Article

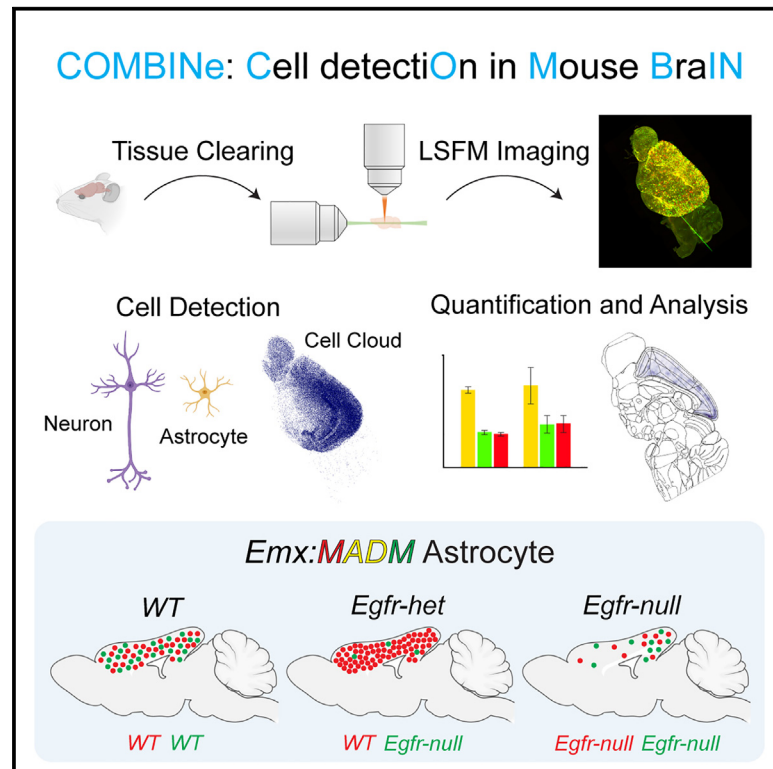# COMBINe enables automated detection and classification of neurons and astrocytes in tissue-cleared mouse brains

Graphical abstract

Highlights

- COMBINe performs automated cell detection based on morphology and color

- COMBINe accurately detects out-of-focus cells in tissue-cleared brains

- COMBINe overcomes slight misalignments between color channels

- Hierarchical analysis reveals regional effect of sparse *Egfr* deletion on astrocytes

Authors

Yuheng Cai, Xuying Zhang, Chen Li, H. Troy Ghashghaei, Alon Greenbaum

Correspondence

greenbaum@ncsu.edu

In brief

Cai et al. present a workflow called COMBINe to automatically detect and quantify all labeled cells in tissue-cleared mouse brains and apply it to study the effects of sparse epidermal growth factor receptor deletion. COMBINe is adaptable for various neuroscience projects that require whole-brain quantification and analysis.

CellPress

Article

COMBINe enables automated detection and classification of neurons and astrocytes in tissue-cleared mouse brains

Yuheng Cai,[1,2] Xuying Zhang,[3] Chen Li,[1,2] H. Troy Ghashghaei,[3] and Alon Greenbaum[1,2,4,5,*]
[1]Joint Department of Biomedical Engineering, North Carolina State University and University of North Carolina at Chapel Hill, Raleigh, NC, USA
[2]Comparative Medicine Institute, North Carolina State University, Raleigh, NC, USA
[3]Department of Molecular Biomedical Sciences, North Carolina State University, Raleigh, NC, USA
[4]Bioinformatics Research Center, North Carolina State University, Raleigh, NC, USA
[5]Lead contact
*Correspondence: greenbaum@ncsu.edu
https://doi.org/10.1016/j.crmeth.2023.100454

MOTIVATION Tissue clearing has been extensively used to explore cellular organizations in intact organs. By imaging entire mouse brains with cellular resolution, the resulting datasets contain terabytes of images to process and millions of labeled cells that need to be counted and classified based on morphology and color. Therefore, automation of data processing and analysis is an urgent need. We present a workflow (COMBINe) to automatically locate and classify labeled cells in such 3D datasets and to quantitatively analyze the regional effects after registration to the referenced Allen Brain Atlas. We applied the approach to study the effects of sparse epidermal growth factor receptor deletion on gliogenesis.

SUMMARY

Tissue clearing renders entire organs transparent to accelerate whole-tissue imaging; for example, with light-sheet fluorescence microscopy. Yet, challenges remain in analyzing the large resulting 3D datasets that consist of terabytes of images and information on millions of labeled cells. Previous work has established pipelines for automated analysis of tissue-cleared mouse brains, but the focus there was on single-color channels and/or detection of nuclear localized signals in relatively low-resolution images. Here, we present an automated workflow (COMBINe, Cell detectiOn in Mouse BraIN) to map sparsely labeled neurons and astrocytes in genetically distinct mouse forebrains using mosaic analysis with double markers (MADM). COMBINe blends modules from multiple pipelines with RetinaNet at its core. We quantitatively analyzed the regional and subregional effects of MADM-based deletion of the epidermal growth factor receptor (EGFR) on neuronal and astrocyte populations in the mouse forebrain.

INTRODUCTION

Recent advances in tissue-clearing methods have enabled three-dimensional (3D) imaging of intact mammalian organs and entire small organisms while preserving comprehensive structural and cellular information.[1–7] With light-sheet fluorescence microscopy (LSFM), high-speed imaging of cleared and immunolabeled tissues can be achieved.[8–11] Yet, challenges remain in analyzing the large resulting 3D datasets with terabytes of images and information on millions of labeled cells. Hence, the automation of data processing and cell detection is paramount for high-throughput cellular profiling of cleared tissues.

Existing pipelines for automated analysis of tissue-cleared mouse brains include BCFind,[12,13] ClearMap,[5] CUBIC,[14] and NuMorph.[15] BCFind achieves fully automated detection of neurons in 3D mouse brains by integrating a U-Net model for semantic deconvolution and a blob detection algorithm for cell identification. ClearMap presents an immunolabeling-enabled tissue clearing method with superior efficiency in data processing, accomplishing automated cell counting, dataset registration, and statistical analysis. The datasets are analyzed using traditional image processing routines (e.g., morphological operations). CUBIC incorporates improved clearing, imaging, and cell-nucleus-detection protocols, with the ability to quantify cells of entire adult mouse brains. NuMorph is an advanced end-to-end data processing tool for accurate cell-type quantification within the mouse cortex, which reaches high precision in nuclei detection. The previous pipelines achieved high efficacy, yet
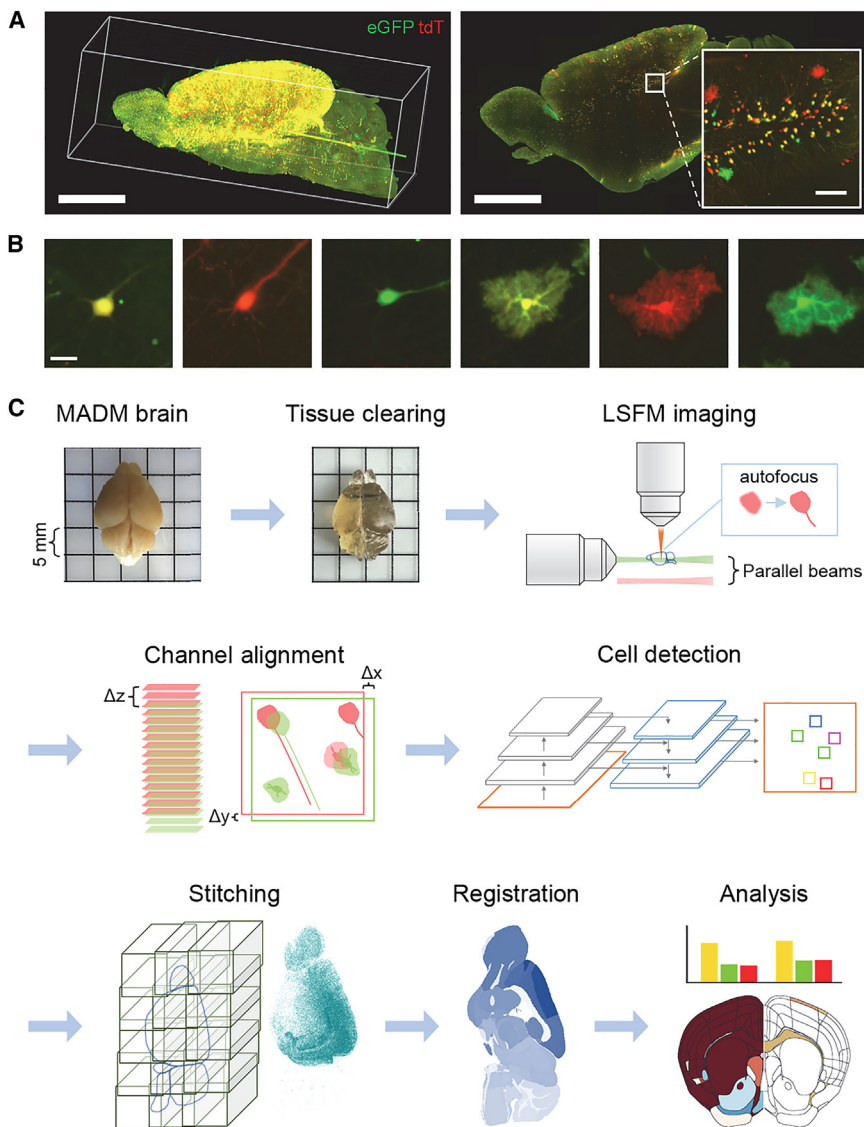
Figure 1. Automated cell detection in cleared MADM mouse forebrain

(A) Left: maximum intensity projection of fluorescence signals from a hemisphere of an *Emx: MADM:+/+* forebrain labeled with antibodies to enhance the endogenous EGFP (green) and tdTomato (tdT; red) signals (scale bar, 3 mm). Right: a virtual 30 μm thick slice from the same cleared sample (scale bar, 2.5 mm) with a zoomed-in area (scale bar, 100 μm). EGFP, enhanced green fluorescent protein; tdT, a red fluorescent protein.

(B) Magnified images of six isolated MADM-labeled neurons and astrocytes expressing EGFP (green), tdT (red), or both (yellow) MADM reporters. Scale bar, 25 μm.

(C) The proposed COMBINe workflow for MADM cell mapping (see also Figure S1).

they focused on single-color channels and/or detection of nuclear signals in relatively low-resolution images. As such, the existing pipelines fail to capture cases where (1) whole cells or cell membranes rather than nuclei are stained, (2) colocalization of color channels is crucial, and (3) the cells can be classified based on their morphology, which requires high-resolution imaging. These requirements are broadly paramount in neuroscience research, especially those that require *in situ* analysis of multiple genetic markers as in Confetti mice[16] or mosaic analysis with double markers (MADM).[17–21] Here, we addressed this critical gap by building on the strengths of existing pipelines while developing an automated workflow that was tested in MADM datasets. In these datasets, the fluorescent expression in neurons and glia can indicate their distinct genotypes.

In the core of our workflow is deep learning, a data-driven approach. Recent adaptation of deep learning[22] in biomedical studies has enabled automated and accelerated image processing.[23–25] In combination with tissue clearing and 3D microscopy, deep learning has been employed in investigation of cancer metastasis in mice,[26] segmentation of mouse brain vasculature,[27] and detection of crown-like structures in adipose tissues.[28] Inspired by these studies, we surveyed the deep learning literature and employed a model (RetinaNet) that shows superior performance in dense object detection and applied it to a neurodevelopmental problem.[29–31]

Using RetinaNet as its core, we developed an automated workflow we refer to as COMBINe to map labeled neurons and astrocytes in the forebrains of three genetically distinct MADM mice (Figure 1). We used a conditional mouse allele for the *Epidermal growth factor receptor* (*Egfr*) in combination with MADM alleles. Using the established *Emx1*[cre] line to restrict recombination to the dorsal telencephalon (cerebral cortices and the hippocampal formation), three genotypes were obtained: *Emx:MADM:+/+*, *Emx:MADM:F/+, and Emx:-MADM:F/F*.[32] We have recently reported on variable effects on specification and differentiation of MADM glia in cerebral cortices of these mice, which has led to the developmental model that EGFR-dependent and EGFR-independent perinatal gliogenesis regulate forebrain development cooperatively.[32] Here, we used the *Egfr* MADM model to test COMBINe. Using tissue clearing and a custom-built LSFM that changes its properties based on the sample, we generated high-quality MADM datasets with cellular resolutions sufficient to morphologically differentiate astrocytes from neurons. Employing RetinaNet, 0.5 million MADM cells across six classes (yellow, red, green: neurons and astrocytes) were located and counted from each MADM forebrain hemisphere. Taking advantage of available packages, stitched volumes and detected cells were aligned to
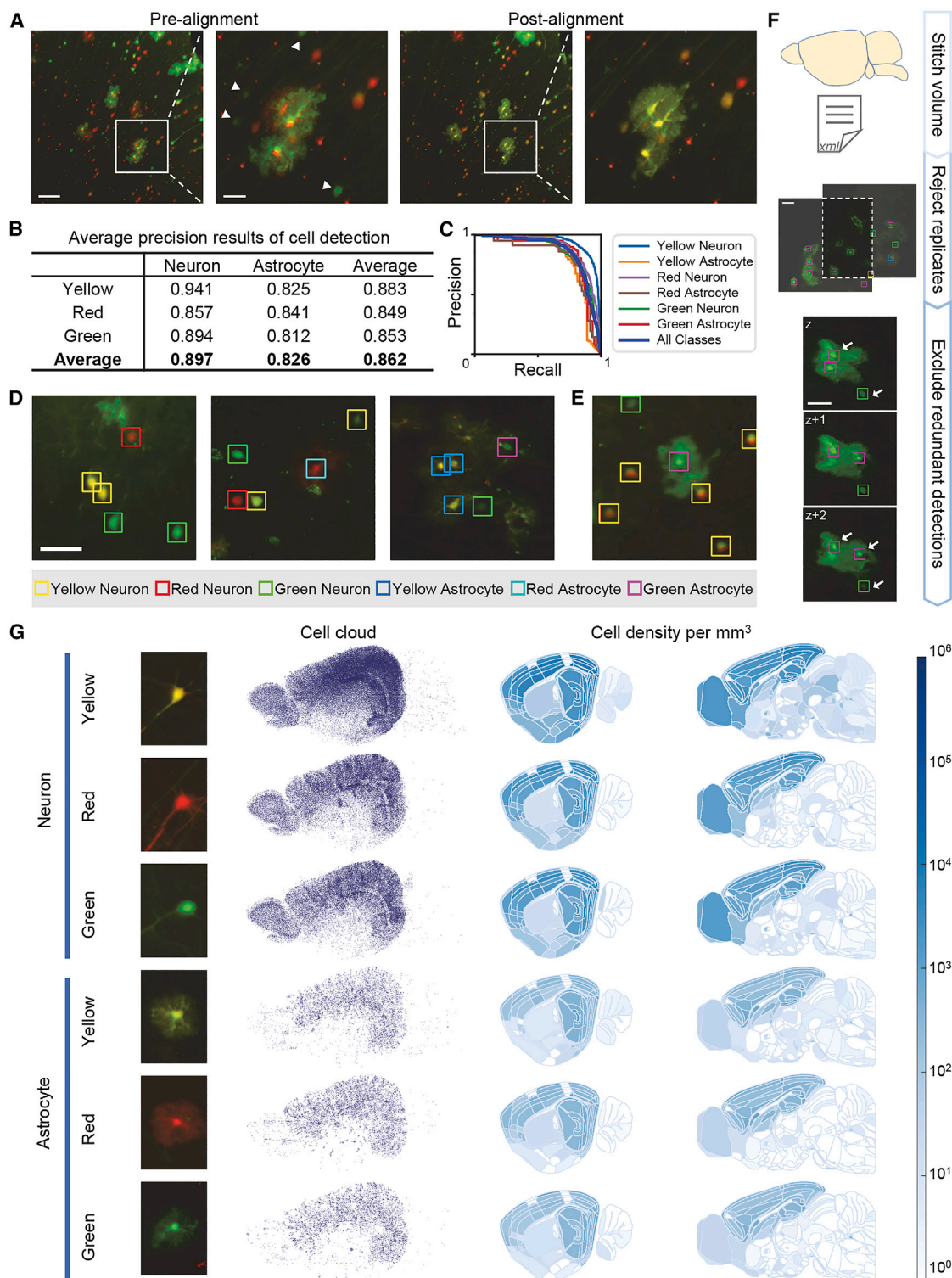
Figure 2. Automated detection, stitching, and registration of MADM cells

(A) Images pre- and post-channel alignment from parallel illuminations (left panel: scale bar, 100 μm). Cells in the wrong z plane (arrowheads) disappear following alignment. Boxed areas are zoomed-in images (scale bar, 30 μm).

(B) An overall average precision of 0.86 was achieved across the six cell classes using RetinaNet.

(C) Precision-recall curves.

(legend continued on next page)

annotations provided by the Allen Brain Atlas (ABA) for regional registration and analysis.[33] Significant variations in cell populations corresponding to the MADM genotypes were confirmed, matching our findings from laborious manual counting and registrations.

RESULTS

Automated cell detection in cleared MADM mouse forebrain

We combined the *Emx1*[cre] with MADM alleles to map neuronal and glial lineages in the forebrain as before.[17,18,20,34] MADM allows simultaneous labeling and genetic manipulation in clones of somatic cells in isolated and sparse populations of progenitors[35] and has been successfully employed in genotype-phenotype studies on neurogenesis and gliogenesis.[21,36] MADM-labeled neurons and glia can possess distinct genotypes, which was tracked by permanent labeling of two mitotically derived daughter cells with two distinct and nonoverlapping fluorescent markers: enhanced green fluorescent protein (EGFP) and/or tdTomato (tdT; Figure 1A). MADM cells in forebrain samples can be counted and classified into six groups based on morphology and color: red (tdT), green (EGFP), and yellow (both reporters coexpressed) neurons and glia (Figure 1B).

The notation *Emx:MADM:+/+* represents wild-type (WT) mice, whereas *Emx:MADM:F/F* indicates mice with complete deletion of *Egfr* in the dorsal forebrain (cerebral cortices and hippocampal formation). Both groups of mice have sparse MADM-labeled populations of cells in the dorsal forebrain with identical genotypes (i.e., *WT* or *Egfr-null*, respectively). The *Emx:MADM:F/+* mice are unique in that the vast majority of cells in the dorsal forebrain are heterozygous for *Egfr* deletion, but spare populations of MADM neurons and glia are genetically guided to have three distinct genotypes reported by red, green, or both (yellow) MADM colors: *WT*, red MADM; homozygous *Egfr-null*, green MADM; and heterozygous *Egfr-null*, yellow MADM (for details on the genetics of these mice, see Zhang et al.[32]). In other words, the red MADM cells in the dorsal *Emx:MADM:F/+* forebrain express EGFR in excess to their bulk environment, and the green MADM cells lack EGFR expression. The yellow MADM cells are genotypically identical to the majority non-MADM background cells, which are all heterozygous for the *Egfr-null* allele. Therefore, a single mouse (*Emx:MADM:F/+*) provided for both overexpression and knockout experimental conditions in the same tissue, which is an advantage of MADM genetics for studying gene-microenvironment interactions.

Increased understanding of the complexities in MADM data in the mouse forebrain inspired the need for classification of MADM-labeled cells based on morphology and color. Therefore, we designed a workflow that spanned tissue clearing to quantitative analysis (Figure 1C). *Emx:MADM* mouse brains were split into hemispheres, cleared using the iDISCO+ protocol,[5] and imaged by a custom-built LSFM.[37–39] The LSFM was equipped with an autofocus algorithm to correct the shift between the light-sheet illumination beam and the detection focal plane, and therefore our images were crisp and focused throughout acquisition. Before imaging, illumination beams of different wavelengths were aligned parallel to ensure accurate data processing (see later in Results). Through imaging with subcellular resolution, the acquired two-channel datasets consisted of ~400 GB data per hemisphere. Channel alignment was achieved using NuMorph,[15] which locally evaluated the translation between the color channels, followed by a RetinaNet model trained to detect six classes of labeled cells simultaneously.[29,30] The aligned tiles and detected cells were then stitched according to TeraStitcher[40] and registered to the ABA by ClearMap[5] for regional analysis in the forebrain. By mixing and matching codes from different packages, we were able to build on their relative advantages (Figure S1). Taking advantage of the above-mentioned tools, we were able to analyze a single *Emx:MADM* hemisphere in 4 days using a regular desktop computer and employing the proposed pipeline. For completeness, glia labeled using *Emx:MADM* forebrains consisted of oligodendrocytes and astrocytes. However, the morphology and size of mature astrocytes were highly distinguishable, whereas oligodendrocytes were small and difficult to detect, and their numbers were small compared with neurons and astrocytes.[32] Thus, we trained our platform to distinguish between neurons and astrocytes.

COMBINe is highly suitable for automated detection, stitching, and registration of MADM cells in the P30 cleared forebrain

RetinaNet was adapted to detect six classes of MADM cells (yellow, red, and green: neurons and astrocytes). RetinaNet is a state-of-the-art object detection network that includes a backbone of a feature pyramid network on top of a ResNet to generate a multiscale feature pyramid, with two subnetworks to regress and classify anchor boxes.[29] The output of RetinaNet is a list of bounding box coordinates, classification labels, and confidence scores for the detected objects. We selected the RetinaNet model because of its superior performance in dense object detection of real-world images[41] and in cell detection of fluorescence microscopy images.[31]

Since local translation was sufficient to register the color channels under parallel beam condition, cell detection was performed after color registration (Figure 2A; see next section). Cell detection was implemented in 2D instead of 3D due to the difficulty of labeling and training, as well as the variance in resolution along different axes in 3D images. We explored the training regime of the network and applied it to fluorescence images of MADM brain sections in a pervious study.[30] We again employed the data augmentation of color swap and saturation to balance the

(D and E) Representative images in which the program made predictions with accurate (D) and inaccurate (E) color registrations (scale bar, 50 μm). See also Figure S2.

(F) Stitching and merging of detected cells in 3D for each tile were according to TeraStitcher parameters. Replicates in overlapped regions across tiles were rejected, and redundant detections across different z planes were excluded according to signal intensities. Scale bars, 50 μm.

(G) Detected cell clouds and corresponding regional densities after registration to the Allen Brain Atlas (ABA) in a single hemisphere of an *Emx:MADM:+/+* forebrain.

color classes and utilized transfer learning to train the network on a pre-trained backbone. After training, the model that performed best on validation data was selected as the inference model. The inference model reached an average precision of 0.862 across six classes on unseen test data (Figure 2B). An average precision of 0.897 was achieved for detection of neurons, while an average precision of 0.826 was achieved for detection of astrocytes. Given a confidence threshold of 0.5, the mean values of precision, recall, and F-score across six classes on test data were 0.848, 0.771, and 0.806 respectively (Tables S1–S3). The adjusted precision-recall curves used to measure average precision across six classes are shown in Figure 2C, with several examples of the model predications shown in Figure 2D. Color classification was robust even in cases with minor misalignment between the two-color channels (Figure 2E). Examples of manual annotation versus automated prediction on a field of view are shown in Figure S2.

After detecting the cells in raw tiles, the displacement parameters between individual tiles were estimated using TeraStitcher,[40] but the high-resolution volume was not stitched. A custom script using the displacement parameters was created for excluding duplicate detections in overlapping regions between adjacent tiles and for removing redundant detections across different z planes (Figure 2F; see STAR Methods). Analysis of a cleared MADM hemisphere was accomplished in 4 days using a standard computer, with generated datasets containing approximately 100,000 images per sample (more than 1.6 million image patches for cell detection). More than 400,000 MADM cells were detected in each cleared hemisphere. A low-resolution stitched volume of the hemisphere was later registered to brain regions annotated by ABA using ClearMap. Detected MADM cells and corresponding regional density maps of a representative *Emx:MADM:+/+* sample are visualized in Figure 2G using a custom script. Nearly all *Emx:MADM:+/+* cells were in the dorsal forebrain consisting of the olfactory bulb, the cortex, and the hippocampus, which was consistent with previous observation on 2D slices.[19] As a quality check for our color registration process, the yellow, red, and green neuron portions of the total population were approximately 50%, 25%, and 25% respectively, which was in line with previous estimates.[35]

In summary, COMBINe can detect and classify cells in an entire MADM mouse brain hemisphere. The raw datasets were acquired using our custom adaptive LSFM, and since commercial LSFMs are not adaptive and do not change their properties on the fly, we have tested the effect of two realistic acquisition scenarios on the performance of the RetinaNet model in the next section (1) when the cells are slightly out of focus and (2) when the illumination beams (one per color) are not perfectly parallel.

RetinaNet accurately detects out-of-focus cells

LSFM images can seem out of focus when the objective focal plane does not perfectly overlap with the illumination beam.[42] Therefore, we tested the performance of our network in images containing blurry cells using the same tiles that were acquired in and out of focus. To our surprise, the RetinaNet model performed well in detection and color classification of cells, although degradation in image quality was observed in the out-

of-focus images. This finding illustrated the robustness of the RetinaNet model and its ability to generalize well in less-than-optimal conditions (Figure 3A). We synthetically defocused the test data according to Yang et al.[43] to compare the model performance with and without defocus and found negligible decrease in performance within 5 μm from the focal plane (Figure S3). We chose to defocus our images synthetically because (1) it allows for direct evaluation of model performance in blurry images; (2) labeling real cells in out-of-focus LSFM datasets is ambiguous, as some cells may seem sharp but very dim, and their accurate z position is difficult to derive; and (3) it will enable future use of this approach to augment existing datasets. It is important to note that while training the RetinaNet model, cells from the edges of the field of view that exhibited lower contrast compared with the center of the field of view were included. This phenomenon of fluctuations in image quality as a function of position within the field of view was an inherent property of using a Gaussian illumination beam. Consequently, the network was trained with variability in image contrast, which may explain the robustness of the network in correctly detecting out-of-focus cells (Figure 3A, parallel beam scenario).

We next tested our pipeline when the illumination beams of different color channels propagated through slightly different light paths during LSFM imaging. This can occur when the microscope setup is out of alignment due to angular drift of the scanning galvo in relation to the room temperature, the shift of the laser beam after long period of imaging (e.g., air bubbles), and more. This situation resulted in only partial overlap of the color channels, i.e., misalignments in some regions even after registration of the color channels using NuMorph. Importantly, misalignment of the color channels influenced the subsequent quantitative analyses (see pie charts in Figure 3B). In WT MADM mice, we expected a color ratio of 50% yellow, 25% red, and 25% green neurons, and we have used this distribution to test the performance of our color registration. This approximate ratio was achieved when the illumination beams were parallel to each other, and the different color channels were registered using local x-y-z translation. However, with slightly tilted illumination beams, this ratio was greatly distorted. We observed fewer yellow cells and more green cells.

Together, these findings suggest that the pipeline and the RetinaNet model were robust and highly effective for detection and color classification, even in cases containing unfocused images or slight misalignments between the color channels, as long as the light-sheet illumination beams were parallel to each other.

COMBINe faithfully captures and detects elevation of red (*WT*) astrocytes and depletion of green (*Egfr-null*) astrocytes in the *Emx:MADM:F/+* forebrain

To test the capacity and accuracy of COMBINe, we applied it to known phenotypes that we had already captured in *Emx:MADM:F/+* forebrains, where neuronal production is largely preserved, while gliogenesis is disrupted during perinatal development.[32] As described earlier, red and green cells in *Emx:MADM:F/+* forebrain correspond to *WT* and *Egfr-null* genotypes, respectively. COMBINe confirmed this effect in that both red neurons and green neurons were present and relatively similar in density, while a significant increase in *WT* red MADM
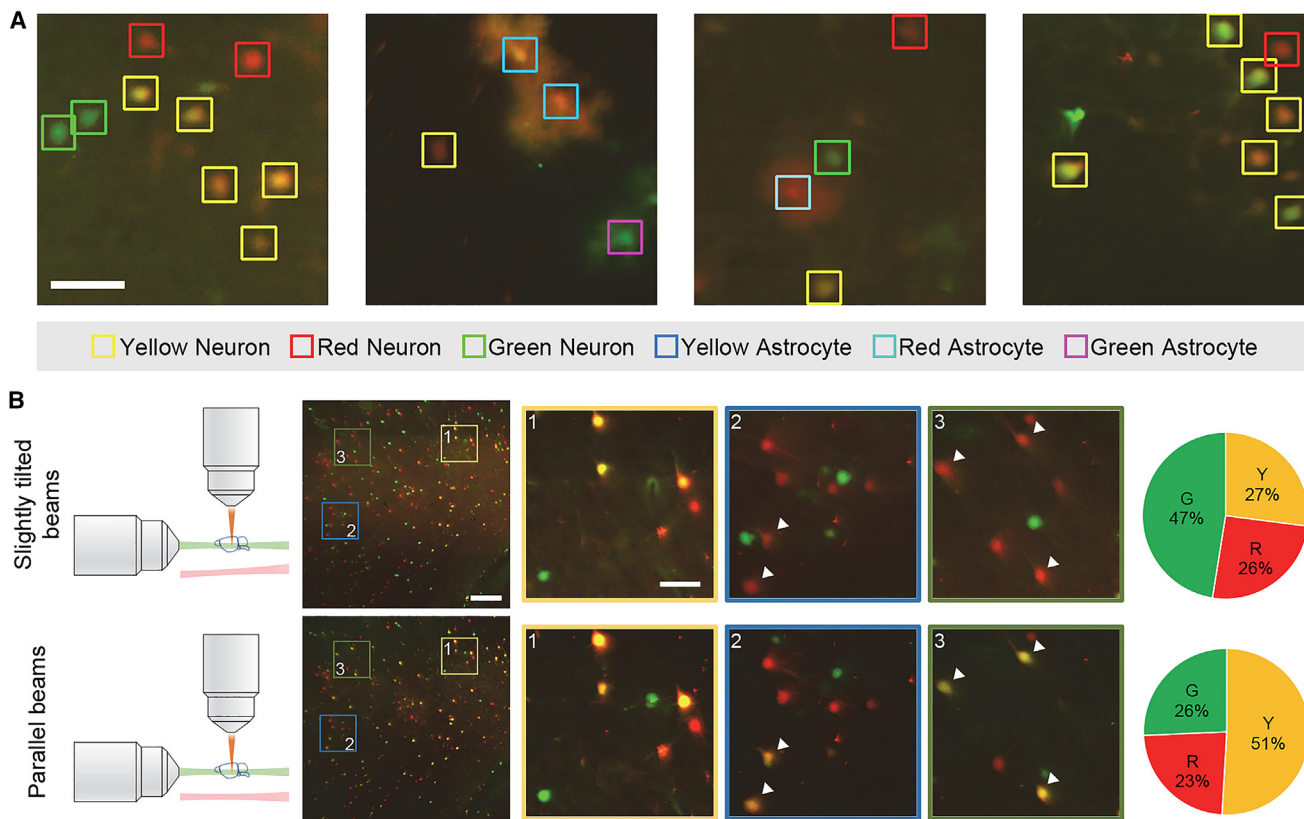
Figure 3. RetinaNet detection is robust for out-of-focus cells

(A) Representative predictions in images of out-of-focus cells (scale bar, 50 μm). See also Figure S3.

(B) Illustrations of LSFM geometry for tilted (top) and parallel (bottom) illuminations. Images from the same field of view (scale bar, 200 μm) are shown with three zoomed-in areas (scale bar, 25 μm). Unparallel beams during imaging result in misalignment between the two channels (e.g., the two-color channels are appropriately overlapped in zoomed region 1 but misaligned in regions 2 and 3 as marked by arrowheads). Erroneous color classification ensues, resulting in fewer yellow cells detected (pie charts).

astrocytes and a severe absence of green *Egfr-null* MADM astrocytes were observable throughout the labeled regions of the forebrain (Figures 4A and 4B). In addition, no significant differences in total volume of the hemisphere or regional volumes were observed, suggesting that heterozygous bulk deletion of *Egfr* or the sparse MADM phenotypes do not cause changes in overall forebrain architecture. Figure 4C shows further region-wise comparison of red and green astrocyte densities between *Emx:MADM:+/+* and *Emx:MADM:F/+* brain hemispheres. Visualized in the left columns are percentage changes of regional cell densities, while adjusted p values obtained from unpaired t tests are displayed on the right (see also Figure S4). Among 1,326 ABA-annotated regions, 31 regions showed significant elevation of *WT* red astrocytes in the *Emx:MADM:F/+* forebrain (Table S4).

Using hierarchical analysis to discover major trends in regional densities of astrocytes

While statistical analysis of all ABA annotated forebrain regions (~1,000) was useful, the average size of each brain region was relatively small. As such, the accuracy of registration to the ABA reference atlas is crucial to identify significantly altered densities of cell types wherein small errors in the registration process

can potentially mask significant differences. Therefore, we sought to merge regions based on anatomical similarities using the hierarchical structure of annotation in the ABA and implementing hierarchical analysis of region-wise statistics (Figures 5A and 5E; e.g., three-level hierarchical annotation of the somatomotor areas). Comparison of forebrain regions at multiple hierarchies revealed significant differences in red MADM astrocyte densities (Figure 5B; see also Table S5 for regions showing significant differences). Among the high-hierarchy regions, the corpus callosum showed significant increase of red astrocyte density in *Emx:MADM:F/+* brain hemispheres compared with the control group (Figures 5C and 5D), which was consistent with manual observation of the acquired images.

Variance in astrocyte distribution between rostral and caudal *Emx:MADM:F/F* cortices is quantitatively captured using COMBINe

Analysis of neurons and astrocytes in the *Emx:MADM:F/F* forebrain hemispheres using COMBINe confirmed a phenotype analyzed in our recent publication in rostral-caudal differences in EGFR-dependent and EGFR-independent astrocyte production.[32] It is important to be reminded that all cells in the
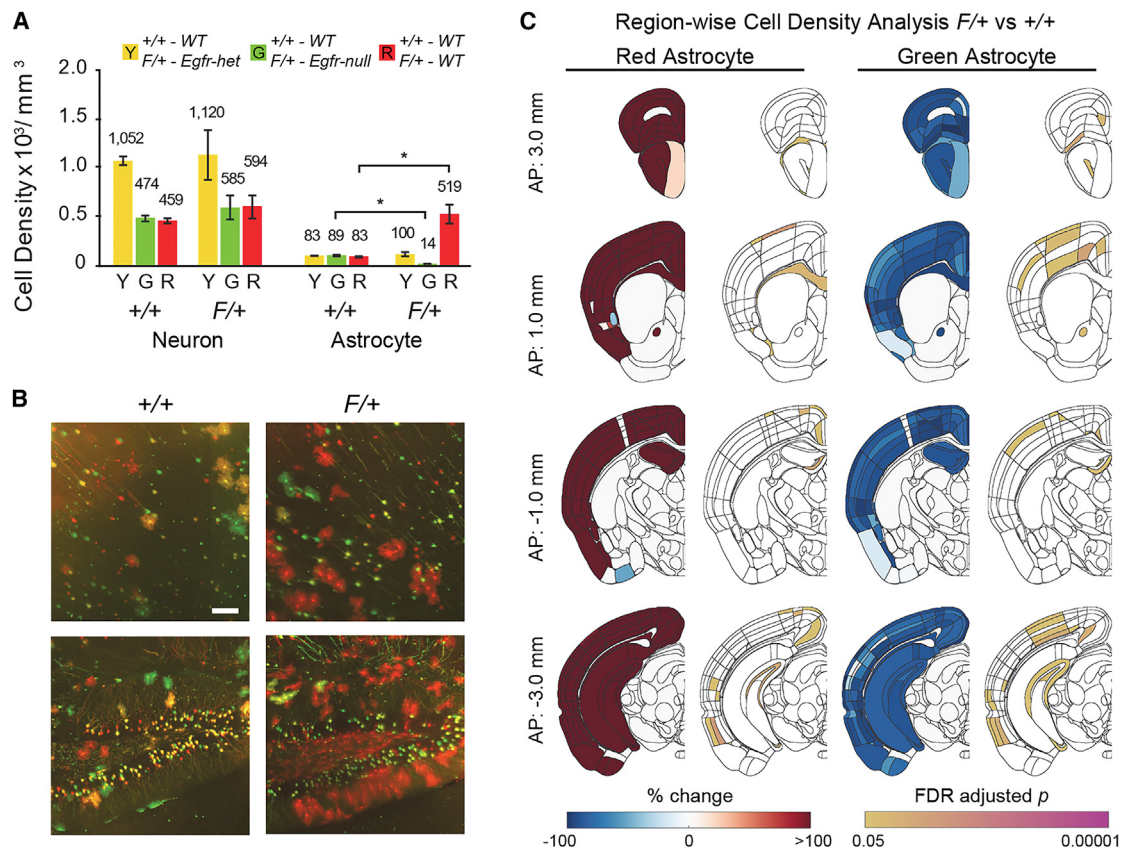
Figure 4. Elevated density of red MADM astrocytes and absence of green MADM astrocytes are faithfully captured by COMBINe in *F*/+ cortices

(A) Cell density comparison between *Emx:MADM:+/+* and *Emx:MADM:F/+* mouse brain hemispheres (mean ± SD, number of animals = 3; *, p < 0.05). In the *Emx:MADM:F/+* forebrain, red and green cells correspond to *Egfr-null* and *wild-type (WT)* genotypes, respectively. While both red and green neurons were present, a substantial increase in *WT* red astrocytes was revealed by COMBINe.

(B) Representative images of the cortex and the hippocampus. Scale bar, 100 μm.

(C) Region-wise analysis of average red and green astrocyte densities between *Emx:MADM:+/+* and *Emx:MADM:F/+* datasets. Left: percentage change in average cell densities of *Emx:MADM:F/+* brain hemispheres compared with *Emx:MADM:+/+* brain hemispheres. Right: adjusted p values (number of animals = 3). See also Figure S4 for analysis of other cell types. AP, anteroposterior distance taken from the bregma; FDR, false discovery rate.

Emx:MADM:F/F forebrain are *Egfr-null* regardless of their MADM color. Figure 6A shows the average region-wise cell density maps onto the *Emx:MADM:F/F* isocortex as defined by the ABA. Hence, we compared the average astrocyte densities of selected regions and found that there was a significant difference in astrocyte distribution between rostral and caudal regions (Figure 6B). While astrocyte densities of control *Emx:MADM:+/+* and *Emx:MADM:F/F* brain hemispheres appeared similar in caudal regions, the rostral regions of *Emx:MADM:F/F* cortices contained few, if any, astrocytes (Figure 6C). As expected, neurons that, in general, do not express EGFR showed similar distributions in the *Emx:MADM:F/F* and *Emx:MADM:+/+* (Figure S5).

To summarize, we automatically calculated and compared the densities of sparse neurons and astrocytes in distinct *Emx1[cre]*-induced MADM brain hemispheres with and without conditional deletion of *Egfr*. Region-wise analysis after registering the samples to ABA guided us to devise a hierarchical analysis method to characterize the variances observed in different cell populations among distinct genotypes. The revealed patterns matched labo-rious manual counting results, which validated the COMBINe approach with high confidence.

DISCUSSION

In this article, we present COMBINe, an automated framework that can be applied to 3D imaging datasets that are captured with multiple color channels. To demonstrate the applicability of COMBINe, we imaged tissue-cleared MADM-labeled mouse forebrains with cellular resolution. We have tested the influence of out-of-focus artifacts on the cell classification accuracy and show that if the illumination beams are parallel, the RetinaNet per-forms adequately. COMBINe uses minimal storage space, as the datasets are not repeatedly saved per pipeline stage (e.g., after stitching and registration with the ABA). We accomplish that by applying a bookkeeping strategy that uses the calculated trans-formation matrices in the stitching and registration parts to move the detected cells' coordinates from one representation to the other. Consequently, only the raw tiles and a low-resolution
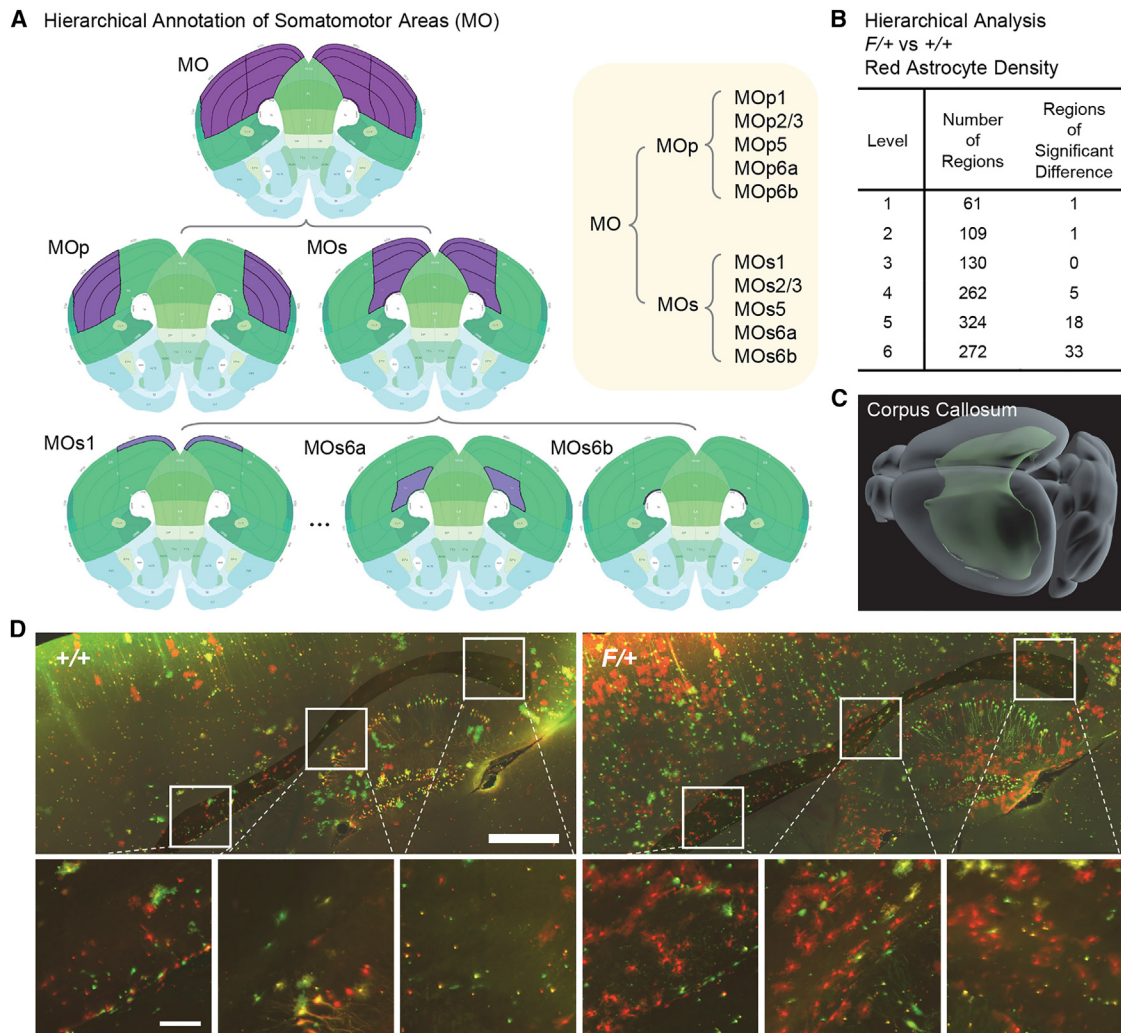
Figure 5. Hierarchical analysis reveals elevation in red MADM astrocytes in the *Emx:MADM:F*/+ corpus callosum
(A) Example of hierarchical annotations of somatomotor (MO) cortical areas according to the ABA. Mop, primary motor area; MOs, secondary motor area.
(B) Results of hierarchical analysis revealed that the corpus callosum contains significantly elevated densities of red MADM astrocytes at hierarchy level 1.
(C) 3D structure of the corpus callosum provided by ABA.
(D) Representative images of the corpus callosum (scale bar, 500 μm) and zoomed-in regions (scale bar, 100 μm).

stitched volume are required to complete the analysis. We also introduce a complementary method to perform the statistical analysis of our results by comparing the cell densities on the hierarchical structure of the brain, provided by the ABA annotation. This method allows us to average the cell densities across larger brain areas that are morphologically and functionally connected and to minimize the penalty for multiple comparisons. COMBINe can be adapted for other neuroscience studies that require quantitation of elements in the mouse brain using tissue clearing and immunofluorescence labeling. By enriching the framework components (e.g., adding features of segmentation and more), COMBINe can be applied to diverse neuroscience questions in need of analyzing 3D mouse brain datasets.

To test the RetinaNet model detection and classification success, we have used the average precision measure, which is a widely used evaluation metric for object detection tasks.[44] The

inference model (specific RetinaNet model) that is employed in the workflow achieved an average precision of 0.86 across six classes, showing its efficacy in cell detection and classification. However, when applied to large datasets, sample-wise quality control and common sense is necessary when utilizing learning-based methods. For example, in this study, the color ratio of neurons is referred to as an internal control to assess the cell detection process. Additionally, in the control brains, assessment that the color ratio is maintained across all tiles is used to assess outlier tiles.

Although we report an average precision of only 86%, such an evaluation is performed based on human-annotated ground truth. Since there is a considerable variability even between human annotators, there are no perfect models for automated analysis. Hence, it is crucial to keep the analysis process consistent and eliminate potential biases.
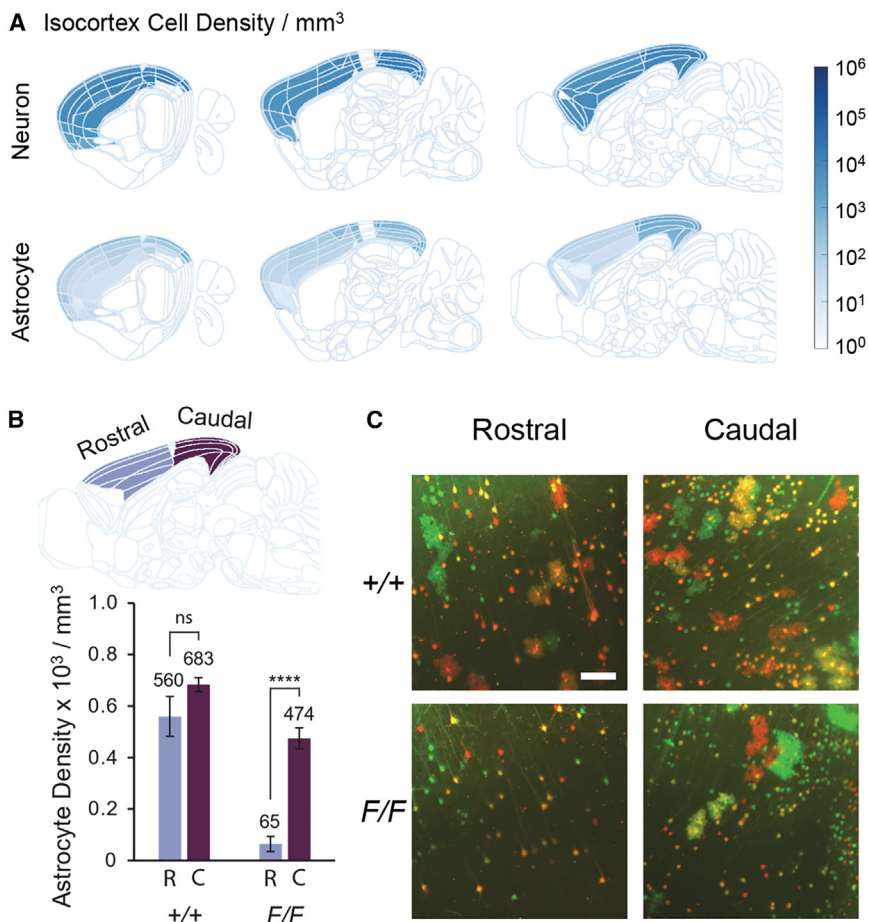
A Isocortex Cell Density / mm³

B

C

Figure 6. Variance in astrocyte distribution between rostral and caudal cortices is quantitatively captured using COMBINe

(A) Average regional cell density maps of *Emx:-MADM:F/F* brain hemispheres (number of animals = 3).

(B) Significant difference in MADM-labeled astrocyte densities was observed between rostral and caudal regions in *Emx:MADM:F/F* brain hemispheres (bar chart: mean ± SD, number of animals = 3; ****, p < 0.0001). See also Figure S5 for analysis of neuron densities.

(C) Representative images of rostral and caudal areas (scale bar, 100 μm). Almost no MADM-labeled astrocytes are observable in rostral areas of *Emx:MADM:F/F* brain hemispheres.

Most of the data processing procedures in the current study were performed in 2D given the large size of the datasets and the differences in resolution along axes. Compared with 3D processing, 2D processing has been intensely investigated and is applicable to diverse image data. Also, labeling and manipulation of data in 2D are easier and more intuitive than in 3D. However, there are situations where 3D processing is favored. For example, when segmenting and tracing nerves, the structural information is mostly maintained in 3D. To summarize, processing in 2D or 3D should be carefully chosen given the specific needs in the analysis.

All in all, performance of COMBINe in analysis of *Emx1*[cre] MADM forebrains is in line with results obtained by manual counting in several ways: (1) restricted distribution of MADM cells, (2) quantitative differences between neurons and astrocytes, and (3) the distribution of different cell colors across the brain. We did not observe volumetric changes when the experimental groups were compared with the control, confirming that sparse alterations using MADM, which are estimated to occur in 1:1,000 cells in the brain,[34,35] have undetectable effects on overall tissue development and homeostasis.

From the biological perspective, COMBINe quantitatively and unbiasedly confirms two important phenotypes in the three *Egfr* genetic backgrounds described here. The findings from *Emx:MADM:F/+* forebrains suggest that when faced with a largely *Egfr-heterozygous* forebrain background, sparse populations of red *WT* MADM cells may overexpand to generate a larger number of astrocytes than their normal capacity during perinatal development when gliogenesis occurs. In contrast, *Egfr-null* MADM cells fail to generate astrocytes in the rostral part of the isocortex. We found similar results in comprehensive study of the role of EGFR in forebrain gliogenesis[32] and in a clonal study using tamoxifen-inducible *Nes-creER:MADM: F/+* alleles to label gliogenic progeny at clonal densities.[21] Since the *WT* glial progenitors are essentially overexpressing EGFR relative to their surrounding cells in the *Emx:MADM:F/+* background, these results constitute a gain-of-function scenario that explains the strong effect on astrocyte production. An unexpected result in our current study was the region-restricted presence of glia in *Emx:MADM:F/F* forebrains since the *F/+* findings suggested that no astrocytes should be produced in the *homozygous Egfr-null* background. However, the astrocytes in this case were largely in the caudal, but not rostral, forebrain regions. This suggests the possibility of an EGFR-independent gliogenic progenitor pool in the caudal aspect of the forebrain appears capable of compensating for the loss of EGFR-dependent glia, which are critical for astrocyte production in rostral forebrain regions.[32] Interestingly, the EGFR-dependent astrocytes appear to occupy neocortical and possibly other neoforebrain areas, whereas the EGFR-independent populations that are present in the *Emx:MADM:F/F* appear to seed evolutionary older paleocortical and hippocampal areas. The precise identity and anatomical domain of the EGFR-independent glia-producing progenitors remain to be determined.

Limitation of study

Currently, deep learning is still considered a black box since the relationships between input data and the neural networks that "learn" patterns are not clearly explainable, which makes it

difficult to evaluate the errors and failures and improve the model performance.[45] Moreover, as deep learning is a data-driven method, additional training is required before applying it to different sample types. Furthermore, the packages utilized/modified in this study were built on different operating systems (e.g., Linux and Windows), and modifying them to run in a different operating system is challenging and inconvenient. In addition, performance of the brain registration approach adopted in this pipeline should be in line with previous publication that tested ClearMap/Elastix relative to other registration approaches. However, one should be aware that the accuracy of the registration to the ABA drops if the sample phenotype includes anatomical defects. This issue highlights the need for the presented hierarchical approach to compare regional changes in the brain but also the continued need for validation by the human expert. Last, some of the steps presented in COMBINe require manual checks to ensure proper functioning. For example, it is highly recommended to manually inspect a few images in each tile after channel alignment to ensure accuracy, and the cell detection model needs to be validated every time when analyzing a new batch of datasets.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Tissue clearing
  - Light-sheet imaging
  - Channel alignment
  - Cell detection
  - Stitching
  - Registration
  - Visualization
- QUANTIFICATION AND STATISTICAL ANALYSIS

AUTHOR CONTRIBUTIONS

Conceptualization, Y.C., A.G., and H.T.G.; methodology, Y.C. and C.L.; investigation, Y.C.; software, Y.C.; visualization, Y.C.; resources, X.Z. and H.T.G.; supervision, A.G. and H.T.G.; writing – original draft, Y.C.; writing – review & editing, Y.C., A.G., and H.T.G.

REFERENCES

1. Cai, R., Pan, C., Ghasemigharagoz, A., Todorov, M.I., Förstera, B., Zhao, S., Bhatia, H.S., Parra-Damas, A., Mrowka, L., Theodorou, D., et al. (2019). Panoptic imaging of transparent mice reveals whole-body neuronal projections and skull–meninges connections. Nat. Neurosci. 22, 317–327. https://doi.org/10.1038/s41593-018-0301-3.

2. Chung, K., and Deisseroth, K. (2013). CLARITY for mapping the nervous system. Nat. Methods 10, 508–513. https://doi.org/10.1038/nmeth.2481.

3. Ertürk, A., Becker, K., Jährling, N., Mauch, C.P., Hojer, C.D., Egen, J.G., Hellal, F., Bradke, F., Sheng, M., and Dodt, H.-U. (2012). Three-dimensional imaging of solvent-cleared organs using 3DISCO. Nat. Protoc. 7, 1983–1995. https://doi.org/10.1038/nprot.2012.119.

4. Nudell, V., Wang, Y., Pang, Z., Lal, N.K., Huang, M., Shaabani, N., Kanim, W., Teijaro, J., Maximov, A., and Ye, L. (2022). HYBRiD: hydrogel-reinforced DISCO for clearing mammalian bodies. Nat. Methods 19, 479–485. https://doi.org/10.1038/s41592-022-01427-0.

5. Renier, N., Adams, E.L., Kirst, C., Wu, Z., Azevedo, R., Kohl, J., Autry, A.E., Kadiri, L., Umadevi Venkataraju, K., Zhou, Y., et al. (2016). Mapping of brain activity by automated volume analysis of immediate early genes. Cell 165, 1789–1802. https://doi.org/10.1016/j.cell.2016.05.007.

6. Richardson, D.S., and Lichtman, J.W. (2015). Clarifying tissue clearing. Cell 162, 246–257. https://doi.org/10.1016/j.cell.2015.06.067.

7. Susaki, E.A., Tainaka, K., Perrin, D., Yukinaga, H., Kuno, A., and Ueda, H.R. (2015). Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. Nat. Protoc. 10, 1709–1727. https://doi.org/10.1038/nprot.2015.085.

8. Kubota, S.I., Takahashi, K., Nishida, J., Morishita, Y., Ehata, S., Tainaka, K., Miyazono, K., and Ueda, H.R. (2017). Whole-Body profiling of cancer metastasis with single-cell resolution. Cell Rep. 20, 236–250. https://doi.org/10.1016/j.celrep.2017.06.010.

9. Moatti, A., Li, C., Sivadanam, S., Cai, Y., Ranta, J., Piedrahita, J.A., Cheng, A.G., Ligler, F.S., and Greenbaum, A. (2022). Ontogeny of cellular organization and LGR5 expression in porcine cochlea revealed using tissue clearing and 3D imaging. iScience 25, 104695. https://doi.org/10.1016/j.isci.2022.104695.

10. Ueda, H.R., Ertürk, A., Chung, K., Gradinaru, V., Chédotal, A., Tomancak, P., and Keller, P.J. (2020). Tissue clearing and its applications in neuroscience. Nat. Rev. Neurosci. 21, 61–79. https://doi.org/10.1038/s41583-019-0250-1.

11. Zhao, S., Todorov, M.I., Cai, R., -Maskari, R.A., Steinke, H., Kemter, E., Mai, H., Rong, Z., Warmer, M., Stanic, K., et al. (2020). Cellular and molecular probing of intact human organs. Cell 180, 796–812.e19. https://doi.org/10.1016/j.cell.2020.01.030.

12. Frasconi, P., Silvestri, L., Soda, P., Cortini, R., Pavone, F.S., and Iannello, G. (2014). Large-scale automated identification of mouse brain cells in

confocal light sheet microscopy images. Bioinformatics *30*, i587–i593. https://doi.org/10.1093/bioinformatics/btu469.

13. Silvestri, L., Müllenbroich, M.C., Costantini, I., Di Giovanna, A.P., Mazzamuto, G., Franceschini, A., Kutra, D., Kreshuk, A., Checcucci, C., Toresano, L.O., et al. (2021). Universal autofocus for quantitative volumetric microscopy of whole mouse brains. Nat. Methods *18*, 953–958. https://doi.org/10.1038/s41592-021-01208-1.

14. Matsumoto, K., Mitani, T.T., Horiguchi, S.A., Kaneshiro, J., Murakami, T.C., Mano, T., Fujishima, H., Konno, A., Watanabe, T.M., Hirai, H., and Ueda, H.R. (2019). Advanced CUBIC tissue clearing for whole-organ cell profiling. Nat. Protoc. *14*, 3506–3537. https://doi.org/10.1038/s41596-019-0240-9.

15. Krupa, O., Fragola, G., Hadden-Ford, E., Mory, J.T., Liu, T., Humphrey, Z., Rees, B.W., Krishnamurthy, A., Snider, W.D., Zylka, M.J., et al. (2021). NuMorph: tools for cortical cellular phenotyping in tissue-cleared whole-brain images. Cell Rep. *37*, 109802. https://doi.org/10.1016/j.celrep.2021.109802.

16. Zhou, B., Kaucka, M., Chagin, A.S., and Newton, P.T. (2019). Clonal genetic tracing using the Confetti mouse to study mineralized tissues. J. Vis. Exp. *152*, e60424. https://doi.org/10.3791/60424.

17. Beattie, R., Postiglione, M.P., Burnett, L.E., Laukoter, S., Streicher, C., Pauler, F.M., Xiao, G., Klezovitch, O., Vasioukhin, V., Ghashghaei, T.H., and Hippenmeyer, S. (2017). Mosaic analysis with Double markers reveals distinct sequential functions of Lgl1 in neural stem cells. Neuron *94*, 517–533.e3. https://doi.org/10.1016/j.neuron.2017.04.012.

18. Johnson, C.A., and Ghashghaei, H.T. (2020). Sp2 regulates late neurogenic but not early expansive divisions of neural stem cells underlying population growth in the mouse cortex. Development *147*, dev186056. https://doi.org/10.1242/dev.186056.

19. Liang, H., Hippenmeyer, S., and Ghashghaei, H.T. (2012). A Nestin-cre transgenic mouse is insufficient for recombination in early embryonic neural progenitors. Biol. Open *1*, 1200–1203. https://doi.org/10.1242/bio.20122287.

20. Liang, H., Xiao, G., Yin, H., Hippenmeyer, S., Horowitz, J.M., and Ghashghaei, H.T. (2013). Neural development is dependent on the function of specificity protein 2 in cell cycle progression. Development *140*, 552–561. https://doi.org/10.1242/dev.085621.

21. Zhang, X., Mennicke, C.V., Xiao, G., Beattie, R., Haider, M.A., Hippenmeyer, S., and Ghashghaei, H.T. (2020). Clonal analysis of gliogenesis in the cerebral cortex reveals stochastic expansion of glia and cell autonomous responses to Egfr dosage. Cells *9*, e122662. https://doi.org/10.3390/cells9122662.

22. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature *521*, 436–444. https://doi.org/10.1038/nature14539.

23. Bayramoglu, N., Kannala, J., and Heikkila, J. (2016). Deep learning for magnification independent breast cancer histopathology image classification. In 2016 23rd International Conference on Pattern Recognition (ICPR) (IEEE), pp. 2440–2445. https://doi.org/10.1109/ICPR.2016.7900002.

24. Hollandi, R., Szkalisity, A., Toth, T., Tasnadi, E., Molnar, C., Mathe, B., Grexa, I., Molnar, J., Balind, A., Gorbe, M., et al. (2020). nucleAIzer: a parameter-free deep learning framework for nucleus segmentation using image style transfer. Cell Syst. *10*, 453–458.e6. https://doi.org/10.1016/j.cels.2020.04.003.

25. Sità, L., Brondi, M., Lagomarsino de Leon Roig, P., Curreli, S., Panniello, M., Vecchia, D., and Fellin, T. (2022). A deep-learning approach for online cell identification and trace extraction in functional two-photon calcium imaging. Nat. Commun. *13*, 1529. https://doi.org/10.1038/s41467-022-29180-0.

26. Pan, C., Schoppe, O., Parra-Damas, A., Cai, R., Todorov, M.I., Gondi, G., von Neubeck, B., Böğürcü-Seidel, N., Seidel, S., Sleiman, K., et al. (2019). Deep learning reveals cancer metastasis and therapeutic antibody targeting in the entire body. Cell *179*, 1661–1676.e19. https://doi.org/10.1016/j.cell.2019.11.013.

27. Todorov, M.I., Paetzold, J.C., Schoppe, O., Tetteh, G., Shit, S., Efremov, V., Todorov-Völgyi, K., Düring, M., Dichgans, M., Piraud, M., et al. (2020). Machine learning analysis of whole mouse brain vasculature. Nat. Methods *17*, 442–449. https://doi.org/10.1038/s41592-020-0792-1.

28. Geng, J., Zhang, X., Prabhu, S., Shahoei, S.H., Nelson, E.R., Swanson, K.S., Anastasio, M.A., and Smith, A.M. (2021). 3D microscopy and deep learning reveal the heterogeneity of crown-like structure microenvironments in intact adipose tissue. Sci. Adv. *7*, eabe2480. https://doi.org/10.1126/sciadv.abe2480.

29. Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2018). Focal loss for dense object detection. Preprint at arXiv. https://doi.org/10.48550/arXiv.1708.02002.

30. Cai, Y., Zhang, X., Kovalsky, S.Z., Ghashghaei, H.T., and Greenbaum, A. (2021). Detection and classification of neurons and glial cells in the MADM mouse brain using RetinaNet. PLoS One *16*, e0257426. https://doi.org/10.1371/journal.pone.0257426.

31. Waithe, D., Brown, J.M., Reglinski, K., Diez-Sevilla, I., Roberts, D., and Eggeling, C. (2020). Object detection networks and augmented reality for cellular detection in fluorescence microscopy. J. Cell Biol. *219*, e201903166. https://doi.org/10.1083/jcb.201903166.

32. Zhang, X., Xiao, G., Johnson, C., Cai, Y., Horowitz, Z.K., Mennicke, C., Coffey, R., Haider, M., Threadgill, D., Eliscu, R., et al. (2023). Bulk and mosaic deletions of Egfr reveal regionally defined gliogenesis in the developing mouse forebrain. iScience *26*, 106242. https://doi.org/10.1016/j.isci.2023.106242.

33. Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., et al. (2020). The allen mouse brain common coordinate framework: a 3D reference atlas. Cell *181*, 936–953.e20. https://doi.org/10.1016/j.cell.2020.04.007.

34. Hippenmeyer, S., Youn, Y.H., Moon, H.M., Miyamichi, K., Zong, H., Wynshaw-Boris, A., and Luo, L. (2010). Genetic mosaic dissection of Lis1 and Ndel1 in neuronal migration. Neuron *68*, 695–709. https://doi.org/10.1016/j.neuron.2010.09.027.

35. Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D., and Luo, L. (2005). Mosaic analysis with Double markers in mice. Cell *121*, 479–492. https://doi.org/10.1016/j.cell.2005.02.012.

36. Laukoter, S., Beattie, R., Pauler, F.M., Amberg, N., Nakayama, K.I., and Hippenmeyer, S. (2020). Imprinted Cdkn1c genomic locus cell-autonomously promotes cell survival in cerebral cortex development. Nat. Commun. *11*, 195. https://doi.org/10.1038/s41467-019-14077-2.

37. Li, C., Moatti, A., Zhang, X., Troy Ghashghaei, H., and Greenabum, A. (2021). Deep learning-based autofocus method enhances image quality in light-sheet fluorescence microscopy. Biomed. Opt Express *12*, 5214–5226. https://doi.org/10.1364/BOE.427099.

38. Li, C., Rai, M.R., Ghashghaei, H.T., and Greenbaum, A. (2022). Illumination angle correction during image acquisition in light-sheet fluorescence microscopy using deep learning. Biomed. Opt Express *13*, 888–901. https://doi.org/10.1364/BOE.447392.

39. Moatti, A., Cai, Y., Li, C., Sattler, T., Edwards, L., Piedrahita, J., Ligler, F.S., and Greenbaum, A. (2020). Three-dimensional imaging of intact porcine cochlea using tissue clearing and custom-built light-sheet microscopy. Biomed. Opt Express *11*, 6181–6196. https://doi.org/10.1364/BOE.402991.

40. Bria, A., and Iannello, G. (2012). TeraStitcher - a tool for fast automatic 3D-stitching of teravoxel-sized microscopy images. BMC Bioinf. *13*, 316. https://doi.org/10.1186/1471-2105-13-316.

41. Zou, Z., Shi, Z., Guo, Y., and Ye, J. (2019). Object detection in 20 Years: a survey. Preprint at arXiv. https://doi.org/10.48550/arXiv.1905.05055.

42. Tomer, R., Lovett-Barron, M., Kauvar, I., Andalman, A., Burns, V.M., Sankaran, S., Grosenick, L., Broxton, M., Yang, S., and Deisseroth, K. (2015). SPED light sheet microscopy: fast mapping of biological system structure and function. Cell *163*, 1796–1806. https://doi.org/10.1016/j.cell.2015.11.061.

43. Yang, S.J., Berndl, M., Michael Ando, D., Barch, M., Narayanaswamy, A., Christiansen, E., Hoyer, S., Roat, C., Hung, J., Rueden, C.T., et al. (2018). Assessing microscope image focus quality with deep learning. BMC Bioinf. *19*, 77. https://doi.org/10.1186/s12859-018-2087-4.

44. Everingham, M., Eslami, S.M.A., Van Gool, L., Williams, C.K.I., Winn, J., and Zisserman, A. (2015). The pascal visual object classes challenge: a retrospective. Int. J. Comput. Vis. *111*, 98–136. https://doi.org/10.1007/s11263-014-0733-5.

45. Belle, V., and Papantonis, I. (2021). Principles and practice of explainable machine learning. Front. Big Data *4*, 688969.

46. Hossin, M., and Mn, S. (2015). A review on evaluation metrics for data classification evaluations. International Journal of Data Mining & Knowledge Management Process *5*, 01–11. https://doi.org/10.5121/ijdkp.2015.5201.

47. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-RFP (1:200)	Rockland	Cat# 600-401-379; RRID: AB_2209751
Chicken anti-GFP (1:200)	Aves Labs	Cat# GFP-1020; RRID: AB_10000240
Cy3-AffiniPure Donkey Anti-Rabbit	Jackson ImmunoResearch Labs	Cat# 711-165-152; RRID: AB_2307443
Alexa Fluor 647-AffiniPure Donkey Anti-Chicken	Jackson ImmunoResearch Labs	Cat# 703-605-155; RRID: AB_2340379
Deposited Data		
Raw imaging data	This paper	Dryad: https://doi.org/10.5061/dryad.fj6q57400
Experimental models: Organisms/strains		
Mouse: $Emx1^{cre}$:$MADM11^{TG/GT}$:$Egfr^{+/+}$, $Egfr^{F/+}$ and $Egfr^{F/F}$	Zhang et al.[32]	N/A
Software and algorithms		
COMBINe	This paper	Zenodo: https://doi.org/10.5281/zenodo.7746971
Python	Python Software Foundation	RRID: SCR_008394
MATLAB R2021a	Mathworks	RRID: SCR_001622
NuMorph	Krupa et al.[15]	https://github.com/ok37/numorph
Keras RetinaNet	Fizyr	https://github.com/fizyr/keras-retinanet
TeraStitcher	Bria and Iannello[40]	http://abria.github.io/TeraStitcher/
ClearMap	Renier et al.[5]	https://github.com/ChristophKirst/ClearMap2
Imaris 9.5	Oxford Instruments	RRID: SCR_007370

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Alon Greenbaum (greenbaum@ncsu.edu).

Materials availability
This study did not generate new unique reagents.

Data and code availability
- An example mouse brain dataset has been deposited at Dryad and is publicly available as of the date of publication. The DOI is listed in the key resources table.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. The DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice were housed in a 12-hour light:dark cycle with *ad libitum* access to food and water. All procedures were performed under the regulations and approval from Institutional Animal Care and Use Committee at North Carolina State University. *Emx1^{cre}: MADM11^{TG/GT}:Egfr^{+/+}*, *Egfr^{F/+}* and *Egfr^{F/F}* mice were generated using breeding schemes previously described[21]. Six males and three females were sacrificed at P30. Three brain hemispheres collected from three mice of each genotype were tissue cleared and analyzed. Sex was not considered as a variable in this study.

METHOD DETAILS

Tissue clearing
The iDISCO+ protocol[5] was used to stain and clear the mouse brain hemispheres. Brain hemispheres were dehydrated at room temperature (RT) using methanol gradients diluted in ddH_2O: 20%, 40%, 60%, 80%, 100%, 100%; 1 h each. Brain hemispheres were

then chilled at 4 °C for 2 h and submerged in the mixture of dichloromethane and methanol (v:v=2:1) overnight at RT. After washing twice in 100% methanol at RT, brain hemispheres were chilled at 4°C for 2 h and decolorized in chilled fresh 5% H_2O_2 in methanol overnight at 4°C. After bleaching, brain hemispheres were rehydrated with methanol series diluted in ddH_2O (100%, 80%, 60%, 40%, 20%), followed by one wash with PBS and two washes with PBS with 0.02% Triton X-100, 1 h each at RT. Next, brain hemispheres were permeabilized with permeabilization solution (0.2% Triton X-100, 0.3 M Glycine, 20% DMSO in PBS) for 2 d at 37°C and blocked with blocking solution (0.2% Triton X-100, 6% donkey serum, 10% DMSO in PBS) for 2 d at 37°C. Brain hemispheres were then incubated with chicken anti-GFP (1:200, Aves Labs) and rabbit anti-RFP (1:200, Rockland) in PTwH (0.2% Tween-20, 10 mg/L heparin in PBS) with 5% DMSO (v/v) and 3% donkey serum (v/v) for 7 d at 37°C. After primary staining, brain hemispheres were washed 5 times with PTwH, 2 h each until the next day. Next, brain hemispheres were incubated with donkey anti-chicken AlexaFluor 647 (1:200, Jackson ImmunoResearch) and donkey anti-rabbit Cy3 (1:200, Jackson ImmunoResearch) in PTwH with 3% donkey serum for 5 d at 37°C, followed by 5 washes with PTwH. For optical clearing, brain hemispheres were dehydrated with methanol series diluted in ddH_2O (20%, 40%, 60%, 80%, 100%), 1 h each at RT, and incubated in the mixture of dichloromethane and methanol (v:v=2:1) for 3 h at RT. Brain hemispheres were then submerged in dichloromethane 15 minutes twice at RT. All the incubation steps above were performed with gentle shaking. Brain hemispheres were cleared with dibenzyl ether at RT.

Light-sheet imaging
Cleared mouse brain hemispheres were imaged using a custom-built light-sheet fluorescence microscope, of which the setup was outlined in.[37–39] Each brain hemisphere was imaged by illumination of 561 nm and 640 nm respectively, with a voxel size of 0.65 × 0.65 × 10 μm³. During image acquisition, every 1 mm in z direction, autofocus method would be applied to correct for the drift between the light-sheet plane and the detection focal plane.

Channel alignment
When the illumination beams of different wavelengths are parallel, images of different color channels can be matched through translation in x, y, and z-axes. NuMorph[15] was employed for local channel alignment on each imaged tile (MATLAB). First, displacement in z-planes was determined by maximizing phase correlation between two channels. After adjustment in z, color registration was performed in x- and y-axes.

Cell detection
Both the training data and the validation data contained fluorescence images from different imaging modalities, e.g., slide scanner, confocal fluorescence microscopy, and LSFM.[30] 3361 individual cells (2497 neurons and 864 astrocytes) were labeled in the training data. 1164 individual cells (896 neurons and 268 astrocytes) were labeled in the validation data. Test data was generated from light-sheet datasets that were analyzed in this paper. 1664 individual cells (1462 neurons and 202 astrocytes) were labeled in the test data.

A RetinaNet[29] model was trained to detect cells of six classes: yellow, red, green neurons and astrocytes. The repository cloned from the source (https://github.com/fizyr/keras-retinanet) was adapted to this work (https://github.com/yccc12/COMBINe/keras-retinanet). Transfer learning was adopted with a pre-trained ResNet50 being used as the backbone. Data augmentation strategies such as geometrical transforms, color swap and saturation simulation were applied.[30] The initial learning rate was 0.0001 and the batch size was four. Using an Adam optimizer, the model was trained for 50 epochs. The model with the best performance on validation data was selected as the inference model. Average precision measure[44] was adopted to evaluate model performance. The average precision measure is widely used in object detection tasks, and it is very similar to the area under the precision-recall curve. Other measures such as precision, recall, and F-scores[46] were reported based on a confidence threshold of 0.5.

Because stitching was likely to generate artifacts, cell detection was performed per image tile, during which images were cropped into 512 × 512 image patches with an overlap fraction of 0.125 for inference. Detections in overlapped areas were merged based on intersection over union and confidence score.

Stitching
TeraStitcher[40] advanced mode was used to stitch the acquired datasets, with misalignment across tiles being corrected. Parameters output from TeraStitcher were adopted to stitch detected cells of each tile. Duplicate detections in overlapping regions between adjacent tiles were rejected. Redundant detections across different z-planes were merged based on signal intensities when the output bounding boxes between adjacent z-planes were found overlapped, which was determined by an intersection over union above 0.5 between two bounding boxes. Stitched datasets and detected cells were visualized in Imaris 9.5 (Oxford Instruments).

Registration
ClearMap[5] in Linux was utilized to register stitched datasets to Allen Brain Atlas[33] and to map detected cells to annotated brain regions.

Visualization
Cell densities and statistics were visualized in MATLAB using a custom script.

QUANTIFICATION AND STATISTICAL ANALYSIS

Regional cell densities between different groups were compared in MATLAB using unpaired t-tests assuming unequal variances. *p*-values were adjusted by the Benjamini–Hochberg procedure[47] to control the false discovery rate (< 0.05) for simultaneous multiple testing. All statistics were reported as mean ± standard deviation (SD).