcripts per cell.) (*F*) *Pdgfra* expression across oligodendrocyte clusters at P10. (Scale, 0 to 150 transcripts per cell.) (*G*) *Pdgfra* expression across oligodendrocyte clusters at P16. (Scale, 0 to 160 transcripts per cell.) (*H*) *Pdgfra* expression across oligodendrocyte clusters at P21. (Scale, 0 to 110 transcripts per cell.) (*I*) *Mbp* expression across oligodendrocyte clusters at P5. (Scale, 0 to 525 transcripts per cell.) (*J*) *Mbp* expression across oligodendrocyte clusters at P10. (Scale, 0 to 1,500 transcripts per cell.) (*K*) *Mbp* expression across oligodendrocyte clusters at P16. (Scale, 0 to 550 transcripts per cell.) (*L*) *Mbp* expression across oligodendrocyte clusters at P21. (Scale, 0 to 640 transcripts per cell.) (*M*) *Bmp4* expression across oligodendrocyte clusters at P5. (Scale, 0 to 50 transcripts per cell.) (*N*) *Bmp4* expression across oligodendrocyte clusters at P10. (Scale, 0 to 110 transcripts per cell.) (*O*) *Bmp4* expression across oligodendrocyte clusters at P16. (Scale, 0 to 60 transcripts per cell.) (*P*) *Bmp4* expression across oligodendrocyte clusters at P21. (Scale, 0 to 50 transcripts per cell.)

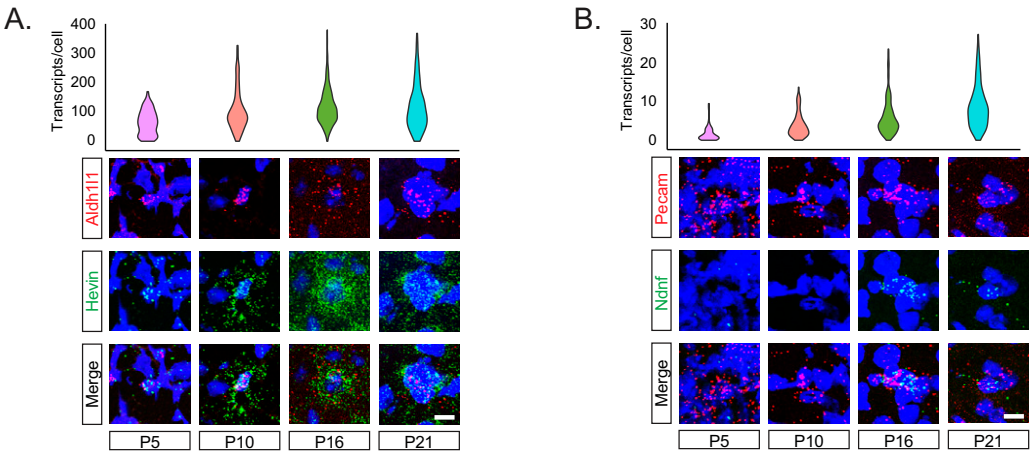

Fig. S3. Validation of cell type-enriched gene expression in astrocytes and endothelial cells. (*A*) Violin plot based on scRNA-seq data (*Top*) and confocal FISH images (*Bottom*) showing developmental regulation of *Hevin* (green) in astrocytes identified by expression of *Aldh1l1* (red). Ages are listed below. (Scale bar, 5 μm.) (*B*) Violin plot based on scRNA-seq data (*Top*) and confocal FISH images (*Bottom*) showing developmental regulation of *Ndnf* (green) in endothelial cells identified by expression of *Pecam* (red). Ages are listed below. (Scale bar, 5 μm.)

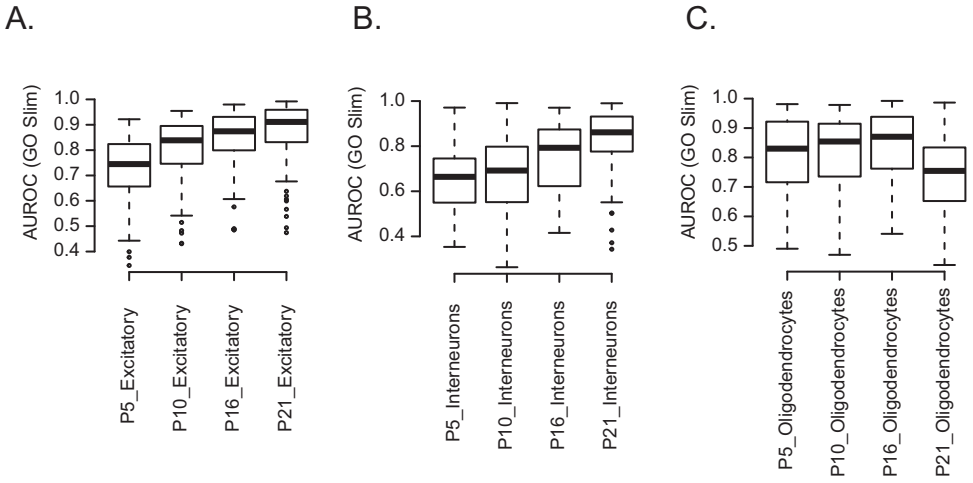

Fig. S4. Comparison of the scRNA-seq dataset with the Allen Brain Atlas dataset. Boxplots of area under the receiver operating characteristic (AUROC) scores for excitatory neurons (*A*), interneurons (*B*), and oligodendrocytes (*C*) based on the MetaNeighbor analysis between developmental LGN data and Allen Brain Atlas adult LGN data. Cell types are divided by developmental stage. Boxes represent quartiles, middle lines represent the median, and whiskers extend to 1.5 times the interquartile range.