

# A three-dimensional single-cell-resolution whole-brain atlas using CUBIC-X expansion microscopy and tissue clearing

Tatsuya C. Murakami<sup>1</sup>, Tomoyuki Mano<sup>1,2</sup>, Shu Saikawa<sup>3</sup>, Shuhei A. Horiguchi<sup>4,5</sup>, Daichi Shigeta<sup>1</sup>, Kousuke Baba<sup>6,7</sup>, Hiroshi Sekiya<sup>8</sup>, Yoshihiro Shimizu<sup>9</sup>, Kenji F. Tanaka<sup>10</sup>, Hiroshi Kiyonari<sup>11</sup>, Masamitsu Iino<sup>8,12</sup>, Hideki Mochizuki<sup>12</sup>, Kazuki Tainaka<sup>1,13</sup> and Hiroki R. Ueda<sup>1,2,4\*</sup>

**A three-dimensional single-cell-resolution mammalian brain atlas will accelerate systems-level identification and analysis of cellular circuits underlying various brain functions. However, its construction requires efficient subcellular-resolution imaging throughout the entire brain. To address this challenge, we developed a fluorescent-protein-compatible, whole-organ clearing and homogeneous expansion protocol based on an aqueous chemical solution (CUBIC-X). The expanded, well-cleared brain enabled us to construct a point-based mouse brain atlas with single-cell annotation (CUBIC-Atlas). CUBIC-Atlas reflects inhomogeneous whole-brain development, revealing a significant decrease in the cerebral visual and somatosensory cortical areas during postnatal development. Probabilistic activity mapping of pharmacologically stimulated Arc-dVenus reporter mouse brains onto CUBIC-Atlas revealed the existence of distinct functional structures in the hippocampal dentate gyrus. CUBIC-Atlas is shareable by an open-source web-based viewer, providing a new platform for whole-brain cell profiling.**

The mammalian brain is composed of various cellular circuits of different physiological functions. Comprehensive analysis of such complex cellular circuits in the entire mammalian brain is one of the fundamental challenges in neuroscience. Toward this goal, a number of mammalian brain atlases, including ones for nonhuman primates, have been constructed<sup>1–8</sup> and provided useful anatomical platforms. These mammalian brain atlases can be used (i) as a platform for mapping different cellular populations over the entire brain when combined with genome-wide expression data by *in situ* hybridization<sup>9</sup> and microarrays<sup>10,11</sup>, (ii) as a platform for mapping cellular activities by measuring gene expression of immediate early genes<sup>12–14</sup> and (iii) as a platform for mapping cellular connections when combined with neural projection analysis based on adeno-associated virus<sup>15</sup> or rabies virus<sup>16,17</sup>.

Three-dimensional (3D) reconstitution of imaging data acquired from 2D serial sections has been a common strategy for mammalian brain atlases<sup>1–3,8</sup>. Although the image-based atlases can provide high-resolution information, the enormity of the data prevents flexible editing and update of the atlas. Considering that the cell is the basic unit of life, a cell-based atlas would provide an attractive alternative to image-based atlases. A compact atlas of this type may be editable via an open-source platform and therefore easily updated by overlaying various cellular functions (for example, activity, gene expression, cell type and neural connection).

To construct a single-cell-resolution mouse brain atlas, it is essential to accurately identify whole cells in the whole brain. To this end,

an appropriate imaging technique should provide both subcellular resolution and whole-brain-scale coverage. The recent development of potent tissue-clearing methods, including BABB, 3DISCO and uDISCO, CLARITY/PACT-PARS, CUBIC and others<sup>18–24</sup>, enables rapid single-cell-resolution imaging of an adult mouse brain when combined with light-sheet fluorescence microscopy (LSFM). However, rapid subcellular-resolution imaging over the entire brain is still challenging because of the limited resolution of current LSFM techniques and insufficient transparency of tissues.

Expansion microscopy, developed by the Boyden group<sup>25</sup>, and another expansion protocol developed by the Gradinaru group<sup>26</sup> provide a new approach to improving optical resolution by physically expanding tissues. The Chung group further applied expansion microscopy to whole organs<sup>27</sup>. These polymer-based expansion microscopy methods may allow subcellular-resolution imaging of samples. In addition to the improved resolution, tissue expansion is expected to provide optical advantages in tissue transparency because the refractive indices (RI) inside tissues would become more homogenous. However, since excess expansion makes tissue too fragile<sup>28</sup>, moderate tissue expansion is called for.

In this study, we developed an editable, point-based mouse brain atlas with single-cell resolution by an intensive tissue-clearing method combined with tissue expansion. Describing the whole mouse brain as an ensemble of cellular points with less than 3 GB of data was achieved by subcellular-resolution imaging of a nuclear-stained brain and automatic detection of cells. With reference to

<sup>1</sup>Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. <sup>2</sup>International Research Center for Neurointelligence (WPI-IRCN), UTIAS, The University of Tokyo, Tokyo, Japan. <sup>3</sup>Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan. <sup>4</sup>Laboratory for Synthetic Biology, RIKEN Quantitative Biology Center, Osaka, Japan. <sup>5</sup>Department of Systems Science, School of Engineering Science, Osaka University, Osaka, Japan. <sup>6</sup>Department of Neurology, Graduate School of Medicine, Osaka University, Osaka, Japan. <sup>7</sup>Department of Kampo Medicine, Graduate School of Medicine, Osaka University, Osaka, Japan. <sup>8</sup>Department of Pharmacology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. <sup>9</sup>Laboratory for Cell-Free Protein Synthesis, RIKEN Quantitative Biology Center, Osaka, Japan. <sup>10</sup>Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan. <sup>11</sup>Animal Resource Development Unit and Genetic Engineering Team, RIKEN Center for Life Science Technologies, Kobe, Japan. <sup>12</sup>Division of Cellular and Molecular Pharmacology, Nihon University School of Medicine, Tokyo, Japan. <sup>13</sup>Brain Research Institute, Niigata University, Niigata, Japan. \*e-mail: [uedah-tky@umin.ac.jp](mailto:uedah-tky@umin.ac.jp)

anatomical annotations from the Allen Brain Atlas<sup>8</sup>, this single-cell-based brain atlas, CUBIC-Atlas, allowed us to count the number of cells in the annotated anatomical region. Region-specific cell counting in the developing mouse brain revealed a significant decrease in cell numbers in the cerebral visual and primary somatosensory cortical areas during early postnatal development. By probabilistically mapping pharmacologically activated neural cells onto CUBIC-Atlas, we also discovered a functionally distinct brain region in the hippocampal dentate gyrus. This single-cell-resolution analytical platform, CUBIC-Atlas, is thus suitable for unbiased system-level cellular analysis.

## Results

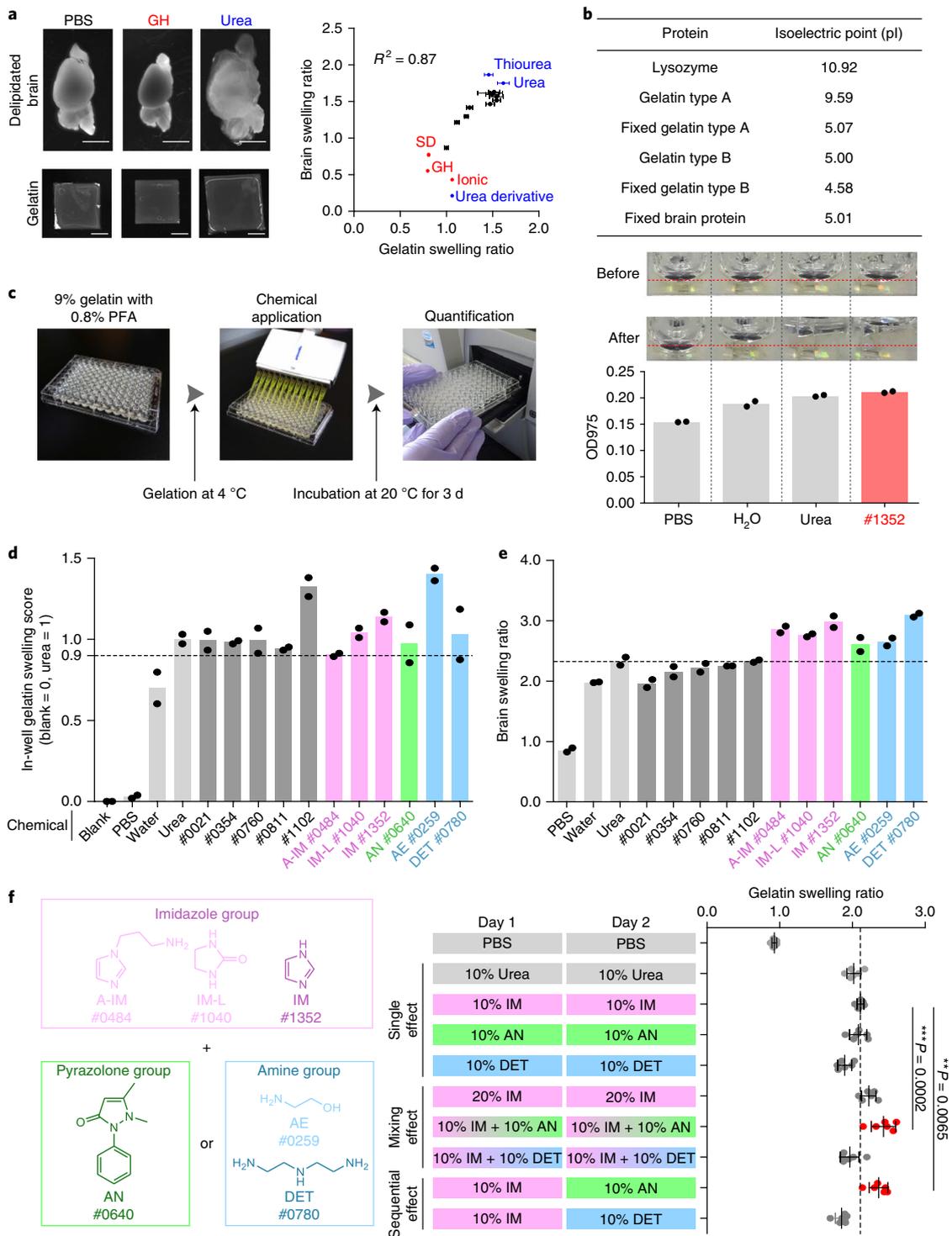
**Identification of swelling reagents by comprehensive chemical screening.** To construct an imaging technique with both subcellular resolution and whole-brain-scale coverage, one feasible approach would be volumetric imaging by tissue expansion. All current expansion microscopy methods are based on hydrogel-embedding of tissues, some of which enable imaging of fluorescent proteins<sup>26,29</sup>. Conventionally, expansion protocols have been limited to tissue slices or cultured cells. Recently, a whole-organ expansion protocol was reported<sup>27</sup>, termed magnified analysis of proteome (MAP), which successfully demonstrates fourfold expansion in one dimension. However, subcellular imaging with whole-brain-scale coverage could not be achieved due to the lack of a hyperhydrative RI matching reagent that can maintain the expansion and the lack of high-resolution objective lenses that can cover an entire expanded brain. Furthermore, because tissue denaturation at high temperature is required, MAP is not able to observe fluorescent proteins without immunostaining. Therefore, expansion protocols that allow both mild whole-organ expansion and retention of fluorescent protein are still needed. Considering that too much expansion causes fragility of tissues<sup>28</sup> and that commercially available high-resolution objective lenses are limited in their working distance (WD; 8-mm XLPLN-series 10× and 25× Olympus lenses), we decided to develop a moderate, ten-fold volume-expansion protocol. Because simple immersion of organs into aqueous chemical solution retains fluorescent protein signals well, and because some aqueous chemical solutions are known to promote tissue swelling<sup>23,30,31</sup>, we endeavored to develop an expansion protocol with aqueous chemical solutions.

To identify the chemicals that achieve both tissue expansion and RI matching, we performed comprehensive chemical screening. For parallelization of screening experiments, we used gelatin instead of brains because of the high similarity between paraformaldehyde (PFA)-fixed gelatin and PFA-fixed delipidated brains in terms of swelling behavior and isoelectric points (Fig. 1a,b and Supplementary Fig. 1). These results prompted us to establish high-throughput screening of swelling agents based on PFA-fixed gelatin gel (type B). To quantitatively evaluate the swelling ability, we measured the water absorption at 975 nm in chemically treated gelatin gel (Fig. 1c and Supplementary Fig. 1). For the chemical screening, we first chose 1,691 chemicals out of over 25,000 commercial chemicals from Tokyo Chemical Industry in terms of their potential water-solubility. The screening revealed 11 chemicals (Supplementary Table 1) with greater swelling ability than that of a conventional swelling chemical, urea (Fig. 1d), and further screening using mouse brains revealed that six of these 11 swelled the delipidated brain more than urea (Fig. 1e). We next asked whether they can promote further swelling if mixed together. To simplify the experiment, we classified these final candidates into three groups (imidazole, pyrazolone and amine groups) based on their chemical structures and chose a representative chemical from each group. The combinatorial swelling effect was tested in the forms of a mixture or a sequential procedure with fixed gelatin blocks. The result suggested significant synergistic effect of imidazole and antipyrine (Fig. 1f). Based on this result,

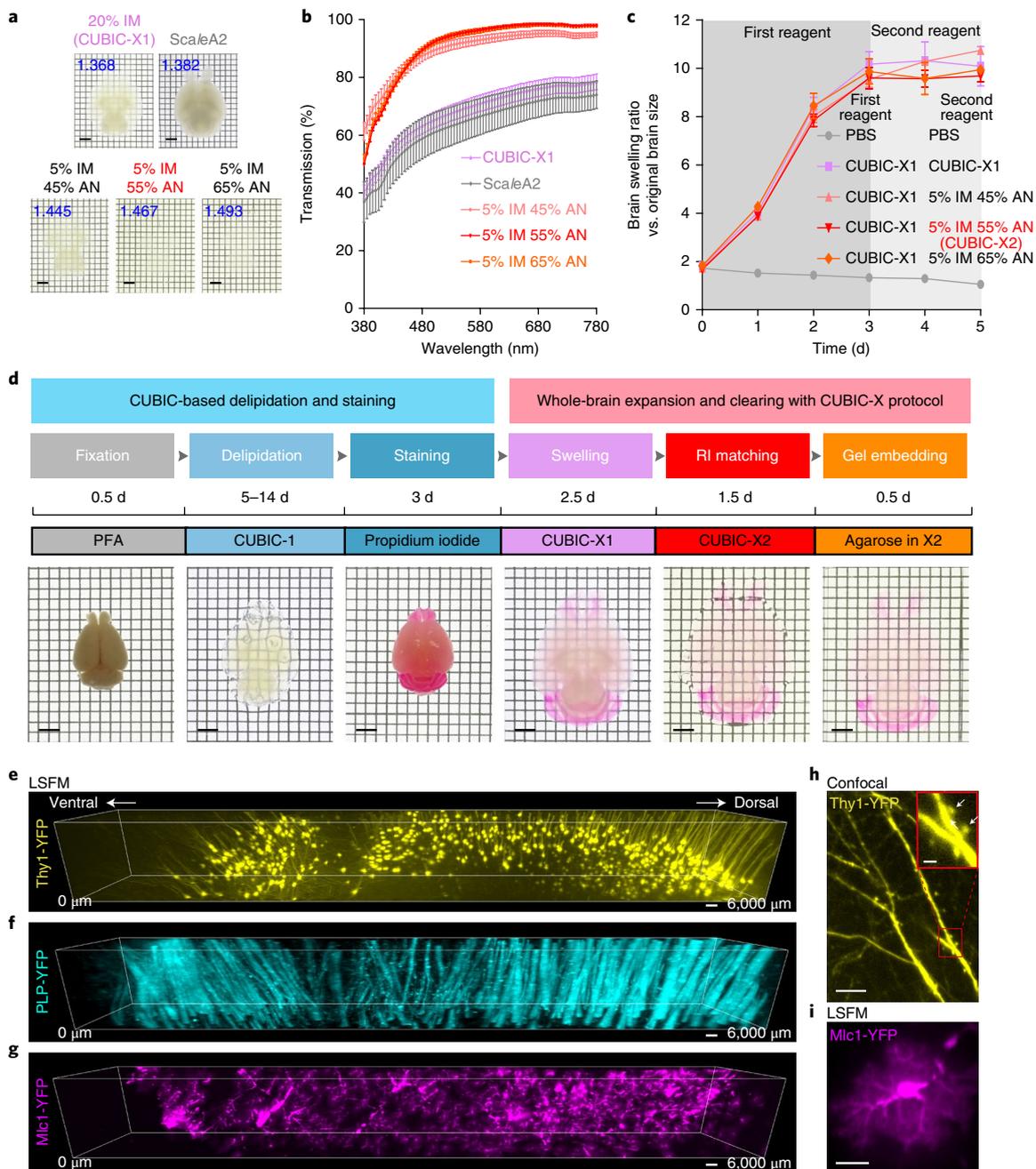
we designed the whole-brain expansion protocol by serially treating brains with imidazole and antipyrine.

**CUBIC-X for whole-brain expansion with retention of native proteins.** To ensure compatibility with 10× volume expansion, tissue clearing and rigid sample mounting on LSFM, we developed the following four-step protocol: (i) tissue delipidation, (ii) volume expansion, (iii) RI matching and (iv) gel embedding. Brains were delipidated by ScaleCUBIC-1, then immersed in aqueous imidazole solutions as the volume expansion step. The 20% imidazole solution, termed CUBIC-X1, was the minimal concentration to achieve the largest expansion volume (Supplementary Fig. 2). To promote the transparency of expanded brains, we next explored a high-RI medium compatible with the expanded brain tissue. Antipyrine, another candidate swelling reagent, displayed remarkable water solubility and high molar refractivity (Supplementary Fig. 2). We found that the mixture of 5% imidazole and varying amounts of antipyrine showed high transparency while maintaining the expansion (Fig. 2a,b). On the other hand, other examined conventional aqueous RI media could not maintain the expanded size or lacked high clearing ability (Fig. 2a,b and Supplementary Fig. 2c). The clearing performance of an imidazole–antipyrine cocktail was also quantitatively confirmed against the USAF resolution chart (Methods and Supplementary Fig. 3d–f). In view of handling, we chose the 5% imidazole + 55% antipyrine cocktail (RI = 1.467) as the RI medium, termed CUBIC-X2. In the final 5-d expansion protocol, we obtained a 10× expanded brain with almost full transparency (Fig. 2c,d). For rigid sample mounting on a microscope, gelation was performed by adding 2% agarose to the CUBIC-X2 medium. CUBIC-X2-based agarose gel was highly transparent and solid (Fig. 2d and Supplementary Fig. 3). A series of expanded tissue-clearing protocols, termed CUBIC-X, was compatible with the fluorescent dye staining of cell nuclei (Fig. 2d). Furthermore, our CUBIC-X protocol was also compatible with fluorescent reporter proteins by using a modified ScaleCUBIC-1 reagent (5% by weight (wt%) Quadrol, 10 wt% Triton X-100 and 25 wt% urea) at the delipidation step (Supplementary Fig. 4). We performed LSFM imaging of brains expressing YFP under Thy1, PLP and Mlcl promoters (see Methods), demonstrating clear imaging of fluorescent proteins deep in the tissues (Fig. 2e–g,i). High-NA (1.0) confocal imaging of Thy1-YFP mouse brain suggested the applicability of CUBIC-X to imaging of subcellular structure with CUBIC-X protocol (Fig. 2h).

**Whole-brain nuclei imaging of a CUBIC-X expanded brain with customized LSFM.** Because of the irregular and diversified morphologies of cells, comprehensive identification of cell bodies in a whole brain is challenging. Because relatively uniform round nuclei facilitate global cell detection and tracing<sup>32</sup>, we decided to identify the locations of nuclei as a surrogate for cell bodies. To this end, propidium iodide, an organic small molecule staining DNA, was applied in the following experiments for whole-brain cell detection. To clarify whether CUBIC-X improves imaging quality, we performed imaging of pre- and post-CUBIC-X brains by commercially available low-magnification LSFM (Methods). The resulting images revealed highly resolved fluorescent signals from the thalamus, located deep in the brain (Supplementary Fig. 5). We then asked whether a CUBIC-X-expanded brain retains original spatial information of nuclei. To verify whether brains hold nuclei during delipidation, we performed sequential observation of the nuclei with 1-mm-thick brain slices from R26-H2B-EGFP mice (Methods). We confirmed that none of the nuclei were lost during delipidation (Supplementary Fig. 6 and Methods). Next, we verified that none of the nuclei were lost during expansion by observing the nuclei in the corresponding regions of pre- and post-CUBIC-X brains (Supplementary Fig. 6 and Methods). Together, we concluded that none of the nuclei were lost during the whole CUBIC-X



**Fig. 1 | Identification of swelling reagents by comprehensive chemical screening.** **a**, Correlation of the swelling ratio between delipidated hemisphere brains ( $n=2$ ) and postfixed type B gelatins ( $n=3$ ) treated with PBS, water or 12 chemicals. These 12 chemicals were chosen from our previous study<sup>23</sup>. GH, guanidine hydrochloride; SD, sodium deoxycholate.  $R^2$  value was calculated with all individual data plots. Scale bars, 5 mm. **b**, Isoelectric points (pIs) of lysozyme, type A gelatin, type B gelatin and extracted brain protein. ‘Fixed’, fixation by PFA. **c**, Chemical screening procedure for swelling reagents from a 1,691-chemical library. Numbers (for example, #1352) indicate chemical identification numbers. OD975, 975-nm absorption. Each bar indicates the mean value. **d**, In-well gelatin gel-based first round of chemical screening ( $n=2$ ) identified 11 chemicals with relative swelling scores greater than 0.9 as a criterion. We highlighted several chemical groups: imidazole (pink), pyrazolone (green) and amine (blue). Each bar indicates the mean value. **e**, Hemisphere brain-based second chemical screening ( $n=2$ ) identified final six chemical candidates. Colors as in **d**. Each bar indicates the mean value. **f**, We classified the six chemical candidates into three groups based on their chemical structure: imidazole (pink), pyrazolone (green) and amine (blue). We evaluated potential additive swelling effects, such as solvent-mixing effects and solvent-substitution effects, by gelatin gel-based assay. A-IM, 1-(3-aminopropyl)imidazole; IM-L, 2-imidazolidinone; IM, imidazole; AN, antipyrine; AE, 2-aminoethanol; DET, diethylenetriamine. \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , one-way ANOVA with Tukey’s post hoc test for multiple comparisons;  $n=6$ . All values are mean  $\pm$  s.d.



**Fig. 2 | CUBIC-X for whole-brain expansion and hyperhydrative RI matching.** **a**, Transmission images of CUBIC-X1-treated whole brains after additional treatment with CUBIC-X1, ScaleA2 and imidazole-antipyrine cocktails. Chemical cocktails of 5% imidazole and 45–65% antipyrine provided almost fully transparent brains without shrinkage. RIs of each chemical are shown in blue. Scale bars, 2 mm. **b**, Transmission curves of the whole-brain samples shown in **a**. Light transmittance around the visible region (380–780 nm) was measured ( $n = 3$ ). **c**, Time course of swelling ratio of delipidated whole brains after chemical treatment ( $n = 3$ ). Brain swelling ratios were evaluated against original brain size. For handling purposes, we chose 5% imidazole + 55% antipyrine cocktail (CUBIC-X2). **d**, Overview of a whole-brain expansion and hyperhydrative RI matching protocol (CUBIC-X protocol) with staining of nuclei. The duration of CUBIC-1 delipidation can be modified to purpose. Scale bars, 2 mm. **e–g**, CUBIC-X was compatible with fluorescent protein imaging of various reporter mouse brains. Thy1-YFP, PLP-YFP and Mlc1-YFP mouse brains were cleared and expanded with CUBIC-X protocol and imaged with LSFM (10 $\times$ , NA = 0.6). Volume-rendered images are shown. Scale bars, 100  $\mu$ m. **h**, Confocal image (25 $\times$ , NA = 1.0) of Thy1-YFP mouse brain. Inset: magnified view showing spines (indicated by arrows). Scale bars, 30  $\mu$ m (main) and 5  $\mu$ m (inset). **i**, An LSFM image of an Mlc1-YFP astrocyte, from **g**. Maximum intensity projection was applied over the 30- $\mu$ m-thick volume. Scale bar, 40  $\mu$ m. Experiments were repeated once (**f**), twice (**e, g, i**) or three times (**h**) with independent brains. Representative images are shown. For **e–i**, a post-CUBIC-X brain was used as a baseline. All values are mean  $\pm$  s.d.

brain preparation procedure. The degree of expansion is not dependent on cell density (Supplementary Fig. 6). We next quantified the amount of distortion at macrostructural scale. By three-dimensionally comparing the pre-CUBIC-X brain and post-CUBIC-X brain,

root-mean-square errors were measured<sup>25</sup>. Root-mean-square errors were estimated at less than  $\sim 20 \mu$ m, which was  $\sim 0.8\%$  of the measurement length (2,500  $\mu$ m) in the investigated regions (Fig. 3a,b, Supplementary Fig. 7 and Methods).

To visualize the CUBIC-X brain in subcellular resolution using LSFM, we implemented a customized LSFM equipped with a long WD detection objective lens (Olympus XLPLN10XSVM, NA=0.6, WD=8 mm; Fig. 3c and Supplementary Fig. 8). A complete list of parts is available in Supplementary Table 2. The imaging region was divided into an  $x$ - $y$  grid, which typically consisted of 17–20 tiles in the  $x$  direction and 18–21 tiles in the  $y$  direction. The sample was scanned in  $z$  direction with a step size of 5  $\mu$ m to obtain a stack of images. To cover an entire brain, we also introduced rotation sequences (Fig. 3d, Supplementary Video 1 and Methods). As a result, we obtained clearly resolved images across the brain (Fig. 3e). We also note that the possible displacements of the imaging sequences were small enough ( $x$ - $y$  tiling:  $\leq 12 \mu$ m,  $\theta$  tiling:  $\leq 1.3\%$ ) to perform the following whole-brain cell detection (Supplementary Fig. 9 and Methods). The resulting whole-brain image data was approximately 14 TB with 16-bit raw TIFF format. We also performed whole-brain imaging of brains expressing YFP under the PLP and Thy1 promoters. These applications demonstrated subcellular-resolution imaging of fluorescent reporter proteins from whole mouse brains (Supplementary Fig. 10).

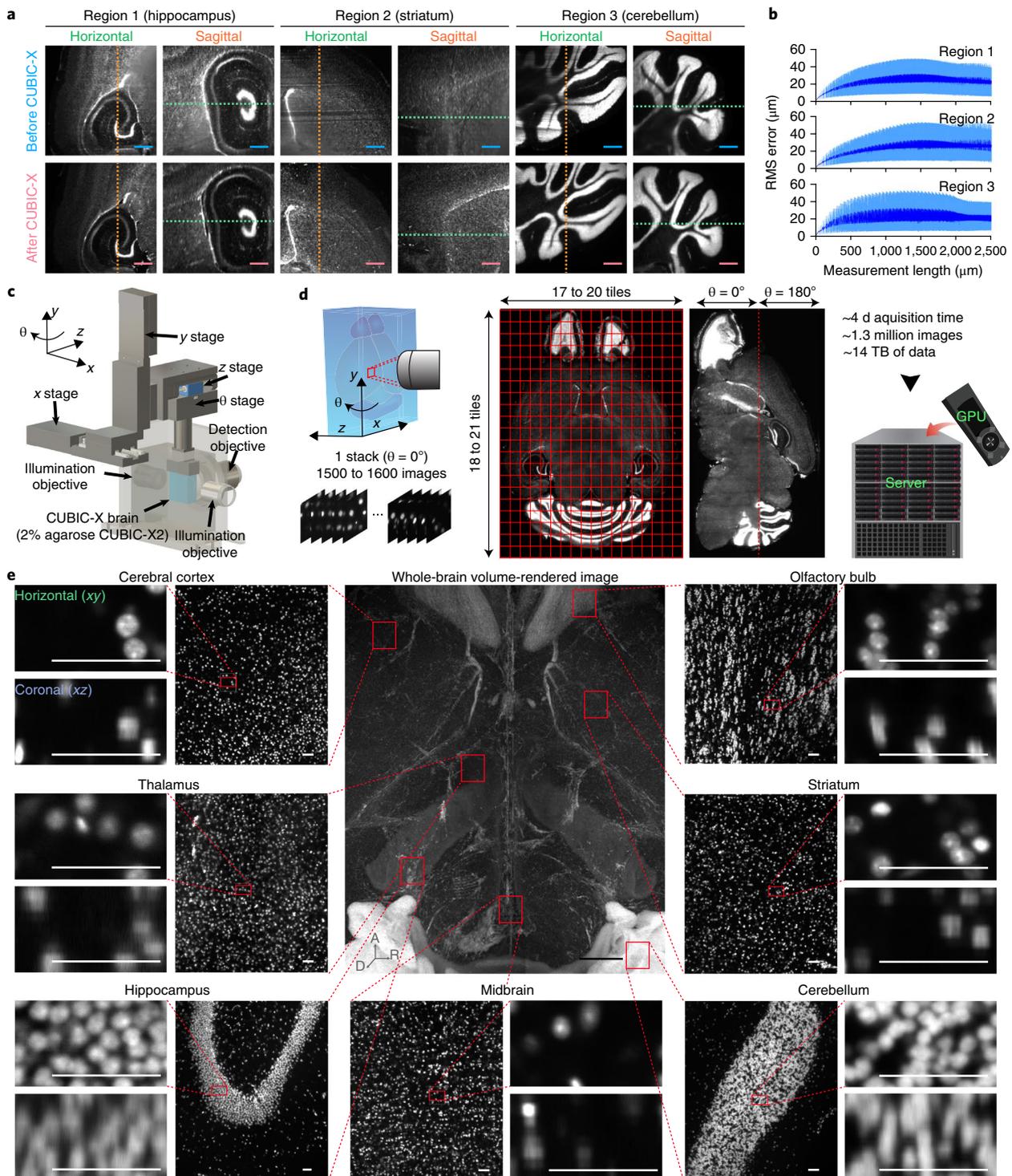
**Construction of a single-cell-resolution mouse brain atlas (CUBIC-Atlas).** The next challenge in constructing a single-cell-resolution atlas is accurately detecting all cells from the acquired images (Fig. 4a). To this end, we devised a cell-detection algorithm based on two-step convolution (Supplementary Fig. 11), which was analogous to a one-step convolution algorithm developed by Shimada et al.<sup>33</sup> In our two-step algorithm, we applied the first convolution with a 2D mean filter of a kernel diameter,  $D_{M1}$ , resulting an image with local maximum values at the center of the cellular nuclei. We limited the pixels to local maximum values, and other pixels were replaced with zero. Then we unified the multiply-detected cells by performing the second convolution upon this image, with a 3D mean filter of a kernel diameter,  $D_{M2}$ . Finally, we detected each cell by finding the local maximum in 3D by automatically comparing determined local thresholds (Supplementary Fig. 12 and Methods). By tuning diameters of two convolution kernels,  $D_{M1}$  and  $D_{M2}$  (Supplementary Fig. 11), we achieved fully automated cell detection. Evaluated by comparing the result to manual detection, accuracy exceeded 94% in interbrain (thalamus), midbrain and cerebrum (except hippocampus and olfactory bulb granular layer), 91% in hippocampal CA1, 90% in olfactory bulb granular layer, 89% in dentate gyrus and 79% in the densest area, the cerebellum granular layer (Supplementary Fig. 13 and Methods). Supplementary Video 2 shows a representative  $z$ -stack image analyzed by our cell-detection algorithm. We did not observe morphology- or size-dependent errors of cell detection with our algorithm. As a result, the spatial information of whole-brain cellular nuclei was represented as an ensemble of cellular points, occupying  $\sim 3$  GB and providing a pointillistic anatomical platform that we termed CUBIC-Atlas. CUBIC-Atlas is an editable platform that can also accommodate multiple types of cellular information in single-cell resolution, such as anatomical area, gene expression, cell type and cellular connections. As an initial step, anatomical information was annotated onto CUBIC-Atlas by adopting anatomical annotations in the Allen Brain Atlas (ABA)<sup>9</sup>. Because the ABA brain was reconstituted from a series of coronal-slice images, we created a registration based on the virtual coronal slices (Supplementary Fig. 14 and Methods). Using this registration, each individual cell in CUBIC-Atlas was assigned an anatomical identity from ABA (Fig. 4b–j, Supplementary Fig. 15 and Supplementary Videos 3 and 4). To examine how accurately CUBIC-Atlas was annotated, a nucleus stained brain of a Thy1-YFP-H mouse was registered onto CUBIC-Atlas. We confirmed that YFP signals of the brain could be accurately mapped to known anatomical locations, such as cerebral cortex layers 5 and 6<sup>34</sup> (Supplementary Fig. 16 and Methods).

**Whole-brain cell counting in mice.** One of the attractive applications of the CUBIC-Atlas is whole-brain cell counting in adult mice. For this purpose, we first sampled brains of three adult C57BL/6N male mice (8 weeks old), which were then cleared, expanded, imaged and profiled using the CUBIC-X pipeline described above (Fig. 4k). As a result, the total cell count in the 8-week-old male mouse brain was found to be  $7.22 (\pm 0.52) \times 10^7$  (coefficient of variation = 7.14%). We also registered and annotated the detected individual cells to anatomical areas in the CUBIC-Atlas (Supplementary Fig. 17), which provided cell-number information for individual brain areas (Fig. 4l and Supplementary Table 3).

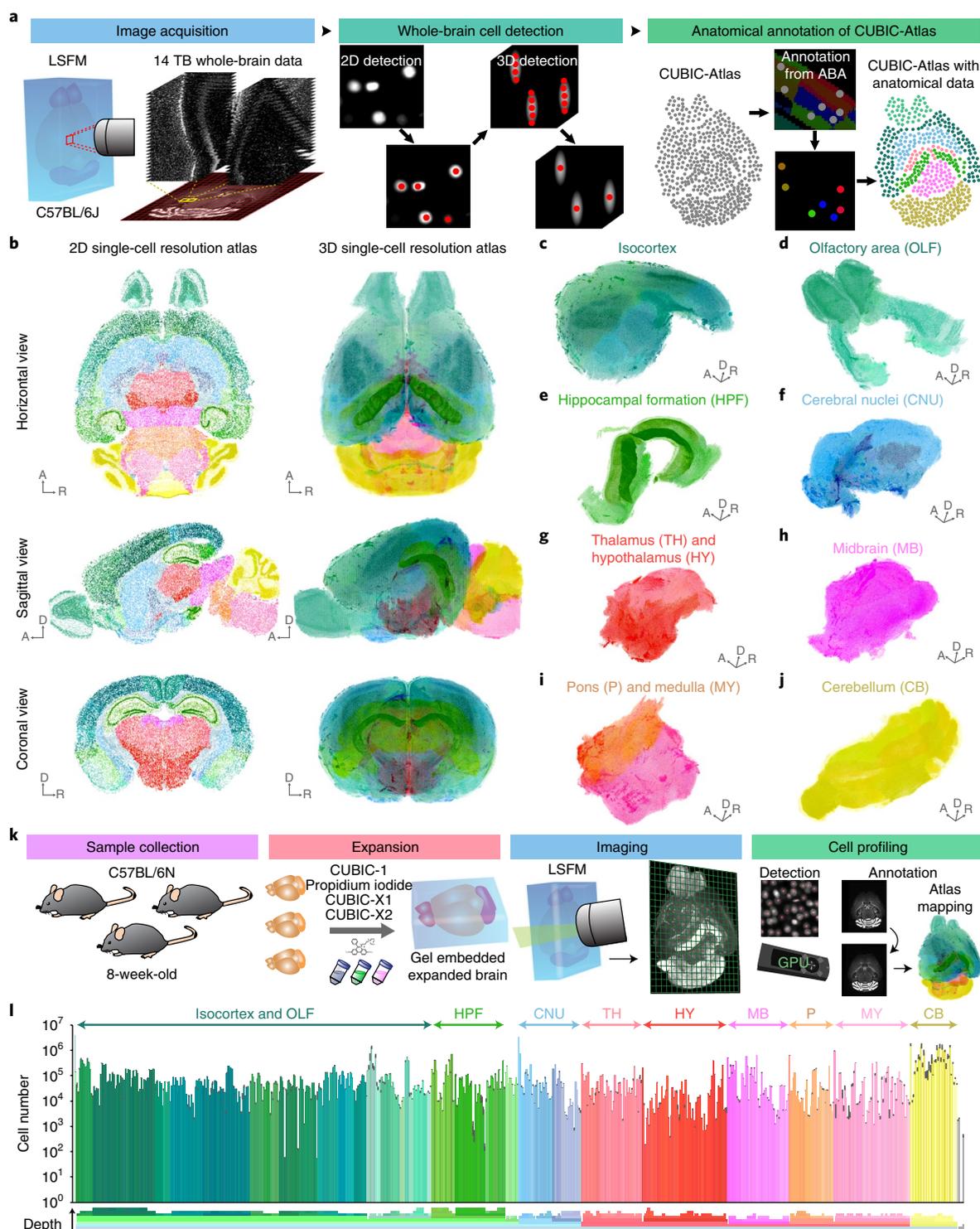
To comprehensively analyze regional progression of cell numbers during postnatal development, we also performed whole-brain cell profiling of 1-week-old, 3-week-old and 6-month-old C57BL/6N male mice according to the CUBIC-X pipeline (Fig. 5a). In this analysis, we focused on the brain areas excluding cerebellum and olfactory bulb, due to their morphological changes during early postnatal development (Supplementary Fig. 18). We first investigated whether anatomical areas of the adult brain atlas could be applied to datasets from younger ages. Automatic annotation of cortical layers based on adult brain atlas was achieved with high accuracy (Supplementary Fig. 19). Thus, we applied the anatomical annotations of adult (P56) mouse brain atlas to datasets from other ages and compared cell numbers in the corresponding areas. To characterize the developmental progression of cell numbers in different brain regions, we analyzed the average cell number and the s.d. over mice aged 1 week, 3 weeks, 8 weeks and 6 months in each area (Fig. 5b,c) and found that variations in developmental progressions over age depended strongly on brain area. Analysis of the coefficient of variation also supported high variability in developmental progression of cell numbers (Fig. 5d,e). To further investigate the developmental progression of cell numbers across the entire brain, we first calculated the normalized cell numbers of each area by setting the cell number averaged over ages to 1.0. To extract the characteristic patterns of developmental progression, we performed hierarchical clustering analysis of the normalized cell numbers (Fig. 5f and Methods). This clustering analysis revealed four major clusters (Fig. 5g,h). The largest cluster (blue cluster) spread over the entire brain (Fig. 5i). Two major regional clusters (pink and yellow) were relatively localized in cortex and olfactory area, respectively, and a fourth (purple) cluster was localized especially in midbrain (Fig. 5i).

We next focused on the differences of the normalized cell number between 1-week-old and 3-week-old mice, in which the normalized cell numbers in purple and yellow clusters markedly increased and decreased, respectively. We first analyzed large brain areas, such as cerebrum, interbrain, midbrain and hindbrain, and found that the midbrain displayed a larger increase in cell number than other areas (Fig. 6a), which was consistent with the enriched distribution of the purple cluster in the midbrain (Fig. 5h,i). We then mapped the normalized cell-number increase between ages 1 week and 3 weeks onto the midbrain, which revealed that the cell number increase was enriched in the anterior part of the midbrain (Fig. 6b,c). During the early postnatal period between 1 week and 3 weeks of age, the normalized cell number significantly increased in the motor-related areas and the behavioral-state-related areas, whereas the normalized cell number did not significantly change in the sensory-related areas (Fig. 6d). These findings in normalized cell numbers were also confirmed by substantial changes of actual cell numbers in these areas (Supplementary Fig. 20).

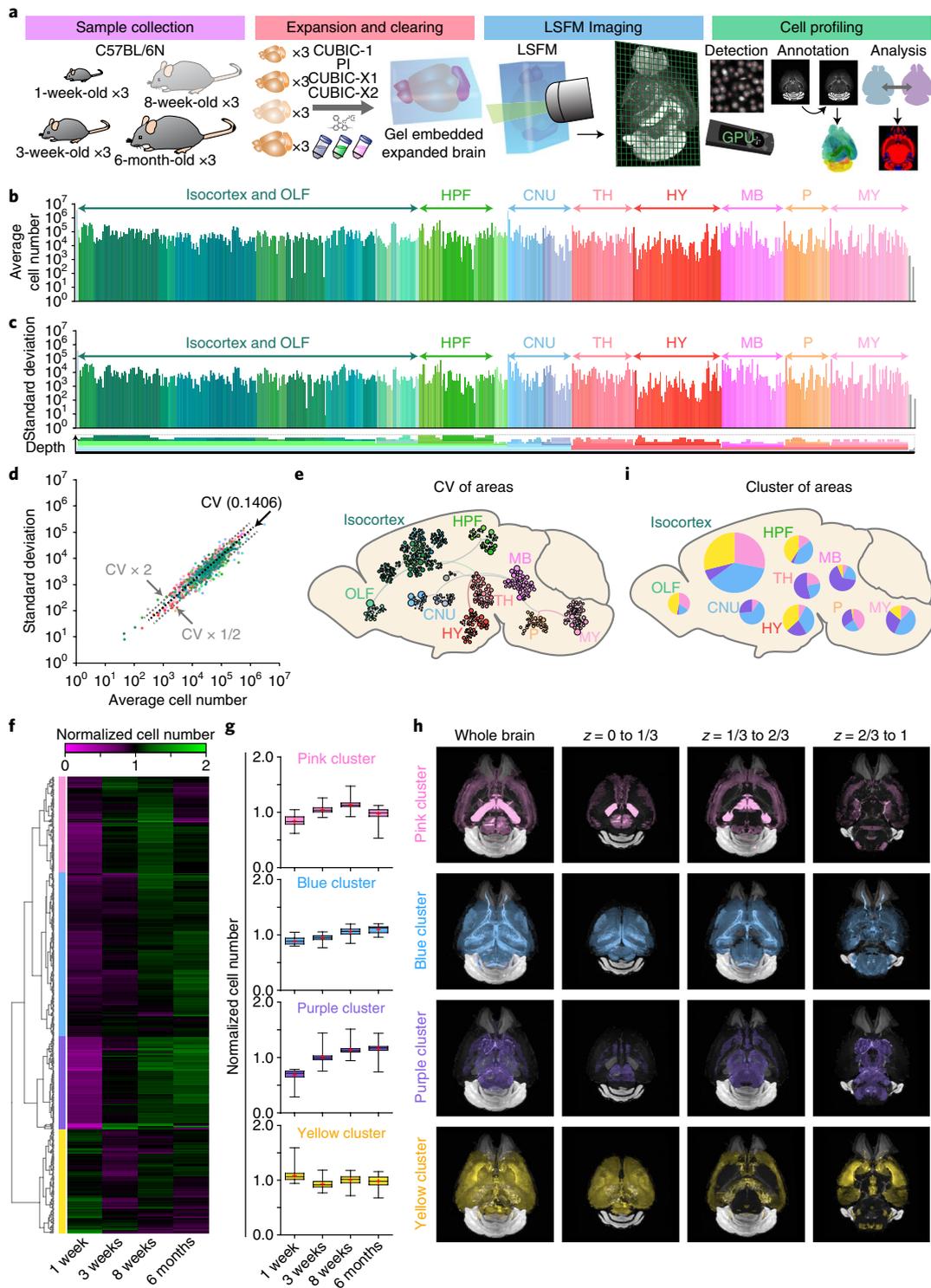
We also expected a decrease in normalized cell numbers in the cortex during the early postnatal period, based on the enriched distribution of the yellow cluster in the cortex (Fig. 5h,i). We focused on the cortical areas excluding layer 1 because cell numbers in this outermost area of the brain could be influenced by the residual arachnoid matter or possible damage during dissection procedures.



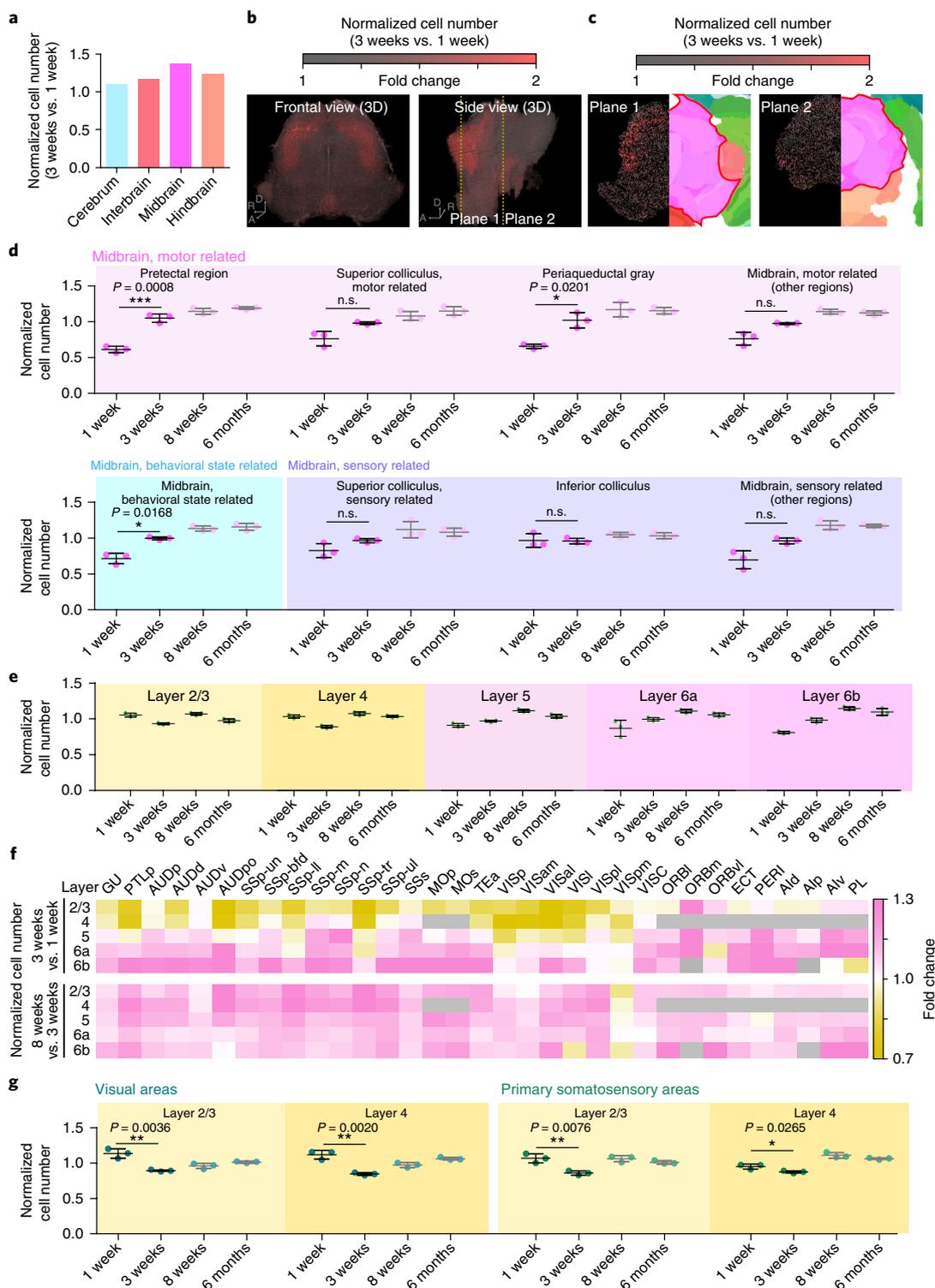
**Fig. 3 | Whole-brain imaging of a CUBIC-X expanded brain with customized LSFM.** **a**, Comparison of a propidium iodide (PI)-stained brain cleared with the modified ScaleCUBIC-2 (pre-CUBIC-X) and a brain expanded by the CUBIC-X treatment (post-CUBIC-X). Three regions (2.5 mm × 2.5 mm × 2.5 mm) of a pre-CUBIC-X brain are shown in the upper row from both horizontal and sagittal views. Corresponding images of a post-CUBIC-X brain, linearly registered to the pre-CUBIC-X brain images, are shown in the lower row. Scale bars, 500 μm, based on the pre-CUBIC-X brain. **b**, Quantification of the nonlinear distortion caused by CUBIC-X treatment. Root-mean-square errors were quantified. Its average (blue) and s.d. (light blue) are plotted against the measurement length. Sparsely sampled pairs of voxels were used (Methods). Note that all lengths are represented using a pre-CUBIC-X brain as baseline. **c**, Image of sample mounting in the customized LSFM. A CUBIC-X brain embedded in 2% agarose CUBIC-X2 is suspended in oil (RI = 1.467). **d**, Overview of whole-brain imaging procedures. **e**, Volume-rendered and single-plane images of a PI-stained CUBIC-X brain. A volume-rendered image of a whole brain is shown in the center. An original single-plane and its magnified (10×) images are shown for cortex, thalamus, hippocampus, midbrain, olfactory bulb, corpus callosum and cerebellum. Both horizontal (*x*-*y*) and coronal (*x*-*z*) views are shown. Experiments were repeated more than ten times with independent brains. Scale bars, 1,000 μm for whole-brain image and 50 μm for others, using a pre-CUBIC-X brain as baseline.



**Fig. 4 | Construction of a single-cell-resolution mouse brain atlas (CUBIC-Atlas).** **a**, Overview of construction of the CUBIC-Atlas. First, we acquired images of highly resolved cellular nuclei throughout the C57BL/6J male mouse brain (image acquisition). Second, we developed an algorithm to comprehensively detect individual nuclei (red dots) without omission or overlap (whole-brain cell detection) to construct the CUBIC-Atlas. We annotated the individual points of CUBIC-Atlas using anatomical annotations from ABA (anatomical annotation of CUBIC-Atlas). **b**, The CUBIC-Atlas. Horizontal, sagittal and coronal view of single-plane images (left) and volume-rendered images (right) of the CUBIC-Atlas. **c–j**, Major anatomical areas in the CUBIC-Atlas. For **b–j**, experiments and annotations were repeated more than ten times with independent brains. A, anterior; D, dorsal; R, rostral. **k**, Overview of whole-brain cell counting in C57BL/6N 8-week-old male mice. **l**, Cell numbers in each brain area, indicated as average  $\pm$  s.d. ( $n=3$ ). Colors represent anatomical areas: hue indicates parental anatomical structures; saturation indicates hierarchical depth within parental anatomical structures. OLF, olfactory areas; HPF, hippocampal formation; CNU, cerebral nuclei; TH, thalamus; HY, hypothalamus; MB, midbrain; P, pons; MY, medulla; CB, cerebellum.



**Fig. 5 | CUBIC-Atlas is applicable to whole-brain cell profiling of mice at various developmental timepoints. a**, Overview of whole-brain cell profiling over time in mice. In addition to the 8-week-old mouse brains described in Fig. 4f, the brain samples of 1-week-old, 3-week-old and 6-month-old C57BL/6N male mice were collected, expanded by CUBIC-X, imaged by the customized LSFM and analyzed with the CUBIC-Atlas. **b,c**, Average cell numbers and s.d. over mice aged 1 week, 3 weeks, 8 weeks and 6 months in each brain area ( $n=3$  mice per age group). Colors represent anatomical areas. Parental hierarchical structures are shown as in Fig. 4f. **d**, Correlations between the average cell numbers and s.d. **e**, Coefficient of variation (CV) values for cell numbers in each area. **f**, Heatmap for hierarchical clustering of normalized cell numbers in each brain area. Mean values of three brains are shown. Four major clusters (pink, blue, purple and yellow) were identified. **g**, Normalized cell numbers in each cluster. Plots were generated with normalized cell numbers from all individual areas of three brains for each age. Box plots indicate median and 25–75% interquartile ranges; whiskers cover full ranges of distributions; dots show averages and center lines show medians. **h**, 3D distribution of each cluster in the mouse brain. **i**, The ratios of clusters in each area. The size of each circle indicates the number of anatomical areas. OLF, olfactory areas; HPF, hippocampal formation; CNU, cerebral nuclei; TH, thalamus; HY, hypothalamus; MB, midbrain; P, pons; MY, medulla. Normalized cell numbers for all individual areas of twelve brains were subjected to clustering analysis (Methods).



**Fig. 6 | CUBIC-Atlas revealed a cell number increase in midbrain and decreases in visual area and primary somatosensory area of cerebral cortex during early postnatal development.** **a**, Normalized cell numbers in large brain areas. **b,c**, Fold-changes of the normalized cell number in the midbrain during early postnatal development (1 to 3 weeks). Fold-changes of the normalized cell number are shown in a volume-rendered image (**b**) and in single-plane images (indicated by yellow dashed lines) with the corresponding anatomical areas (**c**) of the midbrain. **d**, Increases in normalized cell numbers in the midbrain were significant ( $n=3$ ,  $***P < 0.001$ ,  $*P < 0.05$ , independent two-tailed  $t$  test). Background color represents each classification of the midbrain (pink, motor-related areas; blue, behavioral-state-related areas; purple, sensory-related areas). **e**, Normalized cell numbers in each layer of the cerebral cortex ( $n=3$ ). **f**, Fold-changes in the normalized cell number in the cerebral cortex during early postnatal period from age 1 week to 3 weeks (top) or during maturation period from 3 weeks to 8 weeks (bottom). To calculate fold-changes, mean values of normalized cell numbers for each age ( $n=3$ ) were used. **g**, Decreases in normalized cell numbers in visual area and primary somatosensory area of the cerebral cortex were significant ( $n=3$ ,  $***P < 0.01$ ,  $*P < 0.05$ , independent two-tailed  $t$  test). All values are mean  $\pm$  s.d.

As expected, layers 2/3 and 4 displayed a decrease in normalized cell numbers during the early postnatal period (Fig. 6e). We further analyzed the normalized cell number changes of individual cortical areas between ages 1 week and 3 weeks during the early postnatal period and between ages 3 weeks and 8 weeks during the maturation period (Fig. 6f). This whole-cortex cell profiling demonstrated a decrease in normalized cell numbers during the early postnatal period mainly in layers 2/3 and 4 of the gustatory area, posterior parietal association area, auditory area, somatosensory area, motor area, temporal association area and visual area. Since the visual and primary somatosensory areas have been intensively studied in relation to their critical periods<sup>35</sup>, we focused on layers 2/3 and 4 in these areas and analyzed the temporal patterns of normalized cell numbers in these areas (Fig. 6g), which demonstrated a substantial decrease in normalized cell numbers in layers 2/3 and 4 of these areas, as well as a decrease in actual cell numbers (Supplementary Fig. 21). Although a number of studies have been done on the decrease of synapses during early postnatal development in the cerebral cortex<sup>35,36</sup>, this observation was, to the best of our knowledge, the first report of a significant decrease of total cell numbers in layers 2/3 and 4 during early postnatal development.

**Probabilistic annotation of CUBIC-Atlas revealed a functionally distinct structure in the granule cell layer of hippocampal dentate gyrus.** To demonstrate the power of CUBIC-Atlas, we used it to identify functionally distinct brain areas. Because mammalian brains do not have identical cellular structures due to individual differences, we developed a probabilistic method to define areas within the CUBIC-Atlas (Fig. 7a and Methods). As its initial demonstration, we reanalyzed previously published brain image data of Arc-dVenus mice to investigate pharmacologically activated cells (high dVenus signal) in response to chronic administration of the NMDA-receptor inhibitor MK-801<sup>14</sup>. We first confirmed that the activated cells annotated with CUBIC-Atlas more frequently emerged in the cerebral cortex (Fig. 7b and Supplementary Fig. 22), which is consistent with our previous report. By assuming the Gaussian probabilistic distribution around the registered activated cells, we allocated probabilistic activity to each cell within CUBIC-Atlas, then performed *k*-means clustering analysis against all cells in the atlas. The results revealed four clusters distinctly emerging at each circadian time (Fig. 7c). Because the cells of Cluster 1 and Cluster 2 are in the granule cell layer of dentate gyrus (DG-sg; Fig. 7d,e), we further investigated distribution of the clusters over DG-sg. Mapping the cell cluster in the DG-sg revealed functional differences between upper DG-sg, where the majority of dorsal cell population belonged to Cluster 1, and lower DG-sg, where most of cells belonged to Cluster 2 (Fig. 7f,g), suggesting the applicability of CUBIC-Atlas to discovery of previously uncharacterized functionally distinct subareas. Notably, this clear separation of the clusters is consistent with the inhomogeneous characteristics of the hippocampus<sup>37</sup>. Discovery of functional areas is possible even without prior anatomical information. Based on these results, we propose a new approach for exploring and mapping of functionally distinct areas, using the single-cell-resolution atlas.

## Discussion

To clarify the underlying cellular mechanisms in complex neural circuits, an alternative atlas describing whole mouse brain as an ensemble of cellular points is a promising addition to current pixel- or voxel-based brain atlases. Herein we proposed a single-cell-resolution mouse brain atlas, CUBIC-Atlas. To this end, we developed a fluorescent-protein-compatible, intensive tissue-clearing method combined with a tissue expansion protocol, CUBIC-X, which enables seamless imaging of the whole mouse brain at subcellular resolution. Thanks to the high-resolution images, almost all cells in the whole brain could be abstracted by comprehensive and accurate

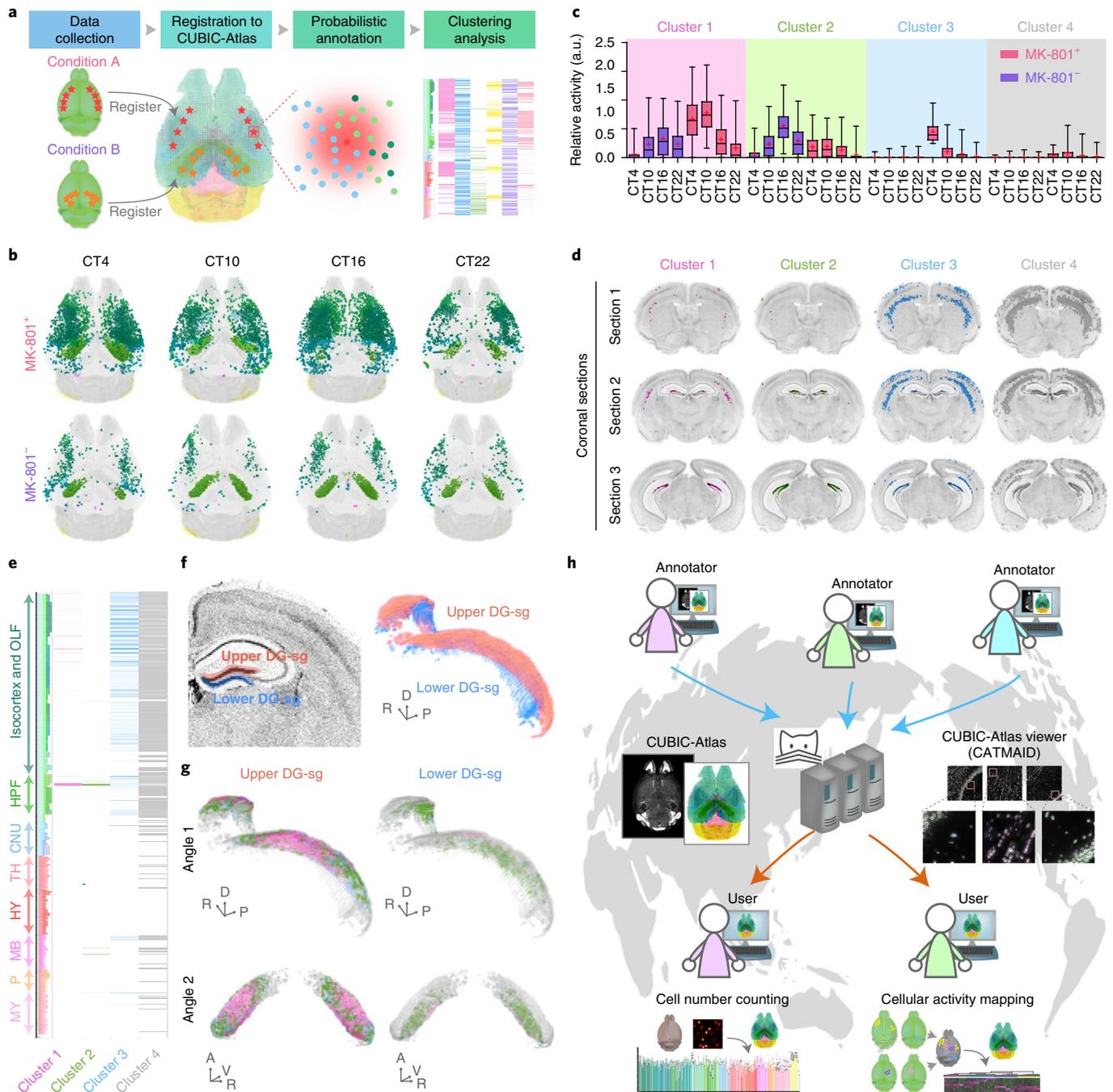
cell detection, and therefore huge datasets were reduced to portable and editable data size (~3 GB). Because CUBIC-Atlas was derived from seamless 3D images, our atlas would be suitable for analysis of volumetric images generated from cleared brains. The protocol may also play a pivotal role in global analysis of gene expression and neural circuits through continuous update of our CUBIC-Atlas by overlaying various cellular functions.

As a test application of the CUBIC-X pipeline and CUBIC-Atlas, we performed whole-brain cell counting. Quantification of cell numbers in the mammalian central nervous system has been one of the most fundamental and technical challenges in neuroanatomy<sup>38,39</sup>. For mice, total cell numbers in brains of various strains have been reported<sup>38,40–43</sup>, ranging from less than  $6.0 \times 10^7$  (ref. <sup>41</sup>) to more than  $15.0 \times 10^7$  (ref. <sup>43</sup>). This large variation in total cell numbers is assumed to be derived from genetic background, age, sex and individual differences. In our study, we reported the total cell numbers of C57BL/6N male mouse brain (at 8 weeks old) to be  $7.22 (\pm 0.52) \times 10^7$ . Our extensive evaluation of accuracy (Supplementary Fig. 13) supports this cell number as an appropriate estimate. We also note that our cell-detection algorithm tends to underestimate the cell number rather than overestimate especially in densely populated areas (Supplementary Fig. 13f); thus, we assume this total cell number is a lower-bound estimate. Further improvement of axial resolution of LSFM will improve accuracy in densely populated areas, such as the granular cell layer of the cerebellum. Since the CUBIC-X pipeline allows not only counting cell numbers but also identifying cellular locations, the application of the pipeline to pathological analysis is a good candidate for additional study. In some neurological disorders, such as Parkinson's disease and depression, neuronal death or dysfunction of neurogenesis in specific part of brains are observed<sup>44,45</sup>. Although local estimations of neuron numbers have been extensively performed, whole-brain comprehensive analysis of cell numbers of these disease model mice is yet to be achieved. Such analysis is promising application of CUBIC-Atlas and will provide insight to neurological disorders.

Our CUBIC-Atlas also revealed a significant decrease in cell numbers in layers 2/3 and 4 of visual and somatosensory areas during early postnatal development (Fig. 6). It has been reported that cell death of presynaptic neurons can be induced by deletion of their postsynaptic target<sup>46,47</sup>. In the visual cortex, the neuronal circuit can be reorganized by synaptic pruning mediated by protease from the extracellular matrix during postnatal critical periods<sup>35</sup>. Therefore, the cell number decreases observed in this study might reflect apoptosis of neuronal cells in those areas due to postsynaptic pruning. Other possible explanations for the decrease in cell numbers are early postnatal glial apoptosis and cortical cellular migration<sup>48</sup>. Since the CUBIC-X protocol is compatible with cellular labeling with fluorescent protein and fluorescent molecule, profiling cell types and tracking cellular migration during critical periods are potential applications of CUBIC-Atlas in future studies.

Another promising application of CUBIC-X protocol is imaging with super-resolution microscopes. As shown in Fig. 2f and Supplementary Fig. 10h, CUBIC-X expanded brains maintain the subcellular structures of neurons. Although the expansion ratio of CUBIC-X is smaller than those of previously reported expansion protocols<sup>25–27</sup>, the availability of fluorescent protein without special anchoring gel matrices or complicated immunostaining procedures could simplify experimental protocols. Notably, the expansion can be performed on a whole-organ scale, thereby allowing researchers to avoid possible disruptions of microscale structure or cellular circuits caused by tissue sectioning.

An unbiased analysis revealing previously unknown areas is one of the most attractive applications of CUBIC-Atlas. In this study, we confirmed the existence of novel subareas in the granule cell layer of DG by probabilistically mapping spatial information of pharmacologically activated cells (Fig. 7). This application not only



**Fig. 7 | Probabilistic annotation of CUBIC-Atlas revealed a functionally distinct structure in granule cell layer of hippocampal dentate gyrus.** **a**, Strategic schema of area discovery with probabilistically annotated CUBIC-Atlas. **b**, Temporal variation in distribution of hyperactivated cells in the whole brain. Original image data were obtained in our previous study<sup>14</sup> and reanalyzed in this study. Arc-dVenus-expressing cells in mice with or without chronic administration of MK-801 are shown. Colors indicate the anatomical area of the detected cells. Eight representative brains are shown. All other brains are shown in Supplementary Fig. 22. **c**, Temporal variation of relative cellular activity in each cluster with or without administration of MK-801. Box plots indicate median and 25–75% interquartile ranges; whiskers indicate second and 98th percentiles; dots show averages and center lines show medians. **d, e**, Spatial distribution of each cluster. **(d)** Each cluster was mapped onto the virtual coronal section of the mouse brain. **(e)** Heat-map view of the cluster distribution with parental hierarchical structures. **f**, Overview of upper and lower DG-sg. Coronal slice view (left) and 3D-reconstructed view (right) are shown. **g**, 3D view of spatial distribution of four clusters in DG-sg. Images are shown from two different angles after volume rendering. When clusters on upper DG-sg are shown, lower DG-sg is shown as gray (left). When clusters on lower DG-sg are shown, upper DG-sg is shown as gray (right). R, rostral; P, posterior; A anterior; D, dorsal; V, ventral. For **c–e, g**, clustering analysis was performed with 20 brains (MK-801<sup>+</sup>,  $n = 3 \times 4$  timepoints, MK-801<sup>-</sup>,  $n = 2 \times 4$  timepoints). **h**, Concept for the CUBIC-Atlas data-sharing platform via an open-source web-based viewer and editor, CATMAID.

demonstrates the feasibility of activity mapping with CUBIC-Atlas but also proposes an approach to redefining anatomical areas. This bottom-up approach to build a mammalian brain atlas is synergistic

with the previously reported histologically annotated atlas. In our study, both information of probabilistic annotation (spatial information of Arc-dVenus expressing cells) and the prior annotation

(DG-sg) lead to the discovery of a functionally distinct subarea in DG-sg. The utilization of analytic tools to handle point-cloud data, such as a point-cloud library<sup>49</sup>, will further assist cellular description of the mouse brain.

Since whole-organ imaging technologies are developing rapidly, our CUBIC-Atlas would contribute to this field by making this single-cell-resolution atlas publicly available. Therefore, we established a universal data-sharing platform for CUBIC-Atlas (<http://cubic-atlas.riken.jp>) via an open-source web-based viewer and annotation editor, the Collaborative Annotation Toolkit for Massive Amounts of Image Data (CATMAID)<sup>50</sup> (Fig. 7h). In CATMAID, a CUBIC-Atlas viewer and editor, annotators can assemble novel annotations such as gene expression profiling, phenotyping of individual cells and functional connections of individual neural circuits. We put our CUBIC-Atlas data and source codes on our webpage (<http://cubic-atlas.riken.jp>) to perform cell detection as well as to map and annotate the detected cells onto CUBIC-Atlas. We thus believe our CUBIC-Atlas will become a useful platform for 3D brain mapping and for the digitalization and decoding of the mammalian brain.

## Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41593-018-0109-1>.

Received: 26 July 2017; Accepted: 6 February 2018;

Published online: 5 March 2018

## References

- Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
- Mikula, S., Trotts, I., Stone, J. M. & Jones, E. G. Internet-enabled high-resolution brain mapping and virtual microscopy. *Neuroimage* **35**, 9–15 (2007).
- Amunts, K. et al. BigBrain: an ultrahigh-resolution 3D human brain model. *Science* **340**, 1472–1475 (2013).
- Calabrese, E. et al. A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. *Neuroimage* **117**, 408–416 (2015).
- Rohlfing, T. et al. The INIA19 template and NeuroMaps Atlas for primate brain image parcellation and spatial normalization. *Front. Neuroinform.* **6**, 27 (2012).
- Johnson, G. A. et al. Waxholm space: an image-based reference for coordinating mouse brain research. *Neuroimage* **53**, 365–372 (2010).
- Papp, E. A., Leergaard, T. B., Calabrese, E., Johnson, G. A. & Bjaalie, J. G. Waxholm Space atlas of the Sprague Dawley rat brain. *Neuroimage* **97**, 374–386 (2014).
- Dong, H. W. *Allen Reference Atlas. A Digital Color Brain Atlas of the C57BL/6J Male Mouse* (Wiley, Hoboken, NJ, USA, 2008).
- Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
- Kasukawa, T. et al. Quantitative expression profile of distinct functional regions in the adult mouse brain. *PLoS One* **6**, e23228 (2011).
- Okamura-Oho, Y. et al. Transcriptome tomography for brain analysis in the web-accessible anatomical space. *PLoS One* **7**, e45373 (2012).
- Kim, Y. et al. Mapping social behavior-induced brain activation at cellular resolution in the mouse. *Cell Rep.* **10**, 292–305 (2015).
- Vousden, D. A. et al. Whole-brain mapping of behaviourally induced neural activation in mice. *Brain Struct. Funct.* **220**, 2043–2057 (2015).
- Tatsuki, F. et al. Involvement of Ca<sup>2+</sup>-dependent hyperpolarization in sleep duration in mammals. *Neuron* **90**, 70–85 (2016).
- Oh, S. W. et al. A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
- Menegas, W. et al. Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. *eLife* **4**, e10032 (2015).
- Vélez-Fort, M. et al. The stimulus selectivity and connectivity of layer six principal cells reveals cortical microcircuits underlying visual processing. *Neuron* **83**, 1431–1443 (2014).
- Dotd, H. U. et al. Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nat. Methods* **4**, 331–336 (2007).
- Ertürk, A. et al. Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nat. Protoc.* **7**, 1983–1995 (2012).
- Pan, C. et al. Shrinkage-mediated imaging of entire organs and organisms using uDISCO. *Nat. Methods* **13**, 859–867 (2016).
- Chung, K. et al. Structural and molecular interrogation of intact biological systems. *Nature* **497**, 332–337 (2013).
- Yang, B. et al. Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* **158**, 945–958 (2014).
- Susaki, E. A. et al. Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* **157**, 726–739 (2014).
- Susaki, E. A. & Ueda, H. R. Whole-body and whole-organ clearing and imaging techniques with single-cell resolution: toward organism-level systems biology in mammals. *Cell Chem. Biol.* **23**, 137–157 (2016).
- Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science* **347**, 543–548 (2015).
- Treweek, J. B. et al. Whole-body tissue stabilization and selective extractions via tissue-hydrogel hybrids for high-resolution intact circuit mapping and phenotyping. *Nat. Protoc.* **10**, 1860–1896 (2015).
- Ku, T. et al. Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nat. Biotechnol.* **34**, 973–981 (2016).
- Chang, J. B. et al. Iterative expansion microscopy. *Nat. Methods* **14**, 593–599 (2017).
- Tillberg, P. W. et al. Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* **34**, 987–992 (2016).
- Hama, H. et al. Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. *Nat. Neurosci.* **14**, 1481–1488 (2011).
- Tainaka, K., Kuno, A., Kubota, S. I., Murakami, T. & Ueda, H. R. Chemical principles in tissue clearing and staining protocols for whole-body cell profiling. *Annu. Rev. Cell Dev. Biol.* **32**, 713–741 (2016).
- Keller, P. J., Schmidt, A. D., Wittbrodt, J. & Stelzer, E. H. K. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* **322**, 1065–1069 (2008).
- Shimada, T., Kato, K., Kamikouchi, A. & Ito, K. Analysis of the distribution of the brain cells of the fruit fly by an automatic cell counting algorithm. *Physica A* **350**, 144–149 (2005).
- Feng, G. et al. Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41–51 (2000).
- Hensch, T. K. Critical period plasticity in local cortical circuits. *Nat. Rev. Neurosci.* **6**, 877–888 (2005).
- Chechik, G., Meilijson, I. & Ruppén, E. Synaptic pruning in development: a computational account. *Neural Comput.* **10**, 1759–1777 (1998).
- Fanselow, M. S. & Dong, H. W. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* **65**, 7–19 (2010).
- Herculano-Houzel, S. et al. Updated neuronal scaling rules for the brains of Glires (rodents/lagomorphs). *Brain Behav. Evol.* **78**, 302–314 (2011).
- von Bartheld, C. S., Bahney, J. & Herculano-Houzel, S. The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *J. Comp. Neurol.* **524**, 3865–3895 (2016).
- Herculano-Houzel, S., Mota, B. & Lent, R. Cellular scaling rules for rodent brains. *Proc. Natl. Acad. Sci. USA* **103**, 12138–12143 (2006).
- Herculano-Houzel, S., Messeder, D. J., Fonseca-Azevedo, K. & Pantoja, N. A. When larger brains do not have more neurons: increased numbers of cells are compensated by decreased average cell size across mouse individuals. *Front. Neuroanat.* **9**, 64 (2015).
- Williams, R. W. Mapping genes that modulate brain development: a quantitative genetic approach. in: *Mouse Brain Development. Results and Problems in Cell Differentiation* (eds. A. M. Goffinet & P. Rakic) 21–49. Springer Verlag, Berlin, 2000).
- Seiriki, K. et al. High-Speed and scalable whole-brain imaging in rodents and primates. *Neuron* **94**, 1085–1100.e6 (2017).
- Dauer, W. & Przedborski, S. Parkinson's disease: mechanisms and models. *Neuron* **39**, 889–909 (2003).
- Sahay, A. & Hen, R. Adult hippocampal neurogenesis in depression. *Nat. Neurosci.* **10**, 1110–1115 (2007).
- Cowan, W. M., Fawcett, J. W., O'Leary, D. D. & Stanfield, B. B. Regressive events in neurogenesis. *Science* **225**, 1258–1265 (1984).
- Vanderhaeghen, P. & Cheng, H. J. Guidance molecules in axon pruning and cell death. *Cold Spring Harb. Perspect. Biol.* **2**, a001859 (2010).
- Rice, D. & Barone, S. Jr. Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environ. Health Perspect.* **108**, 511–533 (2000). Suppl 3.
- Rusu, R.B. & Cousins, S. 3D is here: Point Cloud Library (PCL). *IEEE Int. Conf. Robot.* <https://doi.org/10.1109/ICRA.2011.5980567> (2011).
- Saalfeld, S., Cardona, A., Hartenstein, V. & Tomancak, P. CATMAID: collaborative annotation toolkit for massive amounts of image data. *Bioinformatics* **25**, 1984–1986 (2009).

## Acknowledgements

We thank all lab members at The University of Tokyo and RIKEN QBiC, in particular: A. Millius and W. Kylius for editing, E.A. Susaki and A. Kuno for discovering the

tissue-swelling phenomenon, S. Shoi for helping with statistical analysis, K. Yoshida for helping with the decomposing transformation matrix, and C. Shimizu for supporting swelling experiments. We also thank H. Hayakawa and S. Jiang for supporting the preparation of a C57BL/6J mouse brain, D. Perrin and H. Yukinaga for informatics instruction, J. Kaneshiro, T. Watanabe and Olympus Engineering for helping design the microscope, T. Mitani and K. Matsumoto for reproducibility confirmation, S. Takano, S. Yamazoe and T. Tsukuda for the measurements of zeta potentials and Bitplane for instruction of Imaris 8.1.2. This work was supported by a grant from AMED-CREST (AMED/MEXT, grant number JP17gm0610006, to H.R.U.), CREST (JST/MEXT, to H.R.U.), Brain/MINDS (AMED/MEXT, grant number JP17dm0207049, to H.R.U. and H. M.), Basic Science and Platform Technology Program for Innovative Biological Medicine (AMED/MEXT, grant number JP17am0301025, to H.R.U.), Translational Research Network Program from Japan Agency for Medical Research and development (AMED, to H.M.), World Premier International Research Center Initiative (MEXT, to H.R.U.), a Grant-in-Aid for Scientific Research (JSPS KAKENHI, grant number 16J05041, to T.C.M.), a Grant-in-Aid for Scientific Research (S) (JSPS KAKENHI, grant number 25221004, to H.R.U.), Grant-in-Aid for Challenging Exploratory Research (JSPS KAKENHI, grant number 16K15124, to K.T.), a Grant-in-Aid for Scientific Research on Innovative Areas (JSPS KAKENHI, grant number 23115006, to H.R.U., 15H01558, to H.M., 17H05688, to K.T.), and a Grant-in-Aid from the Naito Foundation (to K.T.).

### Author contributions

H.R.U., T.C.M., T.M. and K.T. designed the study. T.C.M., T.M. and S.S. performed most of the experiments. S.A.H. contributed to CATMAID data sharing. D.S. designed CAD. K.B. and H.M. prepared C57BL/6 mouse brains. H.S., M.I. and K.F.T. produced PLP-YFP and Mlc1-YFP mice. Y.S. produced recombinant fluorescent proteins. H.K. produced R26-H2B-EGFP mice. H.R.U., T.C.M., T.M. and K.T. wrote the manuscript. All authors discussed the results and commented on the manuscript text.

### Competing interests

T.C.M., K.T. and H.R.U. have filed patent application for the CUBIC-X technique. Part of this study was done in collaboration with Olympus Corporation.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41593-018-0109-1>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to H.R.U.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Methods

**Mice.** We used 6-month-old Thy1-YHP-H transgenic female mice<sup>34</sup>, 10-week-old PLP-YFP (PLP-*tTA::tetO-ChR2EYFP*) transgenic male mice<sup>51</sup> and 11-week-old Mcl1-YFP (Mcl1-*tTA<sup>32</sup>::tetO-YC<sup>53</sup>*) transgenic female mice for whole-brain imaging of fluorescent reporter proteins. We used 4-month-old R26-H2B-EGFP male mice<sup>54</sup> (accession number CDB0238K; [http://www.clst.riken.jp/arg/reporter\\_mice.html](http://www.clst.riken.jp/arg/reporter_mice.html)) to observe the influence of delipidation over nuclei, and used 8-week-old C57BL/6J male mice to build the whole-brain atlas with single-cell resolution. C57BL/6N mice were used to analyze whole-brain cell profiling over various ages. One-week-old male mice were purchased from CLEA Japan, and 3-week-, 8-week- and 6-month-old male mice were purchased from Japan SLC, Inc.

For sampling, the mice were deeply anesthetized with pentobarbital. Then the subject was transcardially perfused with PBS to flush the blood vessels, followed by perfusion with 4% paraformaldehyde (PFA) in PBS (pH 7.4) for fixation. Each excised brain was postfixed with the same fixation solution at 4°C overnight. The fixed brain was washed with PBS for several hours to remove PFA. If necessary, the postfixed samples were immersed in 20% sucrose in PBS at 4°C and then stocked in O.C.T. compound at -80°C until use.

All experimental procedures and housing conditions were approved by the Animal Care and Use Committee of the Graduate School of Medicine, the University of Tokyo, or by the Animal Care and Use Committee of the Graduate School of Medicine, Osaka University, and all of the animals were cared for and treated humanely in accordance with the Institutional Guidelines for Experiments using animals.

**Delipidation of brain sample.** For Thy1-YHP-H, PLP-YFP and Mcl1-YFP mouse brains, a modified ScaleCUBIC-1 reagent (5 wt% *N,N,N',N'*-tetrakis(2-hydroxypropyl)ethylenediamine (Quadrol, Tokyo Chemical Industry, T0781), 25 wt% urea (Nacalai Tesque, 35904-45) and 15 wt% Triton X-100 (Nacalai Tesque, 12967-45)) was applied to suppress quenching of fluorescent proteins. For other brains, ScaleCUBIC-1 reagent (25 wt% Quadrol, 25 wt% urea, 15 wt% Triton X-100) was applied<sup>23</sup>. Brain samples were initially immersed in delipidation reagent, diluted by half with water, at 37°C with gentle shaking for 3 h. Then brain samples from Thy1-YHP-H, PLP-YFP and Mcl1-YFP mice were immersed in modified ScaleCUBIC-1 at 37°C with gentle shaking for 8 d. Other brain samples were treated with ScaleCUBIC-1 at 37°C for 5–14 d. Because we noticed that long-term delipidation made the young (less than 3-week-old) mouse brain too soft to handle properly, we used different duration of delipidation depending on age: 5 d for 1-week-old mice, 7 d for 3-week-old mice and 14 d for 8-week-old and 6-month-old mice. The delipidation cocktails were refreshed every 4 d. After delipidation, the brains were thoroughly washed with PBS at room temperature (20°C to 30°C) with gentle shaking overnight.

**Size measurements of gelatin blocks and brains.** We used gelatin type B (Nacalai Tesque, 16631-05) in gelatin-based assays (Fig. 1 and Supplementary Fig. 1). To prepare gelatin blocks, 10% gelatin aqueous solution was prepared by gentle heating until the gelatin powder dissolved, and 1 mL of the solution was dispensed into square dies. Gelatin was allowed to set at 4°C overnight. After gelation, resulting gelatin blocks were fixed by 4% PFA in PBS at 4°C for more than 12 h with gentle shaking and washed with PBS extensively. Gelatin blocks were treated with 10 mL of each chemical at room temperature with gentle shaking for 1 d. We measured the 2D projection area (mm<sup>2</sup>) of the gelatin blocks using ChemiDoc XRS Plus (Bio-Rad) and estimated the volume assuming the gelatin blocks were isotopically swelled or shrunk by each chemical treatment.

To measure the brain swelling ratio (Fig. 1a,e and Supplementary Figs. 1 and 2), brain samples of 8-week-old mice were initially delipidated by ScaleCUBIC-1 as mentioned above. Delipidated brains were cut to make hemisphere brains. Hemisphere brains were treated with 10 mL of each chemical at room temperature with gentle shaking for 1 d. We measured the 2D projection area (mm<sup>2</sup>) of the hemisphere brains using ChemiDoc XRS Plus (Bio-Rad) and estimated the volume, assuming that hemisphere brains were isotopically swelled and shrunk by each chemical treatment. The volume of each hemisphere brain was calculated by simply taking the 3/2 power of the 2D projection area. Custom-written Python code was used to quantify the 2D projection area. Initially, a Sobel filter was applied to give sufficient contrast to the edges, followed by manually adjusting a threshold to determine the contour of the brains. Finally, we measured inner area as a 2D projection area. The swelling ratio was calculated as the ratio of the obtained volume between brain samples before and after chemical treatment.

We measured the 2D projection areas of whole brains instead of hemisphere brains and evaluated the time-courses of swelling ratios of whole-brain samples in Fig. 2c. Volumes were estimated as described above. In this Fig. 2c, we used the size of PFA-fixed whole brain as the original (baseline) brain size.

**Measurement of zeta potential.** Because isoelectric points of gelatin are different depending on the manufacturing process, we measured both gelatin type A (MP Biomedicals, 901771) and gelatin type B (Nacalai Tesque, 16631-05), which are processed under acidic and basic conditions, respectively. We prepared a 0.1% concentration of gelatin in water, which we then fixed by 0.8% PFA in water at 4°C overnight. As a control, we used 0.1% lysozyme aqueous solution (Wako Pure

Chemical Industries Ltd., 122-02673). Brain protein was prepared as follows. Adult mouse brains were removed from the skull, chopped with a blade, suspended and sonicated in extraction buffer (20 mM HEPES, 100 mM NaCl, 0.01 M DTT, and protease inhibitors (1 × cOmplete, EDTA-free Protease Inhibitor Cocktail, Roche, 05056489001)). We centrifuged 20 mL of the resulting brain suspension (from 10 brain samples) at 12,000 g for 15 min, and 1 mL of the supernatant was mixed with 14 mL of cold 80% (v/v) acetone and 20% (v/v) methanol solution. The suspension was incubated at 4°C for 90 min. After centrifugation, the pellet was washed with 20 mL of acetone and adequate volume of methanol, and then dried at 37°C over 30 min. The pellet was fixed with 2 mL of 4% PFA in PBS at 4°C with gentle shaking overnight. After washing in PBS, the pellet was stored in 2 mL of stock solution (20 mM HEPES, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% Na<sub>2</sub>S<sub>2</sub>O<sub>8</sub> and protease inhibitors; pH 7.4). The fixed brain protein was diluted to 1/1,000 in water just before zeta potential measurement. Zeta potential was measured with ZetaSizer Nano ZS (Malvern). pH was adjusted by titrating 0.05 M NaOH or 0.05 M HCl to samples. The measured zeta potentials were plotted in Supplementary Fig. 1 and fitted to fourth-order polynomials. Isoelectric points were determined by calculating a crossing point of the polynomial with the x-axis (Fig. 1b).

**High-throughput screening of swelling reagents.** For high-throughput screening of swelling reagents (Fig. 1c,d), we measured 975-nm absorption of gelatin type B (Nacalai Tesque, 16631-05). We used 975-nm absorption because water absorption of near infrared shows a peak around 975 nm<sup>55</sup>, and this peak is widely used to determine the optical path-length of water<sup>56,57</sup>. We prepared a mixture of 9% gelatin type B and 0.8% PFA; this low concentration of PFA was used because hard fixation by 4% PFA caused less swelling of gelatin, hindering the accurate quantification of chemical swelling abilities. We dispensed 100 µL of this mixture immediately into each well of the 96-well plate. The plates were incubated at 4°C overnight. After gelation, gelatin plates were washed with water. For calibration, 100 µL of water was put onto the gelatin, and the 975-nm absorbance was measured with PowerWave XS (Bio-Tek). After removal of the water, 170 µL of 10% (w/w) chemicals were put onto the gelatin, and bubbles were carefully removed. The plates were sealed and incubated at room temperature for 3 d. After incubation, the supernatants were thoroughly removed by washing the plates with water three times. Then we added 100 µL of water to the gelatin to avoid surface scattering and refraction. The 975-nm absorbance was measured again with PowerWave XS. Because some chemicals produced precipitation in or on gelatin during incubation, we also measured the 900-nm absorbance as an index of scattering. The same experiments were performed twice each. We excluded chemicals that showed more than 0.1 absorbance at 900-nm. We linearly normalized the values of 975-nm absorbance of the gel by the value of blank (0.0) and urea-treated gel (1.0). Then we converted the 975-nm absorbance to in-well gelatin swelling scores (Fig. 1d).

**Evaluation of hyperhydrative RI matching reagent.** Brain samples of 8-week-old mice were delipidated by ScaleCUBIC-1 as described above. After delipidation, brain samples were washed with PBS and then cut to make hemisphere brains. Each hemisphere brain was immersed in 10 mL of CUBIC-X1 (20% imidazole (Tokyo Chemical Industry, I0001)) at room temperature with gentle shaking for 1 d. CUBIC-X1-treated brains were subsequently subjected to 10 mL of RI media at room temperature with gentle shaking for 1 d. We used 60 wt% Histodenz (RI = 1.467; Sigma-Aldrich, D2158)<sup>22,28</sup>, 75 wt% fructose (RI = 1.472; Nacalai Tesque, 16315-55), 100 wt% glycerol (RI = 1.473; Nacalai Tesque, 17018-25), 75 wt% TDE (RI = 1.469; 2,2'-thiodiethanol, Tokyo Chemical Industry, T0202)<sup>59,60</sup>, ScaleA2 (RI = 1.382; 4 M urea, 10% (w/v) glycerol, and 0.1% (w/v) Triton X-100)<sup>30</sup> and ScaleCUBIC-2 (RI = 1.485; 50 wt% sucrose, 25 wt% urea, 10 wt% triethanolamine)<sup>23</sup> in addition to imidazole/antipyrine cocktails. We measured the swelling ratio of the resulting brain samples using ChemiDoc XRS Plus (Supplementary Fig. 2). In Supplementary Fig. 2b, RI values of imidazole/antipyrine cocktails were measured with an Abbe refractometer (ATAGO, DR-A1, Japan).

We captured transmission images of swollen brain samples (Fig. 2a). These samples were prepared as follows. Delipidated brain samples were immersed in 30 mL of CUBIC-X1 at room temperature with gentle shaking for 3 d. CUBIC-X1 was refreshed every 12 h. Then these brains were immersed in 40 mL of each chemical cocktail at room temperature with gentle shaking for 2 d. Each reagent was refreshed daily.

**CUBIC-X protocol.** The brain expansion protocol is composed of the following major steps: delipidation, nuclear staining, swelling, RI matching and gel embedding (Fig. 2d). Delipidation is described above. For nuclear staining, we immersed a delipidated brain sample in 3 mL of 30 µg/mL propidium iodide (Life Technologies, P21493) and 1.5 M NaCl in PBS and incubated it at room temperature with gentle shaking for 3 d. The brain sample was washed with PBS at room temperature with gentle shaking for more than 12 h. Then we immersed the sample in 30 mL of 20% imidazole and incubated it at 4°C (for Thy1-YFP-H, PLP-YFP and Mcl1-YFP brains) or at room temperature (for other brains) with gentle shaking for 2.5 d. Note that we incubated brain samples expressing fluorescent proteins at the lower temperature to avoid possible quenching. Since the expanded brains were fragile, careful handling was required for all following procedures. Each expanded brain was immersed in 40 mL of CUBIC-X2 (5% (w/v) imidazole

and 55% (w/v) antipyrine (Tokyo Chemical Industry, D1876) cocktail) with gentle shaking at room temperature for 1.5 d. The cocktail was refreshed every 12 h. For gel embedding, well-cleared swollen brains were embedded in CUBIC-X2 with 2% agarose (Wako Pure Chemical Industries Ltd., 318-01195). Prior to gel embedding, agarose powder was dissolved in CUBIC-X2 by heating, and then we cooled the mixture to a moderate temperature. Gelation was performed at 4°C overnight. The gel-embedded samples were then dissected for imaging.

**Measurement of light transmittance and scattering.** For Fig. 2b, we measured the light transmittance of the brain samples used in Fig. 2a, from 380 to 780 nm at 5-nm intervals with an integrating sphere (Spectral Haze Meter SH 7000, Nippon Denshoku Industries). Every sample was tested in 2-cm optical path-length cells with each chemical solution. For Supplementary Fig. 3, we measured the light transmittance of water, 2% agarose, CUBIC-X2, 2% agarose CUBIC-X2 and CUBIC-X brain in 2% agarose CUBIC-X2. We calculated the transmission value per centimeter by the square root of the total transmission value at each wavelength.

To measure scattering, we referred to the method described by Schwarz et al.<sup>61</sup>. Images were taken with a macro-zoom microscope (MVX10, Olympus) with an SDFPLAPO1.6XPF (NA = 0.24, Olympus) objective lens. The brains were immersed in ScaleA2 or CUBIC-X2, in glass cases placed on top of a light source with a positive 1951 USAF resolution value (Edmund Optics). CUBIC-X2-treated brains were embedded in gel with 2% agarose before imaging. For quantification, minimum and maximum intensity over element 2, group 2 of the USAF resolution chart was used. The absolute differences between the minimum and the maximum were computed, and the local contrast was calculated by normalizing it to blank controls (ScaleA2 reagent or 2% agarose CUBIC-X2 in CUBIC-X2).

**Evaluation of application of CUBIC-X to fluorescent protein with confocal microscopy.** Purified recombinant EGFP, EYFP and mCherry were used to evaluate quenching of fluorescent proteins by various chemical treatments. Each chemical solution was mixed with EGFP (6.7 µg/mL final concentration), EYFP (8.5 µg/mL final concentration) or mCherry (200 µg/mL final concentration). We measured the fluorescence intensity of each sample with Enspire (Perkin Elmer, USA; Supplementary Fig. 4). To evaluate the availability of CUBIC-X protocol to brain samples expressing fluorescent reporter proteins, we acquired images of the hippocampal area of Thy1-YFP-H mouse brains from the ventral side at 2-µm steps using an upright confocal microscope (FV1200, Olympus) equipped with a 25× objective lens (XLPLN25XGMP, NA = 1.0, WD = 8.0 mm, Olympus). A 473-nm laser was used for excitation. Acquired images were volume-rendered and shown in Fig. 2e.

**Evaluation of distortion during CUBIC-X protocol.** To assess nonlinear deformation of chemically treated brains, we attempted to prepare a transparent brain preserving the original shape as a standard. After delipidation and PI staining, brains were washed with 30 mL of PBS at room temperature with gentle shaking for 12 h, and then immersed in 10 mL of modified ScaleCUBIC-2 (22.5 wt% antipyrine, 22.5 wt% sucrose, 25 wt% urea, 10 wt% triethanolamine (Wako Pure Chemical Industries Ltd., 145-05605) and 80 mM NaCl), diluted by half, at 37°C with gentle shaking overnight. Then the brains were further immersed in 5 mL of modified ScaleCUBIC-2 at 37°C with gentle shaking for 2 d. The modified ScaleCUBIC-2 was refreshed every day. After image acquisition, the brains were washed with 30 mL of PBS at room temperature with gentle shaking overnight, and then immersed in 30 mL of ScaleCUBIC-1 at 37°C with gentle shaking for 1 d. After washing in PBS, samples were expanded with CUBIC-X1, followed by RI matching with CUBIC-X2 and gel embedding as described above.

To acquire images, we used a light-sheet fluorescent microscope (developed by Olympus). The pre-CUBIC-X brain sample was immersed in a mixture (RI = 1.490) of silicon oil TSF4300 (Momentive Performance Materials, RI = 1.498) and mineral oil (Sigma-Aldrich, RI = 1.467), and the post-CUBIC-X brain sample was in mineral oil (Sigma-Aldrich, RI = 1.467) during image acquisition. Images were captured with a 0.63× objective lens (NA = 0.15, WD = 87 mm) with 2× optical zoom for pre-CUBIC-X brains or 1× zoom for post-CUBIC-X brains. A 532-nm laser was used for image acquisition. To cover whole brains, a stage was designed to move in both lateral and axial directions. When the stage was moved axially, the detection objective lens moved in synchrony with it to avoid defocusing.

We measured the volumes of pre- and post-CUBIC-X brains by segmenting the brains from acquired images. Segmentation was performed by combining the binarization and morphological opening of the images. The thresholds for the binarization and the radius of morphological openings were manually determined. Voxel numbers of segmented brain volumes were counted, then converted into mm<sup>3</sup> (Supplementary Fig. 5).

For nonlinear distortion analysis, postexpansion images were registered to pre-expansion images using affine transformation and then nonlinear transformation. We used ANTs (Advanced Normalization Tools)<sup>62</sup> as a registration tool because it has proved successful in a number of image registration tasks<sup>63</sup>. Affine transformation was performed with mutual information, 64 bins and 300,000 samples as the parameters of the ANTs function. We used 100 × 10 × 10 iterations for the nonlinear transformation, using symmetric normalization (SyN)

implemented in ANTs. We used the value in the \*InverseWarp.nii.gz deformation-field file to evaluate nonlinear deformation, because this file contains nonlinear transformation registration elements. To reduce the computational cost, we applied the images to a 3D grid with 20-voxel spacing, and then collected the voxels in the grid. We calculated two values of  $L'$  and  $L$  from each pair of the collected voxels, where  $L'$  is the distance between two voxels in pre-SyN transformed images and  $L$  is the distance between two voxels from post-SyN transformed images (Supplementary Fig. 7). Both  $L'$  and  $L$  were measured using the pre-SyN transformed image as the baseline. We created graphs of  $|L - L'|$  as a function of distance  $L$ , calculating the root-mean-square error and s.d. As shown in Fig. 3b, RMS errors were less than ~20 µm, which is ~0.8% of the measurement length (2,500 µm) in the investigated regions.

**Evaluation of cellular loss during a CUBIC-X protocol.** We confirmed that no nuclei were lost during delipidation process (Supplementary Fig. 6). To visualize the nuclei before delipidation, we used brains of R26-H2B-EGFP mice. The brains were sectioned into 1-mm-thick coronal slices, then immersed in modified ScaleCUBIC-2 before the delipidation process. Fluorescence observation of nuclei was performed with confocal microscopy. Delipidation with modified ScaleCUBIC-1 was performed after observation, and samples were gently shaken at 37°C. The samples were cleared with modified ScaleCUBIC-2 1 h before confocal imaging and delipidated again until the end of the sequential observation period. Because repeated observation with confocal microscopy bleaches EGFP signals, propidium iodide was introduced to enhance the signal of nuclei starting at day 3.

We also confirmed that none of the nuclei were lost during expansion by comparing images taken before and after CUBIC-X treatment. The pre-CUBIC-X brain was prepared by delipidation and nuclear staining and imaged with a confocal microscope. Then the same sample was expanded and imaged with the customized LSM. The post-CUBIC-X brain image was linearly registered to images of the pre-CUBIC-X brain by ANTs using the affine transformation (Supplementary Fig. 6). To evaluate how the intercellular expansion ratio depends on cell density, images of the post-CUBIC-X brain were first registered to images of the pre-CUBIC-X brain by a similarity transformation. Then one pair of 2D images (250 µm × 250 µm) were extracted. Nuclei in the images were segmented by thresholding and converted to circular particles with the 'analyze particles' function in Fiji. One-dimensional cell density was measured as the number of nuclei on a connecting line divided by the length of the connecting line (Supplementary Fig. 6c). The intercellular expansion ratio was measured as  $L_2/L_1$  for a matched pair of nuclei (Supplementary Fig. 6d). Finally, the relationship between the intercellular expansion ratio and the 1D cell density was plotted (Supplementary Fig. 6e).

**Construction of the customized light-sheet fluorescent microscope.** Overviews of the customized light sheet microscope are illustrated in Supplementary Fig. 8. For notation of the  $xyz$  coordinate system, please refer to Fig. 3c,d. A full list of parts is presented in Supplementary Table 2, and CAD drawings are available upon request. The excitation light source was provided by a fiber-coupled CW laser (Omicron, SOLE-3 equipped with  $\lambda = 488, 594$  and 642 nm). After collimation, the beam was split into two paths (right and left excitation) using 50:50 beam splitter (Thorlabs). Left and right paths were essentially identical. The beams struck an electrically tunable lens (ETL, Optotune, EL-10-30) combined with offset lens with  $f = -50$  mm, positioned at the conjugated plane with the back focal plane of the illumination objective through a series of 4f relay lenses. The focal length of the ETL was tuned to adjust the focus point of the light sheet. A graduated iris was placed to adjust the diameter of the beam, which defined the effective NA of the illumination objective (NA = 0.1, WD = 30 mm, Nikon, CFI Plan 4×). The diameter of the iris was set at 3 mm, which corresponds to an effective NA of 0.03. NA = 0.03 corresponds to the minimum beam waist of about 10 µm and Rayleigh length of about 0.64 mm, which ensured a homogeneous light sheet thickness throughout the field of view (in this setup, 1.4 mm in the  $x$  direction). The first galvo mirror (Thorlabs, GVS001) was placed to adjust the  $z$ -position of the light sheet so that the excitation beam illuminated the focal plane of the detection objective. The second galvo mirror (Cambridge Technology, VM500Plus) was used to generate the light sheet, following the a DLMS scheme<sup>32</sup>. The mirror was swept at 500 Hz and the amplitude was adjusted to sufficiently cover the entire field of view (in this setup, 1.65 mm in the  $y$  direction). The detection objective (XLPLN10XSVM, NA = 0.6, WD = 8 mm, Olympus) was inserted directly into the hole located on the side of the custom-made sample chamber, and the chamber was filled with mixture of high-RI silicon oil (RI = 1.498, Momentive Performance Materials, TSF4300) and low-RI silicon oil (RI = 1.396, Momentive Performance Materials, TSF405) whose RI was adjusted to match to the brain samples (RI = 1.467). The samples were mounted on a motorized  $x-y-z-\theta$  stage through a custom-made sample holder ( $x$  and  $y$  stages, Thorlabs, MTS50/M-Z8E;  $z$  stage, Physik Instrumente, M-112.1DG;  $\theta$  stage, Physik Instrumente, M-116.DG). The gel-embedded brain sample was mounted onto the sample holder. The sample fluorescence was imaged with an sCMOS camera (Andor, Neo 5.5) through an  $f = 200$  mm tube lens, after passing through a custom-made filter wheel equipped with  $\phi = 32$  mm filters. Since the detection objective had a large back pupil, use of 32-mm filters was necessary to prevent clipping of the collected photon. All electronic devices were integrated and controlled by custom-written LabVIEW software (National Instruments).

**Tiling sequence in whole-brain imaging.** To image the entire brain with a limited field of view and working distance of the detection objective, we designed the following image acquisition sequence. Schematic of the procedures is presented in Fig. 3d. Initially, the  $x$ - $y$  coordinates of the imaging region were defined by specifying the top-left and right-bottom corner. The program then automatically divided the region into two-dimensional tiles, consisting of rectangular blocks whose size was determined by the field of view of the objective (in this setup,  $x = 1.4$  mm and  $y = 1.65$  mm) and the specified overlap value (10% for data shown in Supplementary Fig. 9b,c; 2% for others). The overlap value was defined to ensure that there would not be any loss of brain regions, and overlap was considered and handled properly in the data analysis. Next, the range of the  $z$ -scan was specified. Images were collected by scanning the sample in the  $z$  direction, with a step size of 5  $\mu$ m. We also implemented an autofocus function, which was executed every 200 images, to compensate for slight shifts of the focal plane of the detection objective due to the movement of the sample. This shift was considered to be caused by the fact that there was a slight RI mismatch between brain sample and immersion oil, and depending on the distance between the sample boundary and objective the light traveled slightly different paths. The autofocus routine swept several different light sheet positions by adjusting the sheet-positioning galvo mirror, and it calculated the image sharpness by applying a Sobel filter and taking the squared sum of the intensity to find the sheet position with maximum sharpness.

After collecting images from the ventral side, the sample was rotated by 180° using the  $\theta$  stage and the second half of the acquisition (dorsal side) was initiated. Since the sample was rotated along the  $y$  axis, it was critical to determine the center of rotation in the  $x$ - $z$  plane. This was done manually by locating identical cells or structures near the cerebral aqueduct from the  $\theta = 0^\circ$  and  $\theta = 180^\circ$  sides. In this way, two blocks of images were merged continuously (Supplementary Fig. 9a). The errors caused by rotation as well as by  $x$ - $y$  tiling were quantified, as described below. The excitation laser power was typically 1 mW (measured before the illumination objective), which was low enough to prevent fluorescence bleaching despite long imaging times. Single-image acquisition took approximately 200 ms (50 ms of exposure, 150 ms of stage motion), achieving an acquisition rate of 5 images/s. Each image was approximately 11 MB ( $2,160 \times 2,560$  pixels, 16 bit), corresponding to 55 MB/s data, which can easily be handled with hard disk drive arrays. The total image acquisition was completed within 4 d, yielding approximately 1.3 million images or 14 TB of data. The acquired data was either recorded on local hard drives or directly transferred to storage in the server through a 10-Gbit Ethernet cable.

**Imaging of whole mouse brains expressing fluorescent proteins at subcellular resolution.** To demonstrate subcellular-resolution imaging on a whole-brain scale, we imaged mouse brains expressing fluorescent proteins PLP-YFP and Thy1-YFP-H. Prior to imaging, these brains were delipidated, PI-stained, expanded and gel-embedded, as described above, and imaged with the customized LSMF. To further analyze the PLP-YFP mouse brain images, we applied a Frangi filter<sup>64</sup> to enhance the fibrous structure (Supplementary Fig. 10d). In the Frangi filter, a 3D Hessian matrix is constructed at each voxel, and the fibrousness and fiber orientation are determined by solving the eigenvalue and eigenvector of the matrix. Prior to filtering, image tiles were stitched together using TeraStitcher<sup>65</sup> to correct the stripes at tile boundaries, and the image resolution was reduced to 5  $\mu$ m on all axes to ensure a homogeneous resolution. The Frangi filter was implemented using Matlab code available online (<https://jp.mathworks.com/matlabcentral/fileexchange/24409-hessian-based-frangi-vesselness-filter>).

**Quantitative measurement of tiling displacements.** We used  $10 \times 10$  ( $x \times y$ ) tiles of 600-image stacks with the 10% overlap of each stack in both  $x$  and  $y$  directions to quantitatively evaluate the possible loss of cell detection originating from tile displacement (Supplementary Fig. 9b,c). We selected  $10 \times 10$  tiles so that pictures would not include space outside the brain. Brain samples from 1-week-old mice were used for image acquisition. Displacement was calculated with stitching software, TeraStitcher<sup>65</sup>. Displacement search regions must be set in the alignment step. We set  $sV$ ,  $sH$ , and  $sD$ , whose searching areas are 50 pixels around 200 slice-images per layer. Because larger searching regions return more correct value, we set the largest values of  $sV$ ,  $sH$  and  $sD$  in our computational setup. In thresholding, we set 0.5 as the threshold parameter to judge whether a tile is stitchable. In the placement step, displacement values were calculated by the minimum spanning tree algorithm. The resulting values in the xml:merging.xml file were analyzed. The displacement values were extracted both from SOUTH\_displacement and EAST\_displacement. We repeated the displacement evaluation twice and summarize it in Supplementary Fig. 9b.

To estimate theta tile displacement, we acquired  $10 \times 10$  ( $x \times y$ ) tiles of 600-image stacks from  $\theta = 0^\circ$  and  $\theta = 180^\circ$  using an identical brain, and then compared the two stacks. Prior to analysis, these datasets were stitched with TeraStitcher using the same parameters described above. Stitched images were further downsampled to a voxel size of  $45 \mu\text{m} \times 45 \mu\text{m} \times 45 \mu\text{m}$  ( $x \times y \times z$ ), and images from  $\theta = 180^\circ$  were registered to  $\theta = 0^\circ$  images with the ANTs similarity transformation. Mutual information with 64 bins was used as the similarity metric. The resulting matrix was decomposed to estimate the rotation angle around the  $x$ ,  $y$  and  $z$  axes (Supplementary Fig. 9c).

**Computational system and data storage.** The image processing server comprises two 12-core CPUs (Intel Corporation, Xeon E5-2590 v3), Intel 256-GB C612 Chipset (16 GB  $\times$  16 GB), DDR4-2133 SDRAM (Samsung, M393A2G40DB0-CPB, ECC Reg.), four GPUs (Nvidia Corporation, GeForce GTX TITAN X), a 12-GB/s SAS RAID card (Avago, MegaRAID SAS 9380-8e) with 42 hard disks (Seagate, ST100NM0033 and ST5000NM0084), two 1-TB SATA hard disks combined in a RAID-1 for the system, and forty 5-TB SATA hard disks combined in a RAID-0 for every ten hard disks for the data. The image processing server is connected to a workstation beside the microscope by a 10-GB Ethernet fiber (Mellanox Technologies, MC3309130-002) via two 10-GbE switches. The storage server comprises a CPU (Intel Corporation, Xeon E5-1620 v3), C612 Chipset (Intel Corporation), 8-GB DDR4-2133 SDRAM (Samsung, M393A1G40DB0-CPB, ECC Reg.) and a RAID controller (Avago, MegaRAID SAS 9361-8i) with thirty-six 6-TB SATA hard disks (HGST, Inc., HUS726060ALE610). One hard disk is a hot spare and other 35 hard disks are combined in a RAID-6. Effective capacity of these hard disks is 198 TB (1 TB for the system and about 197 TB for the data).

**Whole-brain cell-detection algorithms.** Cell detection was performed via our image-processing pipeline, implemented in Matlab for GPU-based 2D cell detection and in Python for CPU-based unification of detected cells to remove duplicated detections in the  $z$  direction. The local thresholds of the subregions were initially determined as shown in Supplementary Fig. 12. Each subregion ( $216 \text{ pixels} \times 256 \text{ pixels}$ ) was determined by dividing each single image ( $2,160 \text{ pixels} \times 2,560 \text{ pixels}$ ) into  $10 \times 10$  ( $x \times y$ ) regions. Modes were calculated using a histogram of pixel intensity with a bin size of 10.

Next, we set the background intensity at  $< 2 \times$  the mode intensity and extracted background signals (blue in Supplementary Fig. 12). Next, we calculated mean intensity and s.d. from the resulting histogram, and then set the image threshold at intensity  $< 2$  s.d. above the mean (red in Supplementary Fig. 12). We determined the range of thresholds both to avoid cell detection in empty regions of the images and to ensure cell detection without loss in dense regions such as the cerebellum. We detected local maxima after applying a 2D mean filter kernel with diameter  $D_{M1}$  in the first 2D cell-detection step (Supplementary Fig. 11a). We determined  $D_{M1}$  as 8.45  $\mu\text{m}$  (13 pixels) to maximize the accuracy of 2D detection among cerebral cortex, hippocampus and olfactory bulb (Supplementary Fig. 11d). Each intensity value and  $x$ - $y$ - $z$  coordinate of detected peaks in each stack were output in CSV file format. This GPU-based 2D cell-detection was completed within 35 h. Among all cells detected in the 2D cell-detection step, we excluded 'isolated' cells that were detected only in a single plane along the  $z$  direction, to remove noise. Next, we unified the multiply-detected cells after applying a 3D mean filter kernel with diameter  $D_{M2}$  in the second 3D-cell unifying step. The depth of the 3D mean filter kernel in the  $z$  direction was set at 20  $\mu\text{m}$  (4 pixels), and the diameter ( $D_{M2}$ ) was determined depending on the surrounding peak number ( $N_p$ ) in each subregion (Supplementary Fig. 11f). This is because optimal values for  $D_{M2}$  are relatively dependent on the density of cells. For example, setting  $D_{M2} = 13.0 \mu\text{m}$  (20 pixels) was the most accurate parameter value for brain regions with intermediate cell densities ( $20 \leq N_p < 40$ ) whereas  $D_{M2} = 18.2 \mu\text{m}$  (28 pixels) and  $D_{M2} = 10.4 \mu\text{m}$  (16 pixels) were more accurate for those with lower ( $N_p < 20$ ) and higher ( $40 \leq N_p$ ) cell densities, respectively. Then we introduced variable  $D_{M2}$  values optimized for the cell density of each region to maximize accuracy (Supplementary Fig. 11e,f). After mean filter application, local maxima in the  $z$  direction were detected.

Accuracy of this fully automated cell-detection protocol among various brain regions was evaluated by comparing the results of automatic detection with those from manual detection (Supplementary Fig. 13). To facilitate the accuracy evaluation, we categorized brain areas into two types, densely populated areas (DG and Cbgr) and sparsely populated areas (other brain areas). For accuracy evaluation of sparsely populated areas, more than 200 cells were manually detected. Accuracy was calculated as the frequency of true-positive detections divided by the sum of frequency of true-positive detection and the penalty value. The penalty value was calculated as the sum of the frequency of false-negative detections, frequency of false-positive detections, and frequency of overlapped detections. The frequency of overlapped detection comprised a count of the number of overlapped detections (for example, 1 for double detection, 2 for triple detection). For densely populated areas, we took subvolume images with a thinner light sheet. The usual low-illumination-NA condition was FWHM = 11.0  $\mu\text{m}$  (measured) with scanning step size = 5  $\mu\text{m}$  while the high-illumination-NA condition was FWHM = 5.2  $\mu\text{m}$  (measured) with scanning step size = 2  $\mu\text{m}$ . Using high-NA images, two people performed independent manual counts and the results were compared with the automatic count of low-NA images. Accuracy was calculated as described above.

To compare our two-step cell-detection algorithm with the direct 3D cell-detection algorithm<sup>33</sup>, we evaluated the computation time and cell-detection accuracy using the three image stacks from striatum, thalamus and cerebral cortex layer 5/6. Each stack included 100 images. All the images were analyzed for computation time, and parts of the images were used for accuracy estimation. For the direct 3D cell-detection algorithm, we applied a cylinder-shaped mean filter kernel with a diameter of 8.45  $\mu\text{m}$  (13 pixels) and a height of 25  $\mu\text{m}$  (5 pixels) and detected local maxima. For the two-step algorithm, we applied a circle-shaped mean filter kernel with a diameter of 8.45  $\mu\text{m}$  (13 pixels) and detected local maxima; then we used a cylinder-shaped mean filter kernel with a diameter of

8.45  $\mu\text{m}$  (13 pixels) and a height of 25  $\mu\text{m}$  (5 pixels) in the second 3D cell-unifying step. The ratio of computation time and accuracy of the cell-detection were compared (Supplementary Fig. 11b).

#### Construction of a single-cell-resolution mouse brain atlas (CUBIC-Atlas).

The CUBIC-X images of C57BL/6J mice brain were registered in the 25- $\mu\text{m}$  Allen Brain Atlas, and detected cells were then annotated. From the Allen Brain Atlas website, we downloaded the atlasVolume file, which is a 3D assembled Nissl volume of the reference brain used for registration, and the 'annotation' file, which contains annotation ID in a reconstituted brain used for annotation (<http://help.brain-map.org/display/mousebrain/API>). We prepared both atlasVolume and 'annotation' in TIFF format.

Initially, we prepared the downsampled images of CUBIC-X brain before registration because conventional registration software is unable to handle large files. The downsampled image size was  $361 \times 428 \times 230$  ( $x \times y \times z$ ) with a voxel size of  $80.12 \mu\text{m} \times 80.10 \mu\text{m} \times 80.00 \mu\text{m}$ , saved as a TIFF file. We fitted the image file to this size by appending pixels with a signal intensity of zero. Then the downsampled image was registered to atlasVolume (we term this the 'ABA brain') with an affine transformation using ANTs software. In the affine transformation, mutual information was used for the similarity metrics, with 64 bins and 300,000 samples as parameters of the ANTs function. The WarpImageMultiTransform command was used to generate an aligned image. We could not smoothly map 3D images of the CUBIC-X brain onto the ABA brain simply by using diffeomorphic registrations such as B-spline transformation or SyN transformation. Thus, we introduced an alternative registration based on the virtual coronal slices. After roughly aligning the CUBIC-X images by affine transformation, we generated 111 overlapping 3D coronal slices with a 2.2-mm thickness at 100- $\mu\text{m}$  steps from both the CUBIC-X and ABA brains (Supplementary Fig. 14a). Then each coronal-slice image of the CUBIC-X brain was nonlinearly registered onto the coronal-slice images of the ABA brain. We used a symmetric normalization model (SyN) as the nonlinear transformation. We used  $100 \times 10 \times 10$  iterations for SyN, with cross-correlation as the similarity metrics.

Because the transformation information was output as affine.txt files for the affine transformation and as InverseWarp.nii.gz files for the nonlinear transformation, we were able to trace the same registration process. In this procedure, each coronal plane of the CUBIC-X brain was mapped 22 times onto the ABA brain. We then calculated the normalized cross-correlations (NCCs) of each CUBIC-X coronal image with the ABA coronal images. Next, we applied individual cells to 22 different deformation fields and calculated 22 different destinations. By referring to NCCs of coronal plane images in the individual destinations, we identified the optimal deformation field displaying the highest NCC. Based on this procedure, we registered individual cells of the CUBIC-X brain to the ABA brain. Overall, it took about 3 d with our image processing server to complete the registration of the whole brain. We note that the coronal slice-based registration demonstrated better performance in annotating detected cells when compared with a simple affine transformation or SyN transformation (Supplementary Fig. 14).

After virtual-coronal-slice-based registration, we next annotated the detected cells by referring to the 'annotation' file. We analyzed the annotation data from all of the detected cells and removed cells annotated out of space in the ABA brain. Based on results of this registration and annotation procedure, we generated a 3D mouse brain atlas with a single-cell resolution and named this point-based brain atlas CUBIC-Atlas.

To generate 3D annotation regions, we applied the Matlab function `alphashape` and input an alpha diameter of 300 as a parameter for each annotated point set. To generate 2D annotation regions, we extracted 100- $\mu\text{m}$ -wide point sets from the selected plane to form the alpha shape, followed by projection to the 2D plane (Fig. 6c and Supplementary Fig. 14d).

**Mapping a Thy1-YFP-H mouse brain onto the CUBIC-Atlas.** To see how accurately anatomical annotations of a CUBIC-Atlas are performed, we overlaid cerebral cortical layers of a Thy1-YFP-H mouse brain onto those of CUBIC-Atlas. A Thy1-YFP-H mouse brain was delipidated, PI stained, expanded and gel-embedded as described above. Following fluorescence imaging of PI and YFP with our customized LSMF, the images were downsampled before registration. The downsampled PI signal image was then registered onto the CUBIC-Atlas with an affine transformation followed by nonlinear SyN transformation. The resulting affine matrix and deformation fields were applied to both the PI signal and the YFP signal. The registered volumetric images of PI and YFP are shown in Supplementary Fig. 16. Cerebral cortical layers of the CUBIC-Atlas are also visualized in pseudocolors, overlaid with the registered images of PI and YFP (Supplementary Fig. 16).

**Whole-brain cell-profiling of mice at different ages.** We performed whole-brain cell profiling of 1-week-old, 3-week-old, 8-week-old and 6-month-old C57BL/6N mice according to the CUBIC-X pipeline (Fig. 5a). Detected cells from the whole brain were registered and annotated to the anatomical regions in the CUBIC-Atlas as a reference. In this step, we applied an affine transformation and the following nonlinear SyN transformation to individual cells in each CUBIC-X image of brains

taken from mice of various ages. After registration, we detected the nearest-neighbor cell as a reference in CUBIC-Atlas to the cell of interest in the obtained images. To reduce computational cost, we performed this search in locally allocated grids ( $120 \mu\text{m} \times 120 \mu\text{m} \times 120 \mu\text{m}$ ) partitioned from the atlas space (Supplementary Fig. 17). We annotated the corresponding anatomical region of the nearest-neighbor cell in CUBIC-Atlas to the region of the cell of interest and excluded the cells that were annotated outside of the brain or allocated outside of the atlas space. To characterize the developmental progression of cell numbers at each age, we first analyzed the average cell numbers and s.d. for 1-week-old, 3-week-old, 8-week-old and 6-month-old mice in each region (Fig. 5b,c) and revealed that variations in cell numbers over different ages strongly depended on brain region. We calculated the coefficient of variation (CV), i.e., normalized s.d. by the average cell number (Fig. 5d), and mapped CV values on the mouse brain schema (Fig. 5e). A dendrogram schema was constructed using `networkX`<sup>66</sup> and `Gephi`<sup>67</sup> software and indicated the hierarchical annotation structure of ABA (Fig. 5e). To investigate developmental changes in cell numbers across the entire brain, we calculated normalized cell numbers for each region by setting the cell number averaged over ages to 1.0. For hierarchical clustering analysis, averaged values of normalized cell numbers within each age group were used. Brain regions were then clustered based on their normalized cell numbers over different ages, using the `heatmap.2` function in the `gplots` R package<sup>68</sup>, with Euclidean distance and the `ward.D2` clustering method (Fig. 5f). Four clusters were determined by the 'cutree' function in the R stats package. This clustering analysis revealed four major clusters (Fig. 5g,h).

To see how accurately brains were anatomically annotated during early postnatal development, we visualized the nucleic images of 1-week-old and 3-week-old mice with the anatomical annotations of detected cells. Cerebral cortical layers of somatosensory areas, posterior parietal association areas and visual areas are visualized with pseudo colors (Supplementary Fig. 19).

**Probabilistic annotation of CUBIC-Atlas.** Hyperactivated cells were detected as in our previous paper<sup>14</sup>, and the detected cells were then registered and annotated as above. Annotated cells were visualized in the brain (Fig. 7b and Supplementary Fig. 22). For probabilistic annotation of the CUBIC-Atlas, a Gaussian probabilistic distribution ( $\sigma = 60 \mu\text{m}$ , using a post-CUBIC-X brain as baseline) was generated around the registered cells. After integrating the probabilistic distribution generated from all registered cells, we annotated the cell in the atlas by referring to the value of the corresponding location in the resulting activity distribution. Probabilistically annotated cells in CUBIC-Atlas were then clustered with the `kmean++` algorithm implemented in Matlab.

**Statistical analysis.** Statistical analyses were performed in R version 3.3.1 or GraphPad Prism. Prior to Student's *t* tests and one-way ANOVA, normality was evaluated by Kolmogorov-Smirnov tests at a significance level of 0.05, and the homogeneity of variance for each group was evaluated by Bartlett's test. In both tests, significance level was set to 0.05. To analyze swelling of gelatin blocks or brains, one-way ANOVA with Tukey's post hoc tests were used when the groups displayed normal distribution with equal variance, or Dunnett's modified Tukey-Kramer pairwise multiple comparison test when groups displayed normal distribution without equal variance. For comparison of computation time, accuracy, normalized cell numbers and cell numbers, Student's *t* tests were used after confirmation of normal distribution with equal variance. In this study,  $P < 0.05$  was considered significant ( $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ; n.s., not significant). Data collection and analysis were not performed blind to the conditions of the experiments. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications<sup>14,23</sup>. The mice used for whole-brain cell counting were randomly chosen from colonies. Experiments were not performed in a blinded fashion.

**Life Sciences Reporting Summary.** Further information on experimental design is available in the Life Sciences Reporting Summary.

**Code availability.** The source codes used in detecting cells as well as mapping and annotating detected cells onto the CUBIC-Atlas are available from our website (<http://cubic-atlas.riken.jp>). Other software used in this study, including custom-written LabVIEW code controlling a custom LSMF, is available upon reasonable request.

**Data availability.** The CUBIC-Atlas can be downloaded from our website (<http://cubic-atlas.riken.jp>). The data supporting the findings of this study are available upon reasonable request.

## References

- Inamura, N. et al. Gene induction in mature oligodendrocytes with a PLP-tTA mouse line. *Genesis* **50**, 424–428 (2012).
- Tanaka, K. F. et al. Flexible Accelerated STOP Tetracycline Operator-knockin (FAST): a versatile and efficient new gene modulating system. *Biol. Psychiatry* **67**, 770–773 (2010).

53. Kanemaru, K. et al. In vivo visualization of subtle, transient, and local activity of astrocytes using an ultrasensitive  $\text{Ca}^{2+}$  indicator. *Cell Rep.* **8**, 311–318 (2014).
54. Abe, T. et al. Establishment of conditional reporter mouse lines at ROSA26 locus for live cell imaging. *Genesis* **49**, 579–590 (2011).
55. Hale, G. M. & Querry, M. R. Optical constants of water in the 200-nm to 200-microm wavelength region. *Appl. Opt.* **12**, 555–563 (1973).
56. Cope, M., Delpy, D. T., Wray, S., Wyatt, J. S. & Reynolds, E. O. A CCD spectrophotometer to quantitate the concentration of chromophores in living tissue utilising the absorption peak of water at 975 nm. *Adv. Exp. Med. Biol.* **248**, 33–40 (1989).
57. Wray, S., Cope, M., Delpy, D. T., Wyatt, J. S. & Reynolds, E. O. Characterization of the near infrared absorption spectra of cytochrome a3 and haemoglobin for the non-invasive monitoring of cerebral oxygenation. *Biochim. Biophys. Acta* **933**, 184–192 (1988).
58. Ke, M. T. et al. Super-resolution mapping of neuronal circuitry with an index-optimized clearing agent. *Cell Rep.* **14**, 2718–2732 (2016).
59. Aoyagi, Y., Kawakami, R., Osanai, H., Hibi, T. & Nemoto, T. A rapid optical clearing protocol using 2,2'-thiodiethanol for microscopic observation of fixed mouse brain. *PLoS One* **10**, e0116280 (2015).
60. Costantini, I. et al. A versatile clearing agent for multi-modal brain imaging. *Sci. Rep.* **5**, 9808 (2015).
61. Schwarz, M. K. et al. Fluorescent-protein stabilization and high-resolution imaging of cleared, intact mouse brains. *PLoS One* **10**, e0124650 (2015).
62. Avants, B. B. et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **54**, 2033–2044 (2011).
63. Murphy, K. et al. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. *IEEE Trans. Med. Imaging* **30**, 1901–1920 (2011).
64. Frangi, A. F., Niessen, W. J., Vincken, K. L. & Viergever, M. A. Multiscale vessel enhancement filtering. *Lect. Notes Comput. Sci.* **1496**, 130–137 (1998).
65. Bria, A. & Iannello, G. TeraStitcher - a tool for fast automatic 3D-stitching of teravoxel-sized microscopy images. *BMC Bioinformatics* **13**, 316 (2012).
66. Hagberg, A.A., Schult, D.A. & Swart, P.J. in *Proceedings of the 7th Python in Science Conference (SciPy2008)* (eds. Varoquaux, G., Vaught, T., Millman, J.) 11–15 (2008).
67. Bastian, M., Heymann, S. & Jacomy, M. in *International AAAI Conference on Weblogs and Social Media* <https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154> (2009).
68. Warnes, G.R. et al. gplots: various R programming tools for plotting data. *R package* version 2 <https://cran.r-project.org/web/packages/gplots/index.html> (2009).

## Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

### ▶ Experimental design

#### 1. Sample size

Describe how sample size was determined.

The sample sizes were determined based on the literatures in the fields. No statistical were used to predetermine sample sizes.

#### 2. Data exclusions

Describe any data exclusions.

For developmental progression analysis, we excluded the cerebellum and olfactory bulb from analysis due to their significant morphological change during early postnatal development. We also excluded cortical layer 1 because the cell numbers of this outermost area of the brain could be influenced by the residual arachnoid matter or the possible damages during the dissection procedures.

#### 3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

CUBIC-X pipeline was repeated more than ten times with independent brains and the results were reliably reproduced. For whole-brain cell counting, three brains for each age were used. The cell numbers and their variations in the areas used for early postnatal developmental analysis are shown in Figure 6 and Supplementary Figs. 20, 21. All three brains show the same trends.

#### 4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

The mice used for whole-brain cell counting were randomly chosen from colonies.

#### 5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

No blinding was done in this study because knowledge of experimental conditions was required during data collection. Analysis was conducted using automated workflow.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

## 6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- Test values indicating whether an effect is present  
*Provide confidence intervals or give results of significance tests (e.g.  $P$  values) as exact values whenever appropriate and with effect sizes noted.*
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on [statistics for biologists](#) for further resources and guidance.

## ► Software

Policy information about [availability of computer code](#)

## 7. Software

Describe the software used to analyze the data in this study.

Terastitcher, MATLAB, ANTs, networkX, Gephi, Imapris, R, Graphpad Prism, Microsoft excel, LABVIEW, Image J and Fusion360 were used. We used custom scripts for the construction of CUBIC-Atlas with anatomical annotation, and for activity mapping on CUBIC-Atlas.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). [Nature Methods guidance for providing algorithms and software for publication](#) provides further information on this topic.

## ► Materials and reagents

Policy information about [availability of materials](#)

## 8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

No unique materials were used.

## 9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

## 10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used .

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used .

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used .

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No commonly misidentified cell lines were used.

## ► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

### 11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

6-month-old Thy1-YHP-H transgenic female mice, 10-week-old PLP-YFP (PLP-tTA::tetO-ChR2EYFP) transgenic male mice and 11-week-old Mcl1-YFP (Mcl1-tTA::tetO-YC) transgenic female mice were used to render whole-brain imaging of fluorescent reporter proteins. 4-month-old R26-H2B-EGFP male mice were used for observation of influence of delipidation over nuclei. 8-week-old male mice of C57BL/6J were used to build the whole brain atlas with single-cell resolution. C57BL/6N mice were used to analyze whole-brain cell profiling over different aged mice. 1-week-old male mice were purchased from CLEA Japan, and 3-, 8-week- and 6-month-old male mice were from Japan SLC, Inc.

Policy information about [studies involving human research participants](#)

### 12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

The research did not involve human research participants.

In the format provided by the authors and unedited.

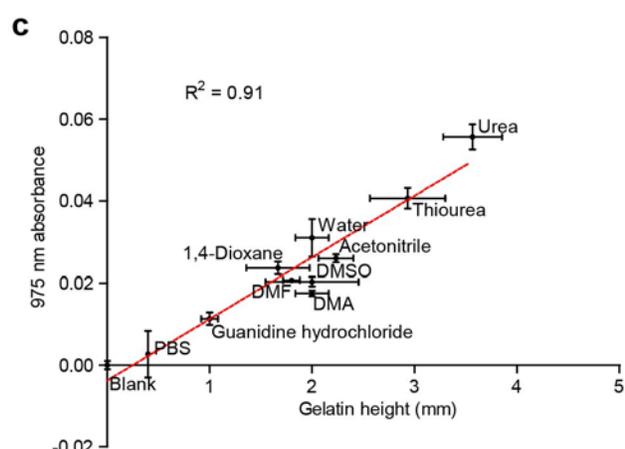
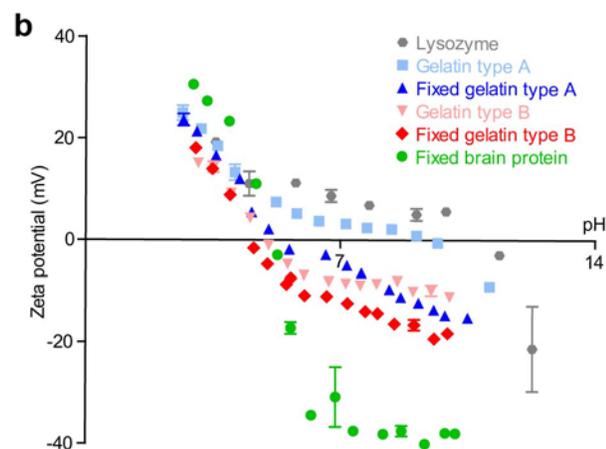
# A three-dimensional single-cell-resolution whole-brain atlas using CUBIC-X expansion microscopy and tissue clearing

Tatsuya C. Murakami<sup>1</sup>, Tomoyuki Mano<sup>1,2</sup>, Shu Saikawa<sup>3</sup>, Shuhei A. Horiguchi<sup>4,5</sup>, Daichi Shigeta<sup>1</sup>, Kousuke Baba<sup>6,7</sup>, Hiroshi Sekiya<sup>8</sup>, Yoshihiro Shimizu<sup>9</sup>, Kenji F. Tanaka<sup>10</sup>, Hiroshi Kiyonari<sup>11</sup>, Masamitsu Iino<sup>8,12</sup>, Hideki Mochizuki<sup>13</sup>, Kazuki Tainaka<sup>1,13</sup> and Hiroki R. Ueda<sup>1,2,4\*</sup>

<sup>1</sup>Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. <sup>2</sup>International Research Center for Neurointelligence (WPI-IRCIN), UTIAS, The University of Tokyo, Tokyo, Japan. <sup>3</sup>Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan. <sup>4</sup>Laboratory for Synthetic Biology, RIKEN Quantitative Biology Center, Osaka, Japan. <sup>5</sup>Department of Systems Science, School of Engineering Science, Osaka University, Osaka, Japan. <sup>6</sup>Department of Neurology, Graduate School of Medicine, Osaka University, Osaka, Japan. <sup>7</sup>Department of Kampo Medicine, Graduate School of Medicine, Osaka University, Osaka, Japan. <sup>8</sup>Department of Pharmacology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. <sup>9</sup>Laboratory for Cell-Free Protein Synthesis, RIKEN Quantitative Biology Center, Osaka, Japan. <sup>10</sup>Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan. <sup>11</sup>Animal Resource Development Unit and Genetic Engineering Team, RIKEN Center for Life Science Technologies, Kobe, Japan. <sup>12</sup>Division of Cellular and Molecular Pharmacology, Nihon University School of Medicine, Tokyo, Japan. <sup>13</sup>Brain Research Institute, Niigata University, Niigata, Japan. \*e-mail: [uedah-ky@umin.ac.jp](mailto:uedah-ky@umin.ac.jp)

**a**

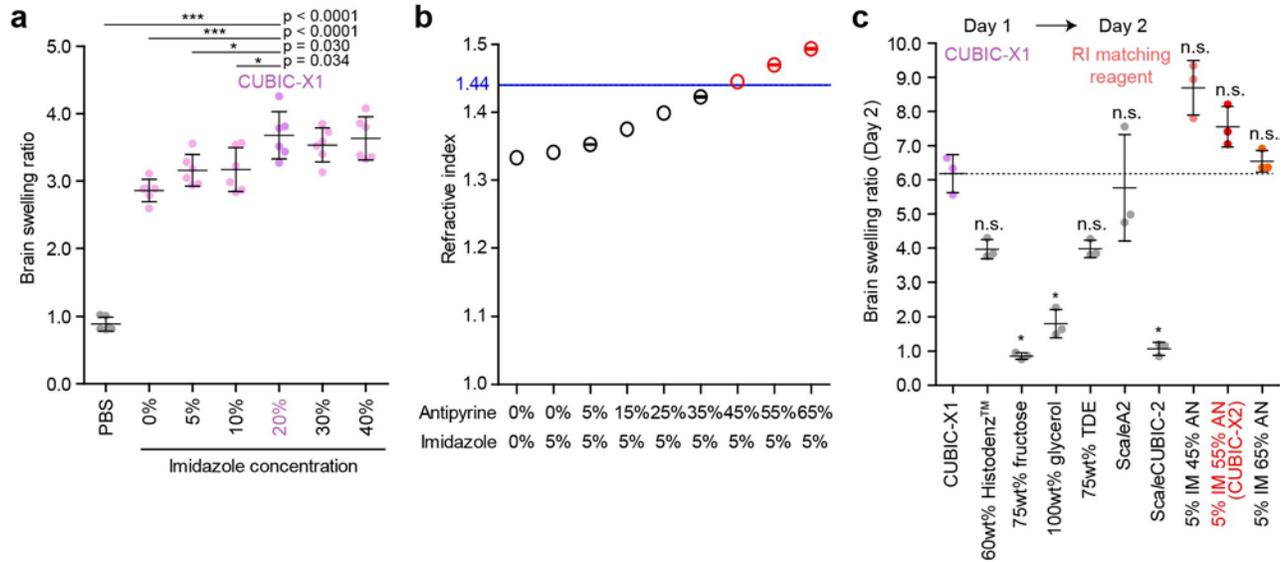
		PBS	Water	Ionic				Tween20	Nonident P-40	Triton X-100	Acetonitrile	DMF	DMA	1,4-Dioxane	DMSO	Urea derivative	
				Guanidine hydrochloride (GH)	Sodium deoxycholate (SD)										Urea	Thiourea	
Brain	Average	0.847	1.514	0.553	0.774	1.219	1.361	1.412	1.541	1.542	1.585	1.614	1.673	1.714	1.896		
	SD	0.021	0.004	0.001	0.003	0.005	0.063	0.005	0.075	0.024	0.012	0.017	0.056	0.037	0.031		
Gelatin	Average	0.999	1.544	0.798	0.808	1.113	1.215	1.251	1.473	1.541	1.496	1.516	1.473	1.616	1.458		
	SD	0.022	0.041	0	0.008	0.024	0.02	0.033	0.046	0.075	0.07	0.061	0.134	0.064	0.047		



**Supplementary Figure 1**

Development of gelatin-based high-throughput screening assay.

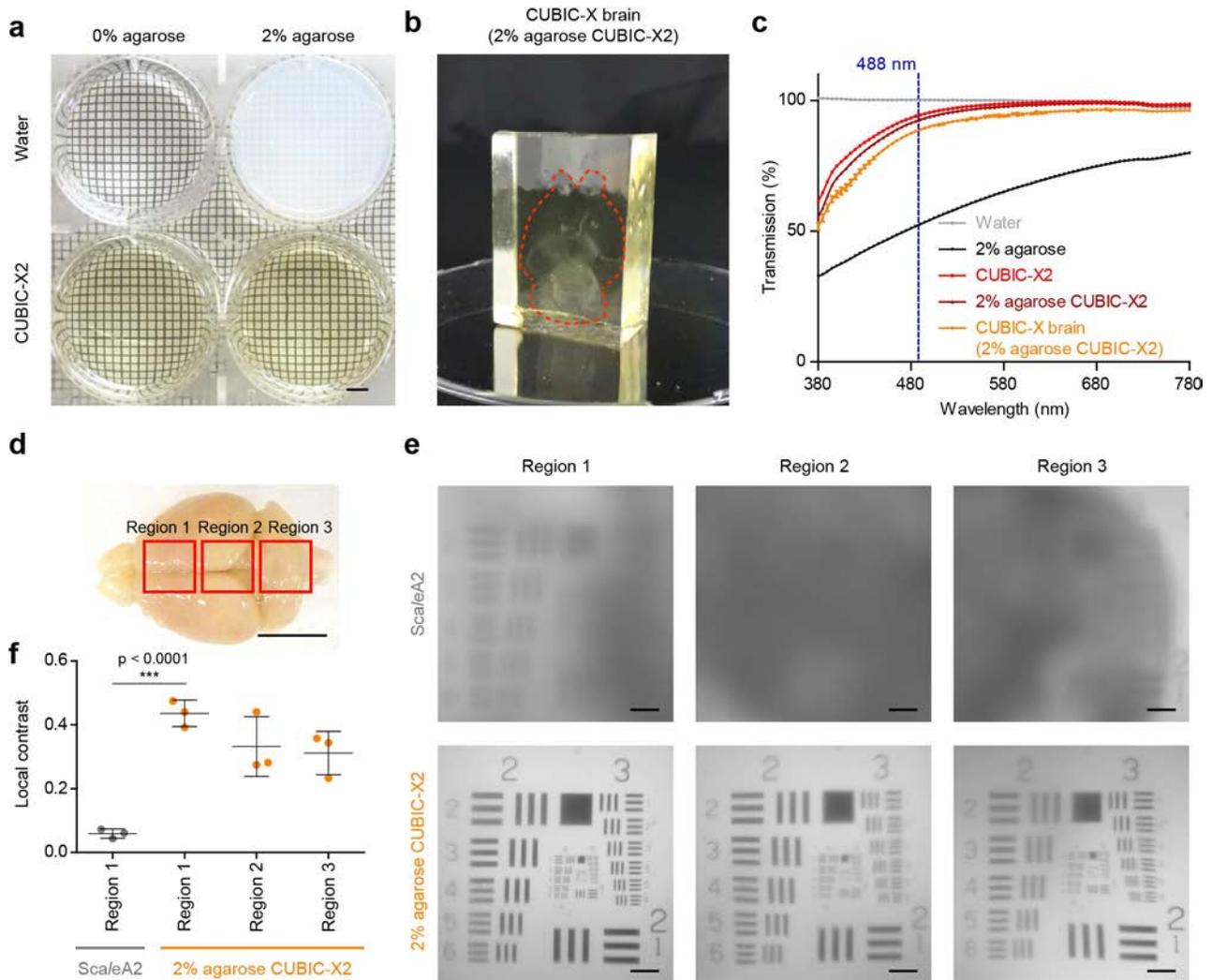
(a) The swelling ratio (the average and SD) of chemically treated brain samples and gelatin gel between before and after chemical treatment shown in Fig. 1a (brain;  $n = 2$ , gelatin;  $n = 3$ ). All chemicals are chosen from our previous study (Online Methods). Urea derivatives (blue color) remarkably swelled brain samples compared to ionic chemicals (red color). (b) The plots of zeta potential against pH for various protein samples. Each zeta potential curve was fitted to fourth-order polynomials. The isoelectric points were determined by calculating crossing points of polynomials with the x-axis. The measurements were repeated three times at each pH point. (c) Correlation of the 975 nm absorbance and apparent height of type B gelatin ( $n = 3$ ) treated with PBS, water, or 8 other chemicals indicated in the panel. The 975 nm absorbance of chemically treated gelatin gels showed a linear correlation with the apparent height of those. R-squared value was calculated with all data plots. All values are mean  $\pm$  SD.



### Supplementary Figure 2

CUBIC-X1 and CUBIC-X2 for whole-brain expansion and hyperhydrative RI matching.

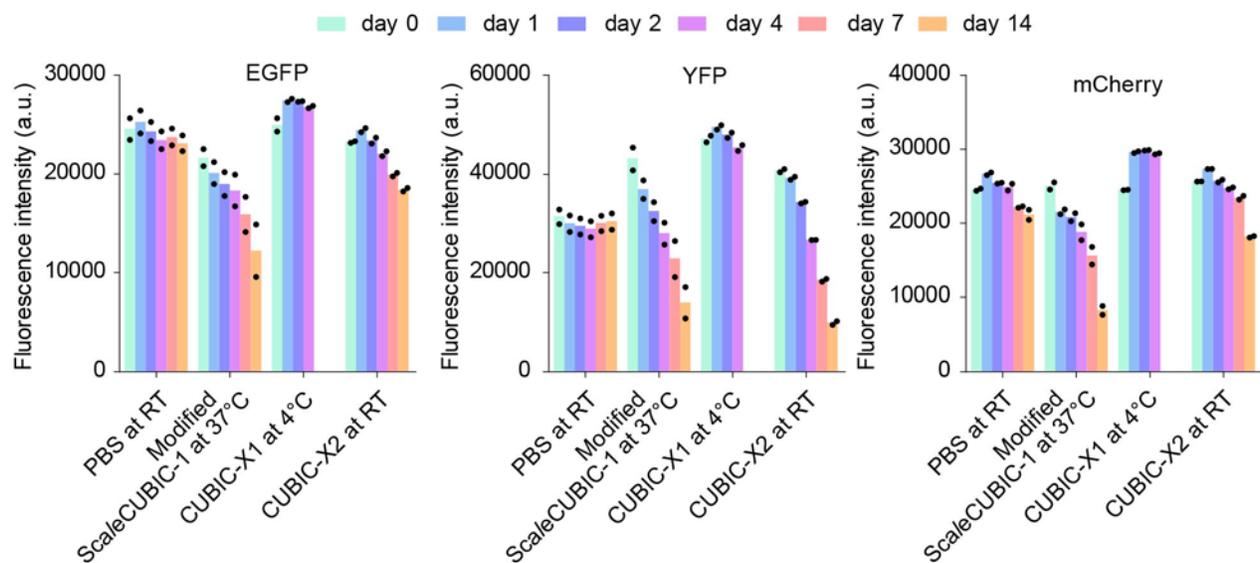
(a) Swelling ratios of delipidated hemisphere brains treated by imidazole aqueous solutions for 1 day. Delipidated hemisphere brains were immersed in various concentration of imidazole aqueous solutions ( $n = 6$ ). The 20% imidazole solution, termed CUBIC-X1, was the minimal concentration to achieve the largest expansion volume (\*\* $p < 0.001$ , \* $p < 0.05$ , one-way ANOVA, Tukey's post hoc test for multiple comparisons). (b) Refractive indices of cocktails of imidazole and antipyrine ( $n = 3$ ). We note that the standard error of several points cannot be shown because those measured values rounded at third decimal place resulted in the identical values. (c) Swelling ratios of hemisphere brains by RI matching media. Delipidated hemisphere brains were treated with CUBIC-X1 for 1 day, and then immersed in various kinds of RI media for 1 day ( $n = 3$ , \* $p < 0.05$ , Dunnett's Modified Tukey-Kramer pairwise multiple comparison test). All values are mean  $\pm$  SD.



### Supplementary Figure 3

Embedding of a CUBIC-X brain into an agarose gel.

(a) Transmission images of water, CUBIC-X2, 2% agarose, and 2% agarose CUBIC-X2. 2% agarose CUBIC-X2 showed high transparency while 2% agarose showed white turbidity. (b) Image of a CUBIC-X brain (2% agarose CUBIC-X2). The gel embedded expanded brain was highly transparent and solid enough to mount on the LSMF. (c) Transmission curves of water, CUBIC-X2, 2% agarose, 2% agarose CUBIC-X2, and CUBIC-X brain (2% agarose CUBIC-X2). Light transmittance around the visible region (380-780 nm) was measured ( $n = 3$ ). (d) Three regions of the brain quantified in e. (e) ScaleA2 and CUBIC-X2 treated CUBIC-X1 expanded brains were placed on USAF chart for quantitative analysis of scattering. The CUBIC-X2 brains were embedded in 2% agarose. (f) Comparison of the local contrasts of the ScaleA2 and CUBIC-X2 treated CUBIC-X1 expanded brains. The local contrasts were quantified by referring local intensity profiles over element 2, group 2 of USAF chart (Online Methods,  $n = 3$ ,  $***p < 0.001$ , independent two-tailed t-test). All values are mean  $\pm$  SD. Scale bars indicate 5 mm (a, d) and 500  $\mu$ m (e).

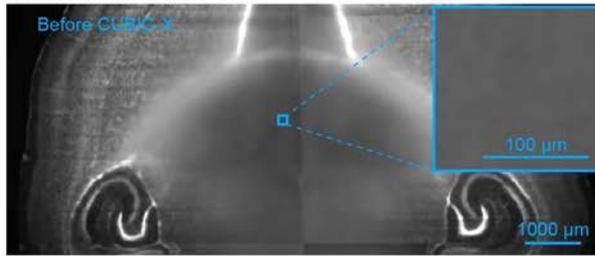


#### Supplementary Figure 4

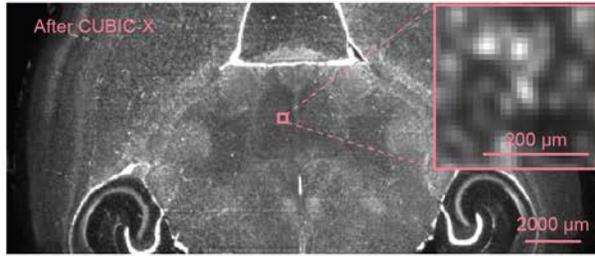
Intensity of fluorescent proteins in various CUBIC solutions.

Quenching test of fluorescent proteins by several CUBIC reagents. The indicated recombinant fluorescent proteins were incubated in each reagent or PBS for up to two weeks. The bars indicate the averages ( $n = 2$ ).

**a** Commercial LSMF, 2× zoom



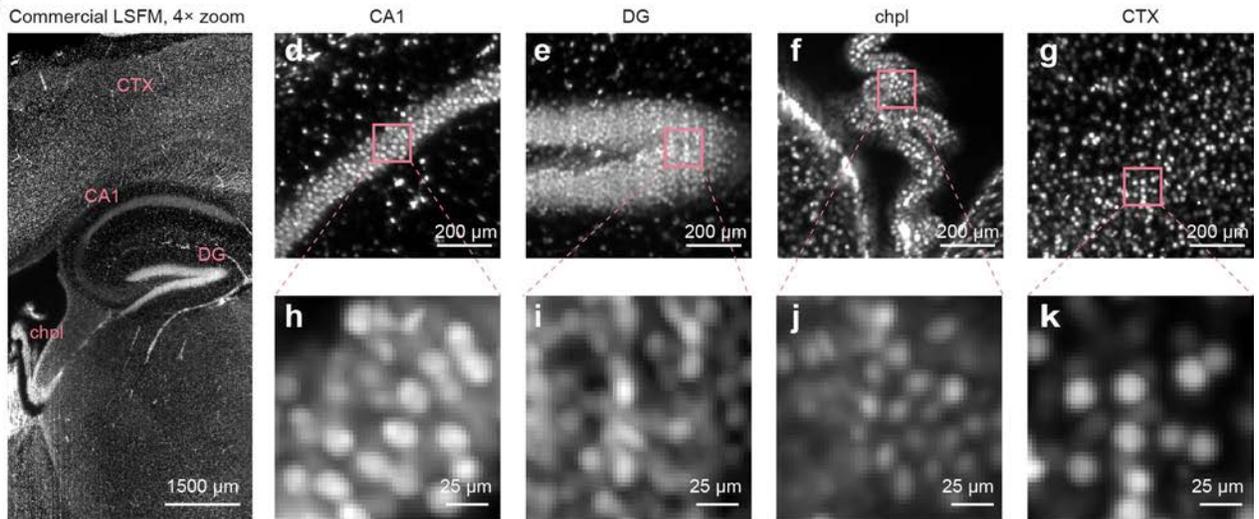
Commercial LSMF, 1× zoom



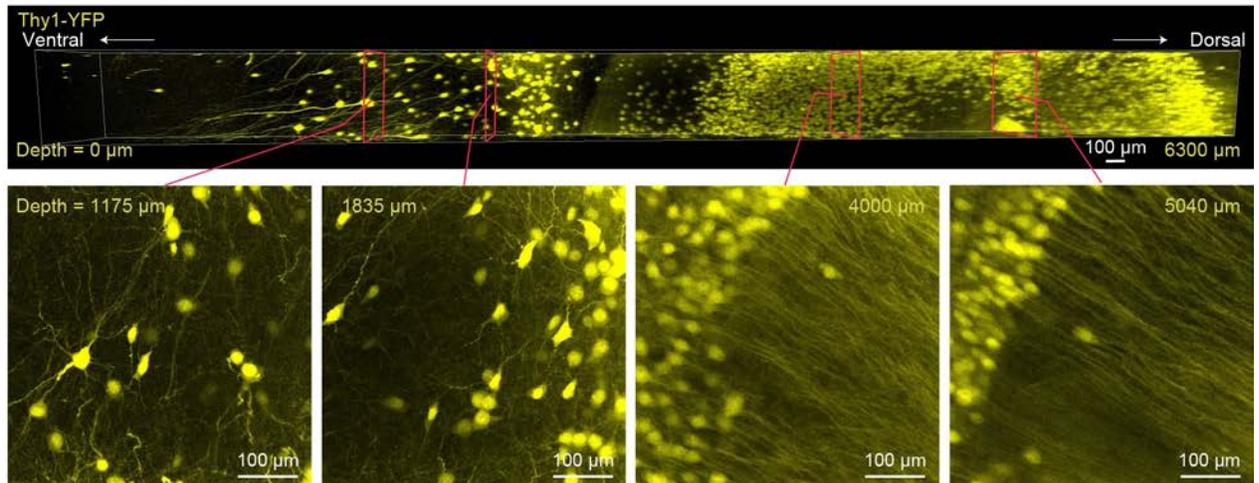
**b**



**c** Commercial LSMF, 4× zoom



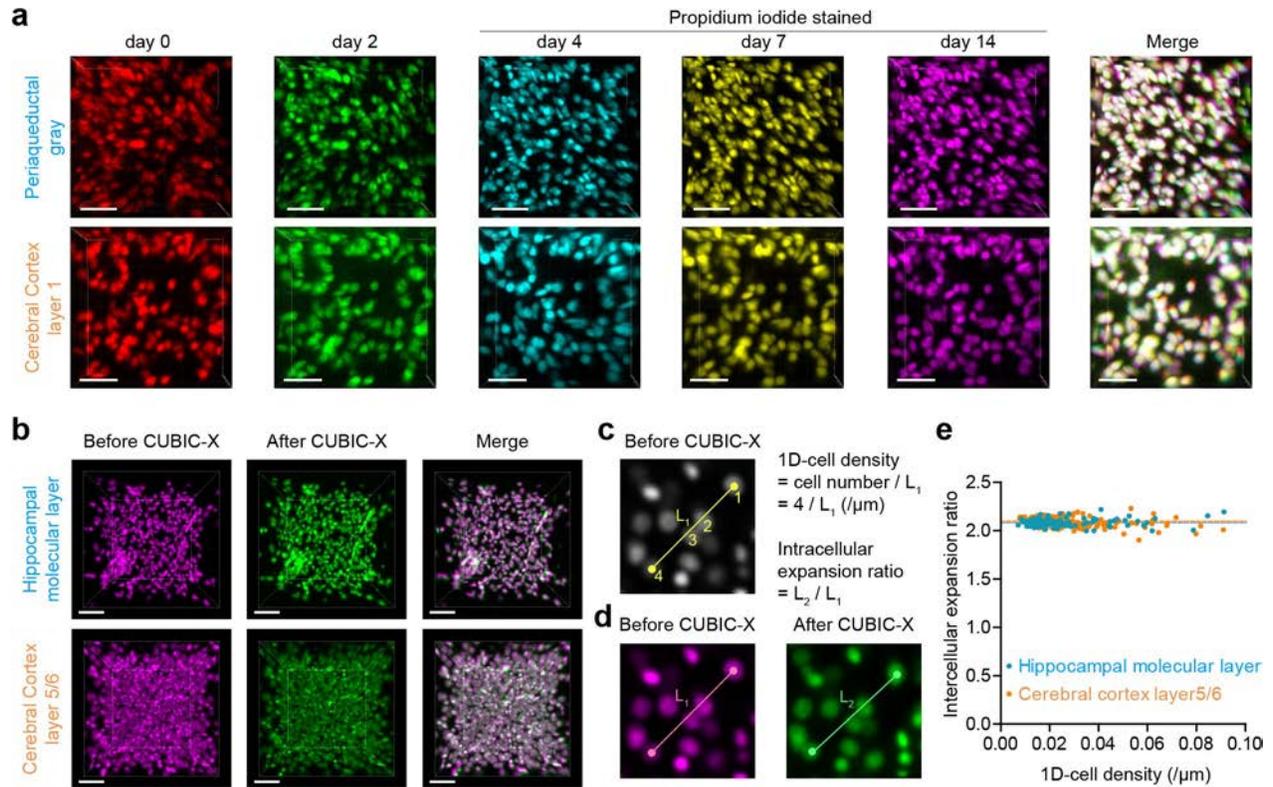
**l** Commercial confocal



## Supplementary Figure 5

CUBIC-X brain images acquired by commercial LSFM and commercial confocal microscopy.

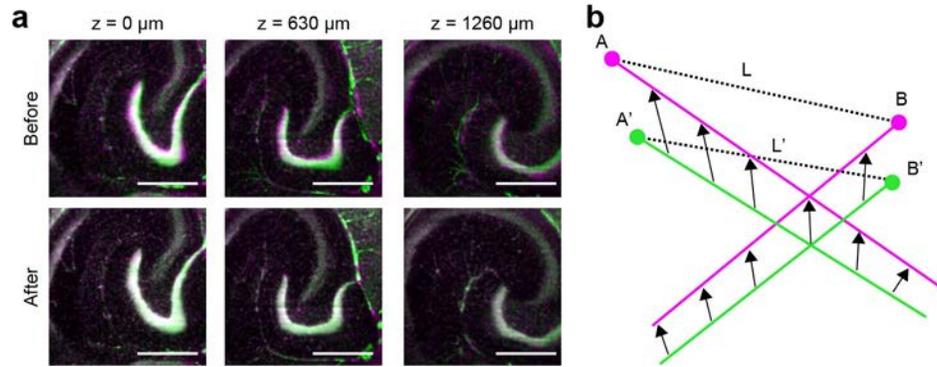
(a) Using a commercial LSFM, a low-magnification images of a PI-stained brain before (upper) and after (lower) CUBIC-X were obtained with 2x and 1x zoom lens, respectively (Online Methods). Insets: magnified views of the brains. (b) Volume-rendered images of a PI-stained brain before (left) and after (right) CUBIC-X in the same scale. Volumes were quantified by segmenting the brains (Online Methods). (c) PI-stained brain after CUBIC-X treatment was imaged with commercial LSFM with 4x magnification. To generate a virtual thin light sheet, four images with different light sheet focus positions were tiled together. DG: dentate gyrus, chpl: Choroid plexus, CTX: Cerebral cortex. (d-g) Magnified views of representative regions in brain. (h-k) Magnified views of d,e,f,g, respectively. (l) Fluorescent protein imaging of a Thy1-YFP-H mouse brain with a commercial confocal microscope. A Thy1-YFP-H mouse brain was cleared and expanded with CUBIC-X protocol, and imaged with confocal microscopy (25x, NA = 1.0). The volume rendered image (an upper panel; a volume size of 0.5 mm × 0.5 mm × 6.3 mm) and images of maximum-intensity projections (lower panels; 225 μm, 225 μm, 10 μm, 10 μm in thickness from left) are visualized. For all panels, experiments were repeated three times with independent brains. The representative images are shown.



**Supplementary Figure 6**

Multi-cellular scale evaluation of distortion and cellular loss during delipidation and expansion procedures.

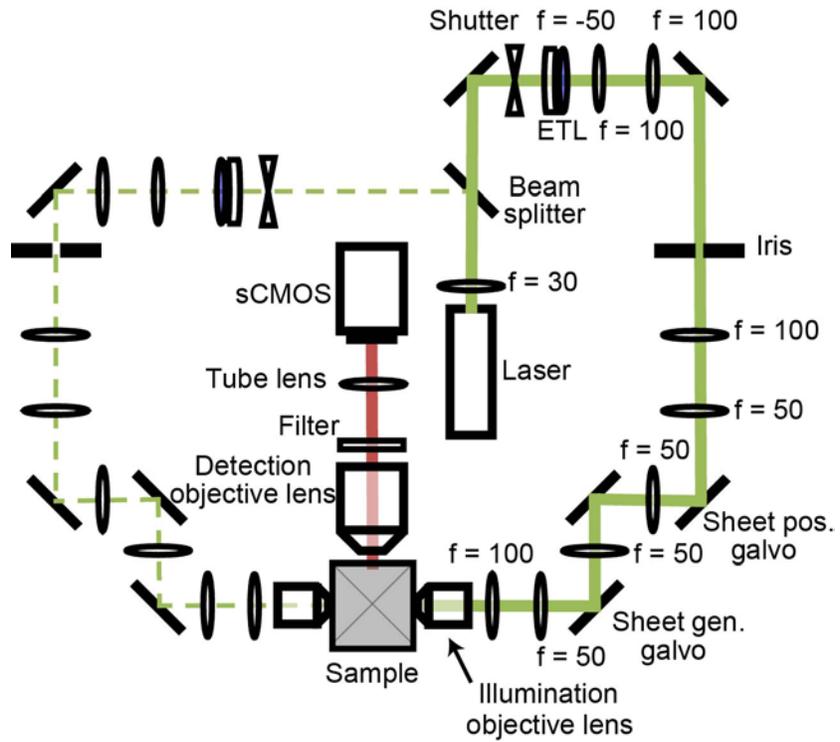
(a) Sequential observation of nuclei during the delipidation process. The sliced brains of R26-H2B-EGFP mice were observed prior to the delipidation process (day 0). The slices were repeatedly observed during the delipidation process (day 2, 4, 7 and 14). Because repeated observation with fluorescence microscopy bleaches EGFP signal, propidium iodide was added at day 3 to enhance the signal of nuclei. Two regions of the brains (periaqueductal gray and cerebral cortex layer 5/6) are shown. Scale bars indicate 25  $\mu\text{m}$  in the scale of a “before-CUBIC-X” brain. (b) Multi-cellular scale comparison of “before-CUBIC-X” brain and “after-CUBIC-X” brain. Two regions (hippocampal molecular layer and cerebral cortex layer 5/6) are shown. The corresponding images of “after-CUBIC-X” brain, which were linearly registered to “before-CUBIC-X” brain images by affine transformation, are also shown (Online Methods). It was confirmed that none of the nuclei were lost in the investigated area. The merged images after registration are shown on the right. Scale bars indicate 50  $\mu\text{m}$  in the scale of a “before-CUBIC-X” brain. In (a) and (b), a perspective 3D rendering was used by Imaris software for visualization. Experiments were repeated three times with independent brain slices or brains. The representative images are shown. (c,d) Schematic procedures of the measurement of 1D-cell density (c) and intercellular expansion ratio (d). 1D-cell density was measured for every pair of cells by counting the number of nuclei between the pair. Intercellular expansion ratio was calculated for every pair of cells as  $L_2/L_1$ . For the analysis, images of “after-CUBIC-X” were registered to images of “before-CUBIC-X” by similarity transformation. (e) The relation between 1D-cell density and Intercellular expansion ratio. Blue and orange dashed lines indicate overall expansion ratios in hippocampal molecular layer and cerebral cortex layer 5/6, respectively.



### Supplementary Figure 7

Global scale evaluation of distortion during expansion procedures.

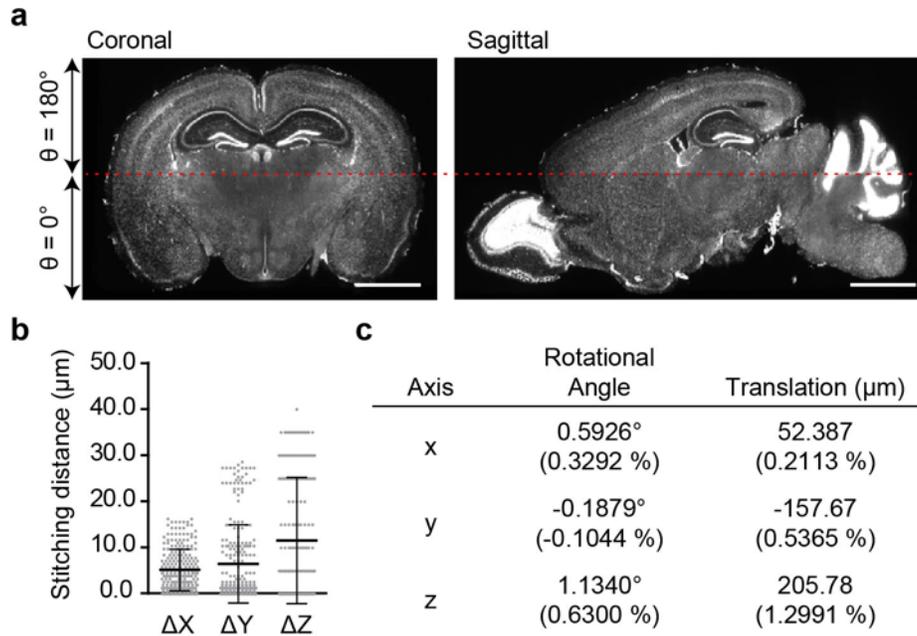
(a) Comparison of merged images of “before-CUBIC-X” (magenta) and “after-CUBIC-X” (green) brain between before- and after-non-linear transformation. After linear registration by affine transformation, three horizontal images were merged in the hippocampal region ( $2.5 \text{ mm} \times 2.5 \text{ mm} \times 2.5 \text{ mm}$ ) of a “before-CUBIC-X” (magenta) brain to the corresponding images of an “after-CUBIC-X” (green) brain in the upper row. After further non-linear transformation, the corresponding images are shown in the lower row. Scale bars indicate  $500 \mu\text{m}$  in the scale of a “before-CUBIC-X” brain. (b) Schematic procedures of the measurement of root-mean-square (RMS) error value in Fig. 3b.  $L'$  and  $L$  were calculated by measuring the Euclidean distance between two points of before non-linear registration (shown as  $A'$  and  $B'$ ) and after non-linear registration (shown as  $A$  and  $B$ ), respectively. Both of  $L'$  and  $L$  were measured in the scale of the before non-linear transformed image.



**Supplementary Figure 8**

Optical layout of the customized light sheet microscope.

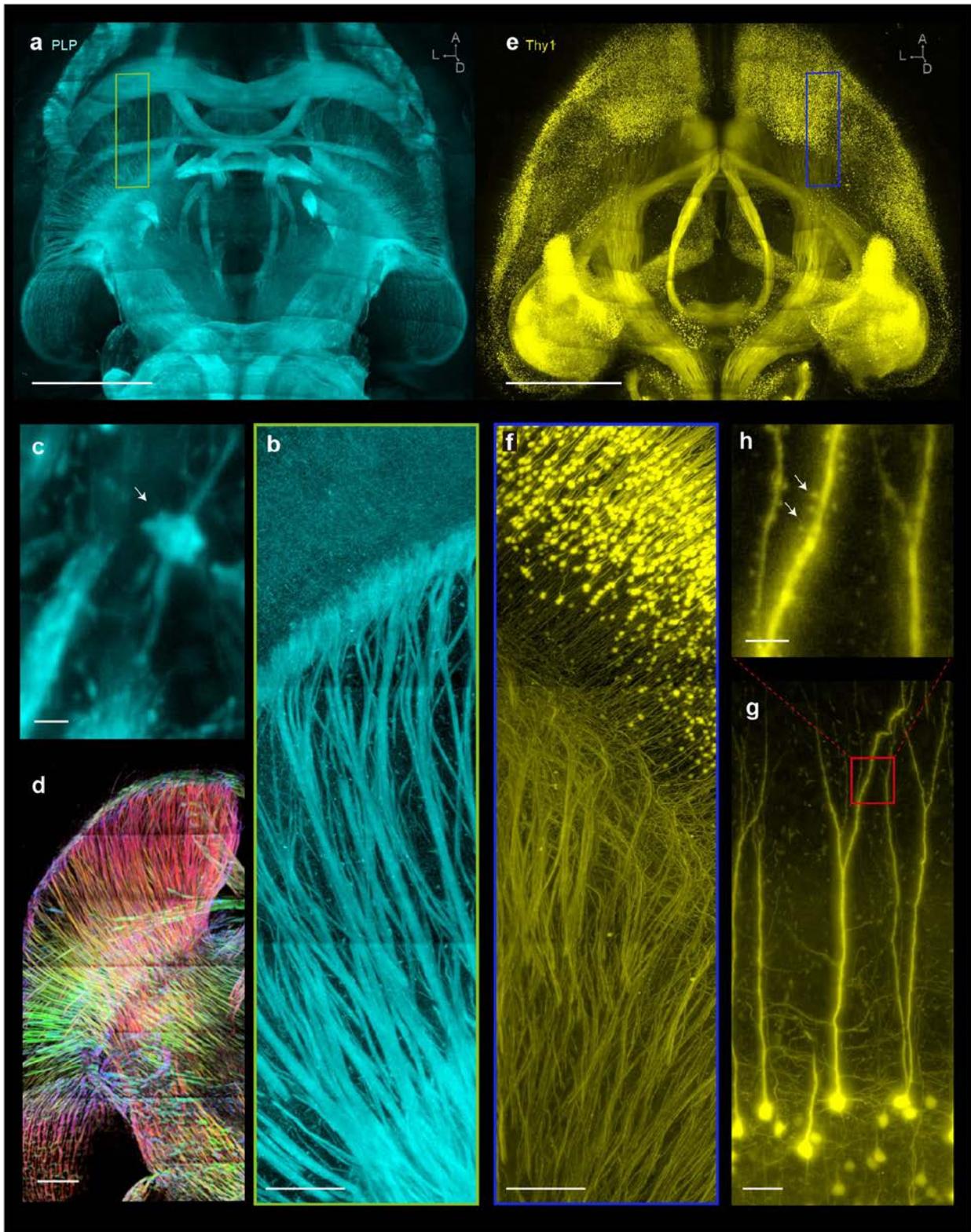
Electrically tunable lenses (ETL) were put for the adjustment of the light-sheet focus position along the x-direction. Galvanometer scanners were introduced for adjusting a light-sheet position along z-direction (shown as “sheet pos. galvo”). Autofocus (i.e. adjustment of the light-sheet z-position) by these galvanometer scanners was performed every 1 mm.



### Supplementary Figure 9

Quantitative measurements of the displacements during the imaging sequence.

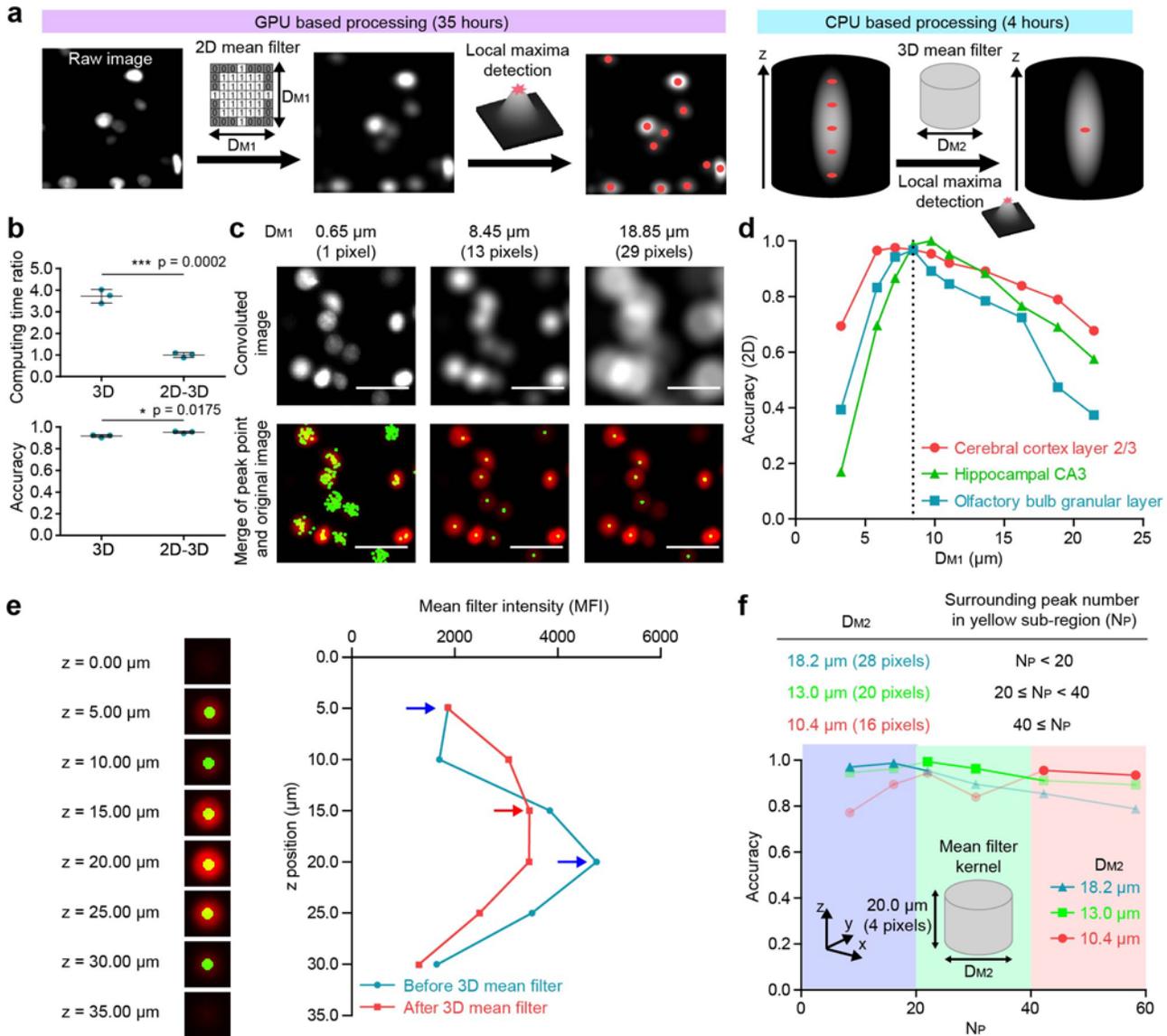
(a) Coronal and sagittal image merged by  $\theta$ -rotation with dorsal ( $\theta = 0^\circ$ ) and ventral ( $\theta = 180^\circ$ ) sides of the CUBIC-X brain. Red dashed line is the merged position. Scale bars indicate 2000  $\mu\text{m}$  in the scale of a “before-CUBIC-X” brain. (b) Quantitative evaluation of stitching displacement errors in x, y, and z-axes in the xy-direction tiling. We prepared  $10 \times 10$  tiles of 600 images for each stack with 10% overlap in both x- and y-direction. Stitching distance was calculated with the stitching software, TeraStitcher. The displacement errors of all possible combinations of the stacks are shown. The average  $\pm$  SD are shown. (c) Quantitative evaluation of theta tile displacement around x, y, and z-axes. We prepared two sets [ventral ( $\theta = 0^\circ$ ) and dorsal view ( $\theta = 180^\circ$ )] of  $10 \times 10$  tiles of 600 images for each stack with 10% overlap in both x- and y-direction. These images were stitched with TeraStitcher, and the resulting stitched images were downsampled. The downsampled ventral images were registered to the downsampled dorsal images by ANTs. Then Euler angle around x, y, and z-axes, and translation were calculated.



## Supplementary Figure 10

Whole-brain imaging of a CUBIC-X expanded brain with customized LSMF resolves multiscale cellular structures.

(a) Volume rendered image of PLP-tTA::tetO-ChR2EYFP mouse brain, imaged by customized LSMF. (b) Magnified view of a in the region enclosed in the green box. The maximum intensity projection along V-D axis spanning 500  $\mu\text{m}$  was applied to generate the image. (c) Detailed view of a from the striatum region. Because CUBIC-X highly retains the fluorescent proteins, subcellular structures of a single oligodendrocyte cell (indicated by a white arrow) can be clearly imaged by LSMF. (d) From the images shown in a, the fiber orientations were color-coded by their orientations, which were determined by applying fiber enhancement filter (Online Methods) (e) Volume rendered image of Thy1-YFP-H mouse brain, imaged by customized LSMF. (f) Magnified view of e in the region enclosed in the blue box. The maximum intensity projection along V-D axis spanning 1000  $\mu\text{m}$  was applied to generate the image. (g) Detailed view of e from the cerebral cortex. (h) A magnified view of g in the region enclosed in the red box. Branching of axons as well as single spines (indicated by white arrows) can be clearly resolved by expanding the sample and using the high-resolution LSMF. In b, c, f, g and h, the image intensity was adjusted by taking the logarithms of raw intensity for better visual contrast. For b and f, three image tiles were stitched together by blending the overlapping regions with linear weight using "Pairwise Stitching" plugin in ImageJ. Scale bars indicate 5 mm (a), 500  $\mu\text{m}$  (b), 20  $\mu\text{m}$  (c), 1 mm (d), 5 mm (e), 500  $\mu\text{m}$  (f), 100  $\mu\text{m}$  (g), and 20  $\mu\text{m}$  (h) in the scale of "after-CUBIC-X" brain.

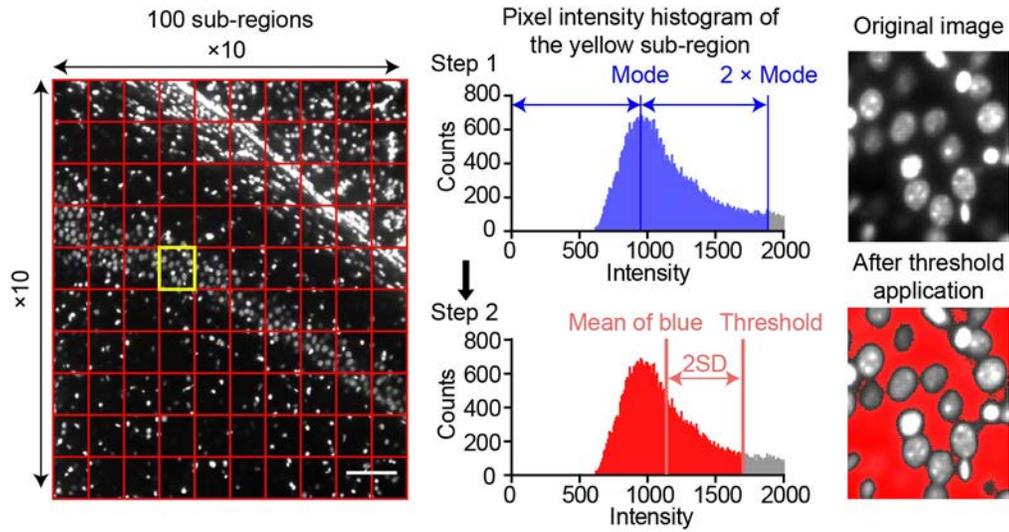


**Supplementary Figure 11**

A two-step convolution algorithm for the detection of cell nuclei in a whole brain.

(a) Schematic of cell detection of nuclear-stained CUBIC-X brain images. To establish two-step cell-detection algorithm, we first detected individual cells in a 2D xy-plane, and then unified the multiply-detected cells in 3D volume. In each step of this algorithm, we detected local maxima after applying a different mean filter. We used a 2D mean filter kernel with a diameter  $D_{M1}$  in the first GPU-based 2D cell-detection step, and 3D mean filter kernel with a diameter  $D_{M2}$  in the second CPU-based 3D cell-unifying step, respectively. (b) Comparison of calculation speed and accuracy in cell detection between 3D-based one-step convolution algorithm and our two-step algorithm (\*\* $p < 0.001$ , \* $p < 0.05$ , independent two-tailed t-test). The average  $\pm$  SD were evaluated in the independent three regions. (c) The effect of the first 2D mean filter in cell detection. While no filters [ $D_{M1} = 0.65 \mu\text{m}$  (1 pixel)] to the raw data resulted in multiple detections of a single cell, large diameter [ $D_{M1} = 18.85 \mu\text{m}$  (29 pixels)] led to the failure in detecting all cells within the image. The accuracy of cell detection depends on mean filter diameter ( $D_{M1}$ ). Scale bars indicate 50  $\mu\text{m}$  in the scale of an “after-CUBIC-X” brain. (d) The plots of the accuracy of 2D cell-detection against  $D_{M1}$  in cerebral cortex layer 2/3, hippocampal CA3, and olfactory bulb granular layer, respectively. We found  $D_{M1}$  of 8.45  $\mu\text{m}$  was the optimal value over the three regions. (e) The effect of the second 3D mean filter in cell detection. After 2D cell detection, a single cell was multiply detected over the several stacked images (left). Then, we attempted to unify the multiply-detected points by detecting local maxima of these points in the z-direction. We succeeded in the convergence of the multiply detected points (red arrow; local maxima of after 3D mean filter) after 3D mean filtering while failing to unify these points before 3D mean filtering blue arrow; local maxima of before 3D mean filter. (f) The plots of the accuracy of 3D cell-unification against

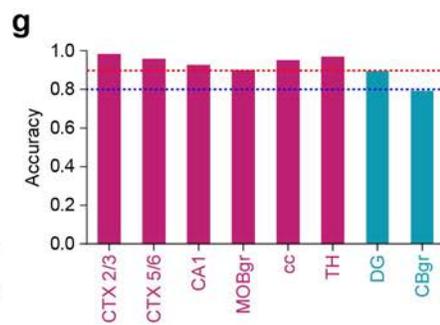
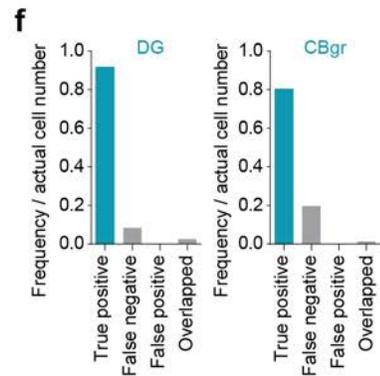
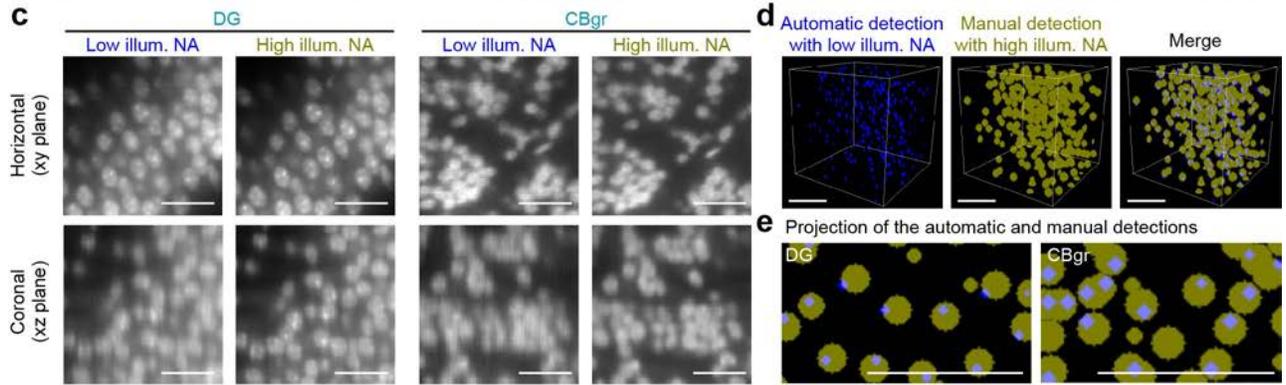
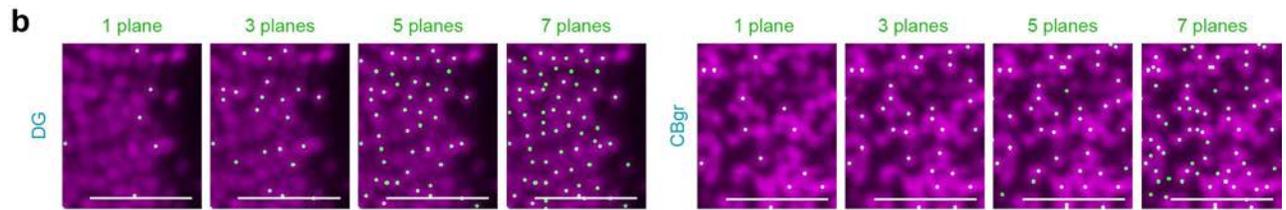
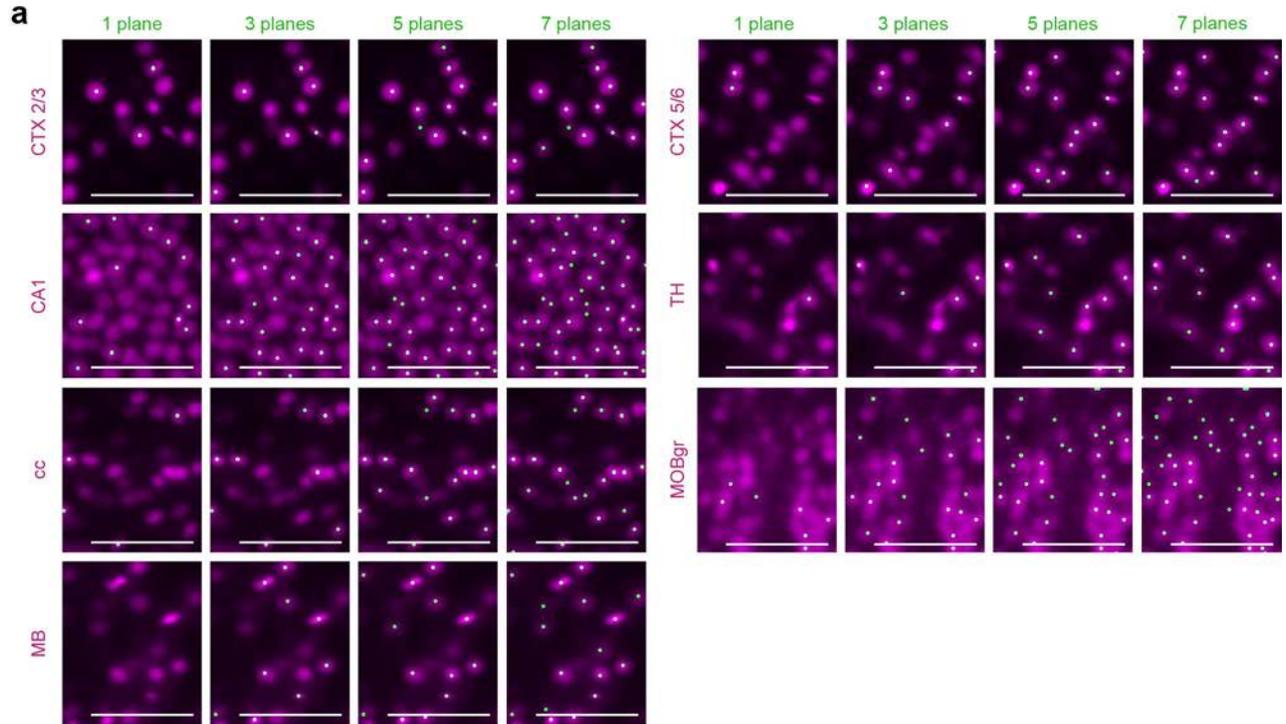
surrounding peak number (NP, Online Methods) in the three different  $D_{M2}$  value. We note that optimal values for  $D_{M2}$  were highly dependent on the density of cells. 13.0  $\mu\text{m}$  for  $D_{M2}$  was the most accurate parameter value for the brain region with an intermediated cell density whereas 18.2  $\mu\text{m}$  and 10.4  $\mu\text{m}$  for  $D_{M2}$  were the more accurate for those with lower and higher cell density, respectively.



### Supplementary Figure 12

The automated determination of an image threshold.

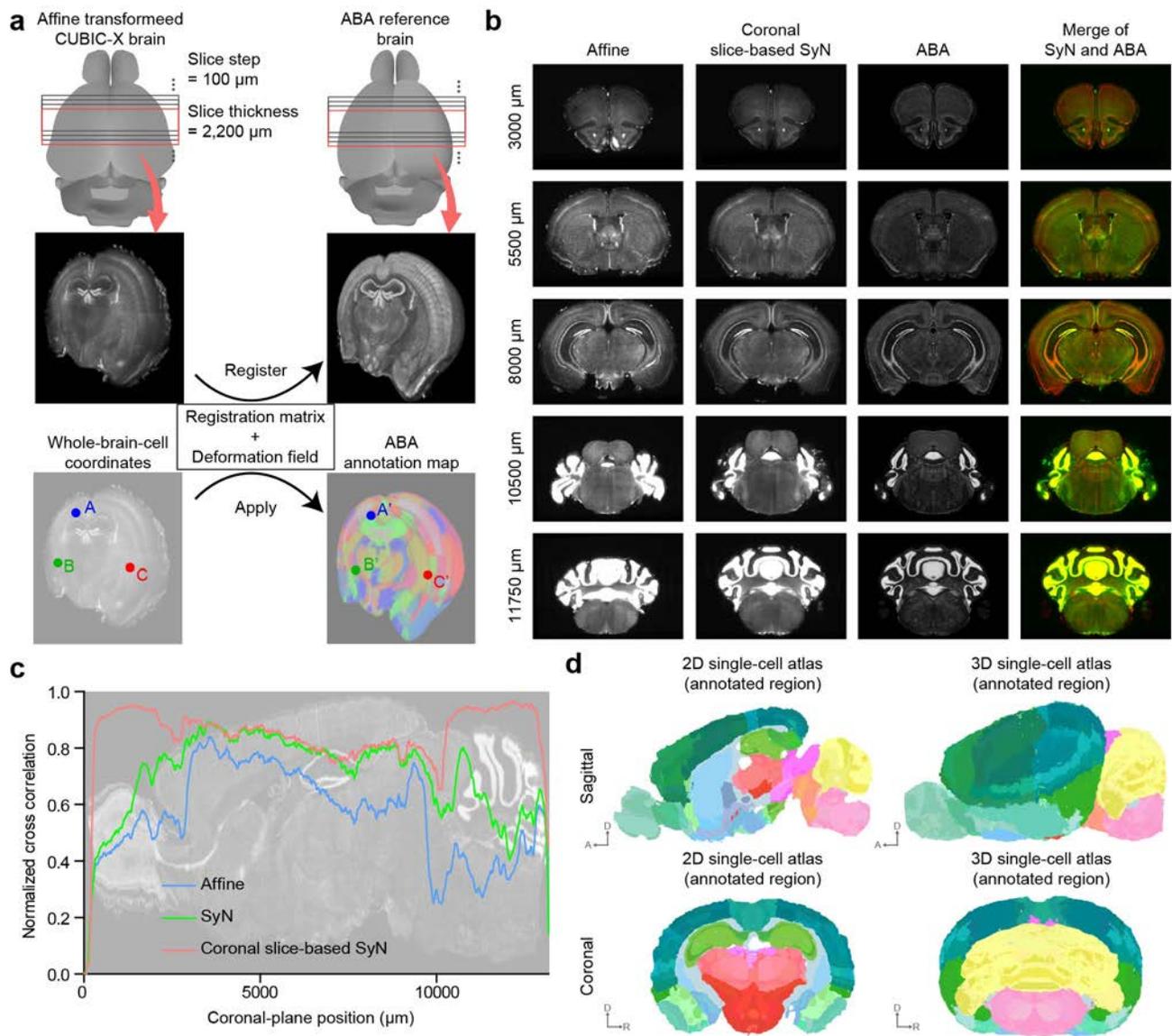
Background threshold was determined by the histogram of pixel intensity in each sub-region (1/100 of the original image) in the 2D cell-detection step. We determined the intensity less than  $2 \times$  mode as background intensity, and then extracted background signals (blue). Next, we calculated mean intensity and SD from the resulting histogram, and then determined the intensity less than 2 SD above the mean intensity as an image threshold (red). Scale bars indicate 200  $\mu$ m in the scale of an “after-CUBIC-X” brain.



### Supplementary Figure 13

The evaluation of accuracy for the automated detection of cell nuclei in a whole brain.

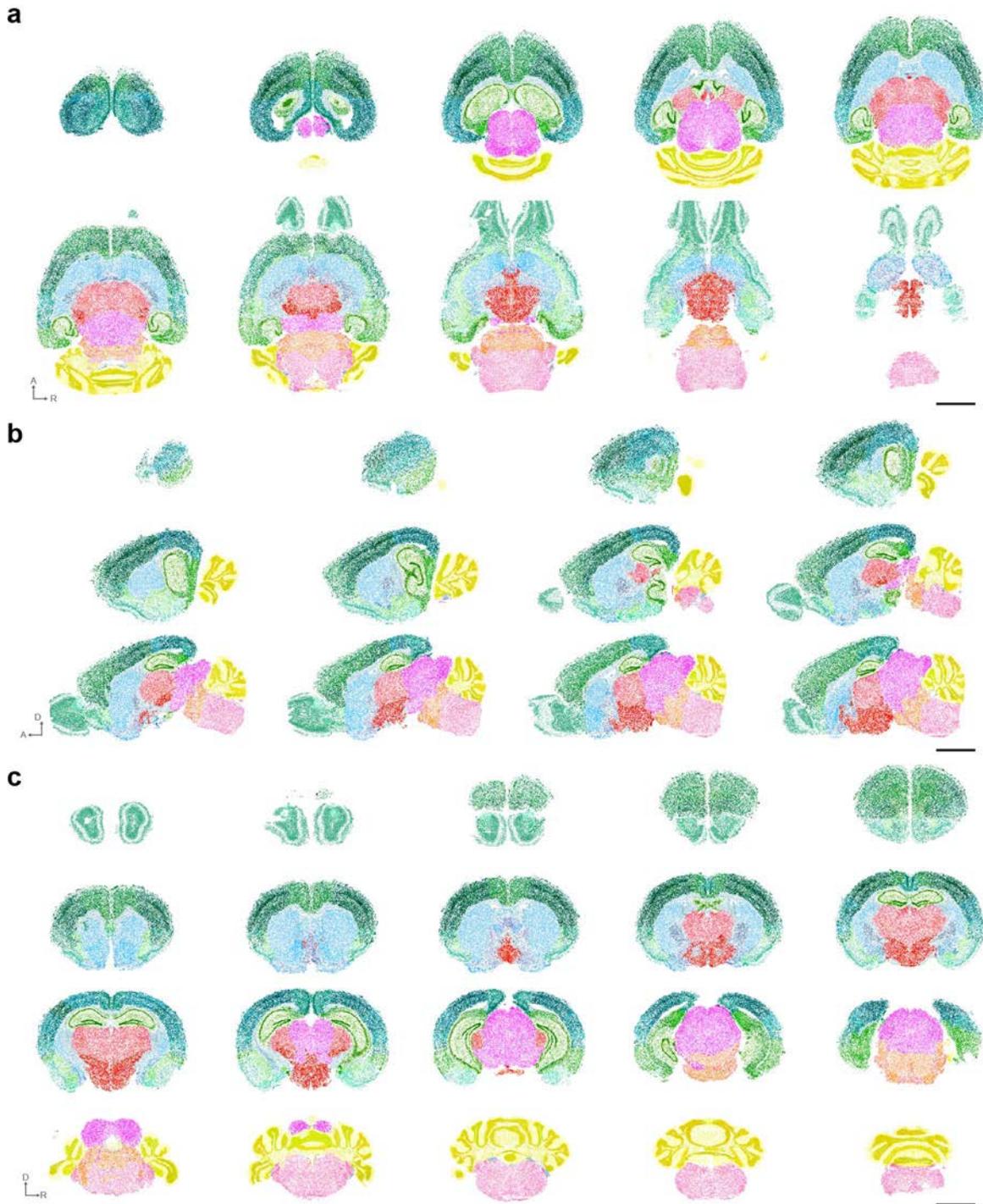
(a) The representative images of cell detection in sparsely populated areas. Because cell detection was performed in a three-dimensional manner, the cellular nuclei shown in single planes (magenta) were repeatedly detected in different axial positions. The lateral positions of detected cellular nuclei (green) in surrounding z-planes (1, 3, 5, or 7) were shown in each image. CTX, cerebral cortex; TH, thalamus; cc, corpus callosum; MOBgr, main olfactory bulb granular layer; MB, midbrain. Scale bars indicate 100  $\mu\text{m}$  in the scale of an "after-CUBIC-X" brain. (b) The representative images of cell detection in densely populated areas. DG, dentate gyrus; CBgr, cerebellum granular layer. For a and b, experiments were repeated more than ten times with independent brains. The representative images are shown. (c) Image comparison of low-illumination-NA condition (measured FWHM = 11.0  $\mu\text{m}$ , scanning step size= 5.0  $\mu\text{m}$ ) and high-illumination-NA condition (measured FWHM = 5.2  $\mu\text{m}$ , scanning step size= 2.0  $\mu\text{m}$ ). Horizontal and coronal sectional images of DG and CBgr are shown. Scale bars indicate 50  $\mu\text{m}$  in the scale of "after-CUBIC-X" brain. Experiments were repeated three times with different brain areas. The representative images are shown. (d) Comparison of the manually annotated nuclei with the automatically detected nuclei. Two people independently annotated the nuclei for manual annotation. Scale bars indicate 50  $\mu\text{m}$  in the scale of "after-CUBIC-X" brain. (e) Projected image of manually annotated nuclei and automatically detected nuclei in DG and CBgr. The projection of 25  $\mu\text{m}$  are shown. Scale bars indicate 50  $\mu\text{m}$  in the scale of "after-CUBIC-X" brain. (f) Details of the accuracy of the automated two-step cell-detection algorithm in DG and CBgr. (g) The accuracy of the cell detection in various brain areas. Red and blue lines indicate 0.9 and 0.8, respectively. For f and g, one and three volumetric images were used for each to evaluate accuracy.



### Supplementary Figure 14

Non-linear registration of coronal slices in a CUBIC-X brain to Allen Brain Atlas.

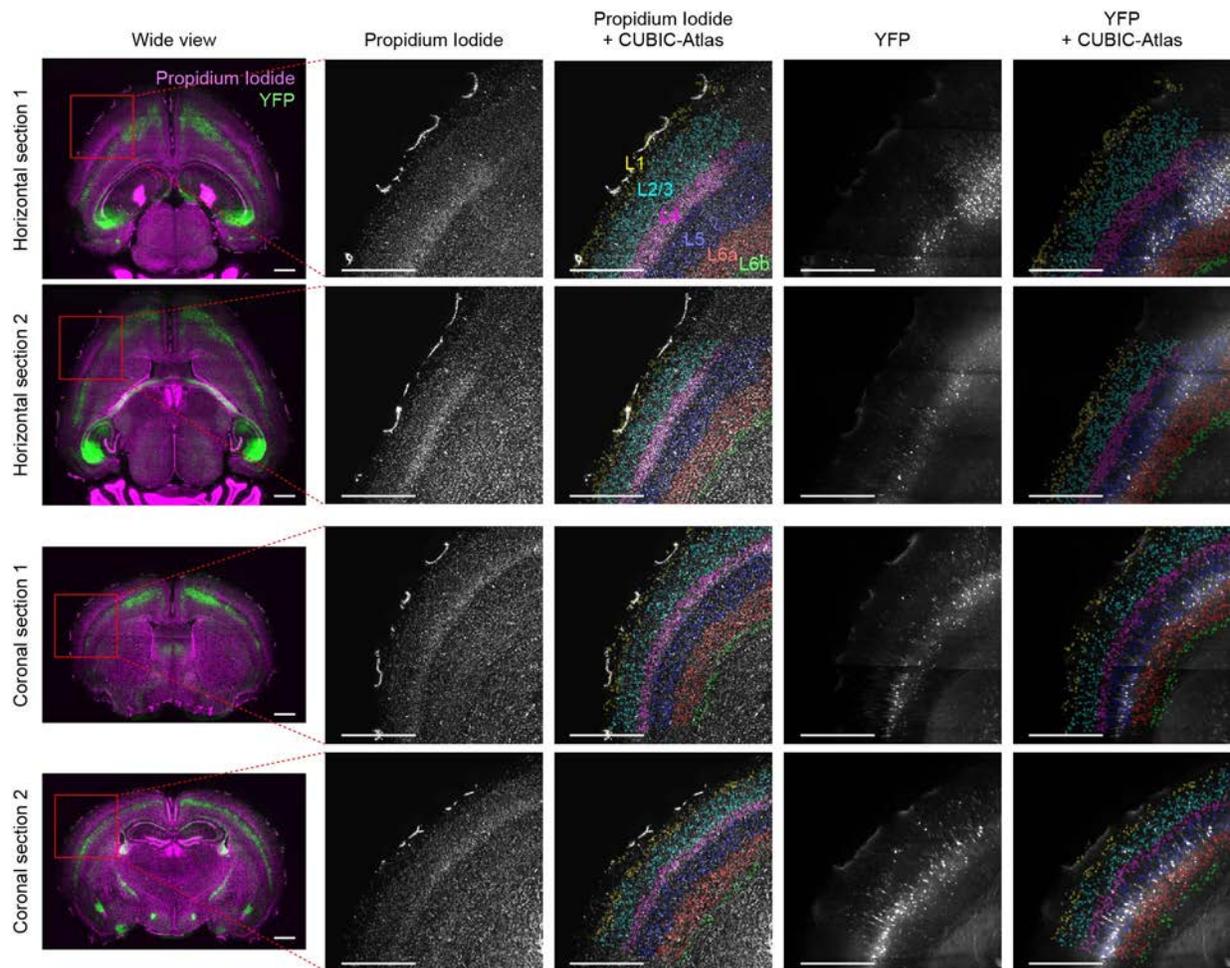
(a) A schematic diagram for the registration of a CUBIC-X brain to an Allen Brain Atlas (ABA) brain based on the virtual coronal slices. After registration of the CUBIC-X brain images by affine transformation, we generated about a hundred overlapping 3D coronal slices with a 2.2-mm thickness at a 100- $\mu\text{m}$  step both from the CUBIC-X and ABA brains. We then registered each coronal-slice image of the CUBIC-X brain onto the coronal-slice images of the ABA brain by non-linear transformation. (b) Coronal single-plane images of a CUBIC-X brain with affine transformation (“Affine”), a CUBIC-X brain with non-linear transformation based on coronal slices (“Coronal slice-based SyN”), an ABA brain (“ABA”), and merged coronal single-plane images of non-linearly transformed CUBIC-X (green) and ABA brains (red). (c) Normalized cross-correlations of a CUBIC-X brain against an ABA brain over entire coronal single-plane positions by straightforward affine transformation (“Affine”, blue), straightforward non-linear transformation (“SyN”, green), and non-linear transformation based on coronal slices (“Coronal slice-based SyN”, red). (d) Sagittal and coronal images of a 2D and 3D single-cell-resolution mouse brain atlas (CUBIC-Atlas) based on an 8-week-old C57BL/6J mouse brain. For non-linear transformation, symmetric normalization (SyN) in ANTs software was used (Online Methods).



**Supplementary Figure 15**

A single-cell-resolution mouse brain atlas (CUBIC-Atlas).

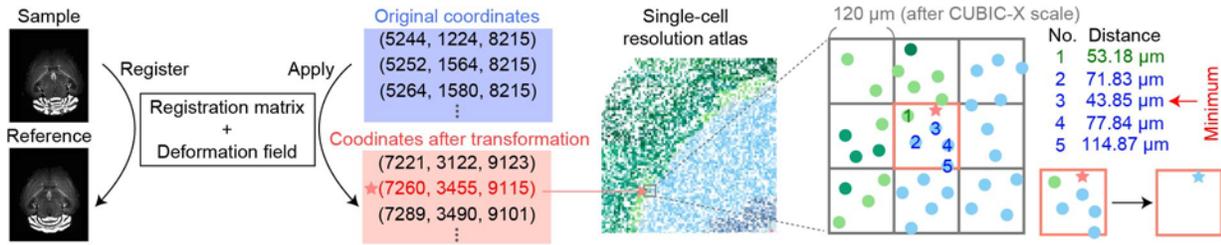
(a-c) Horizontal (a), sagittal (b) and coronal (c) single-plane images of CUBIC-Atlas. Each color indicates a corresponding anatomical area (see also Supplementary Table 3 for color). Scale bars indicate 5 mm in the scale of “after-CUBIC-X” brain. Each slice is shown at 1.495 mm (a), 0.845 mm (b) and 0.585 mm (c) step.



**Supplementary Figure 16**

Layers of Cerebral cortex in a CUBIC-Atlas overlaid with a PI-stained Thy1-YFP-H mouse brain.

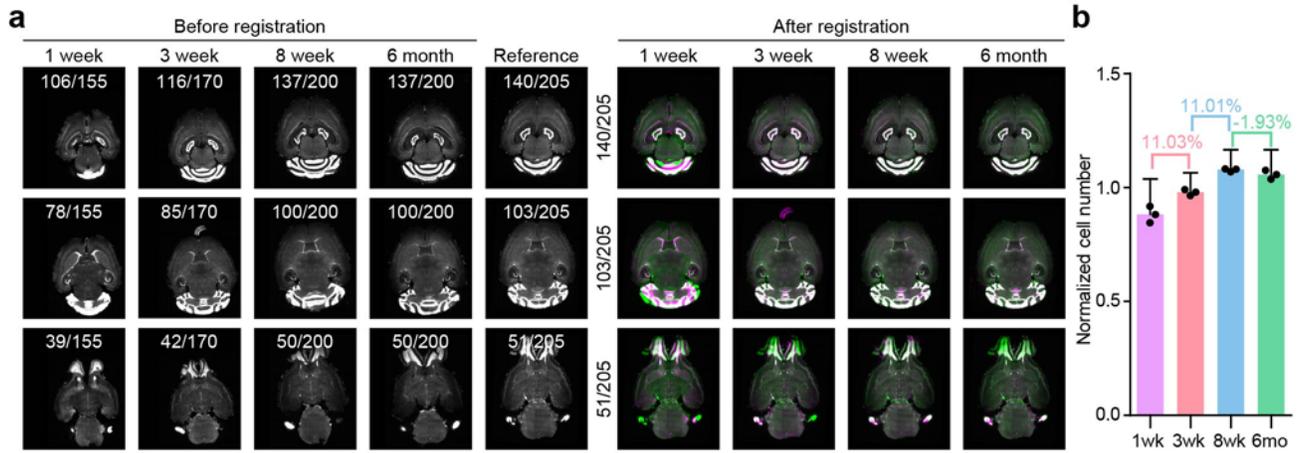
A PI stained Thy1-YFP-H mouse brain was expanded by CUBIC-X and imaged with the customized LSMF. The acquired volumetric images were downsampled, then non-linearly registered to a CUBIC-Atlas. The layer structures of CUBIC-Atlas were overlaid with the images of PI and YFP, respectively. Cerebral cortical layers of somatosensory areas, posterior parietal association areas, and visual areas of a CUBIC-Atlas were visualized with pseudo-color codes. L1, layer 1; L2/3, layer 2/3; L4, layer 4; L5, layer 5; L6a, layer 6a; L6b, layer 6b. Experiments were repeated twice with independent brains and the representative images are shown. Scale bars indicate 2000  $\mu\text{m}$  in the scale of an “after-CUBIC-X” brain.



### Supplementary Figure 17

Anatomical annotation with CUBIC-Atlas.

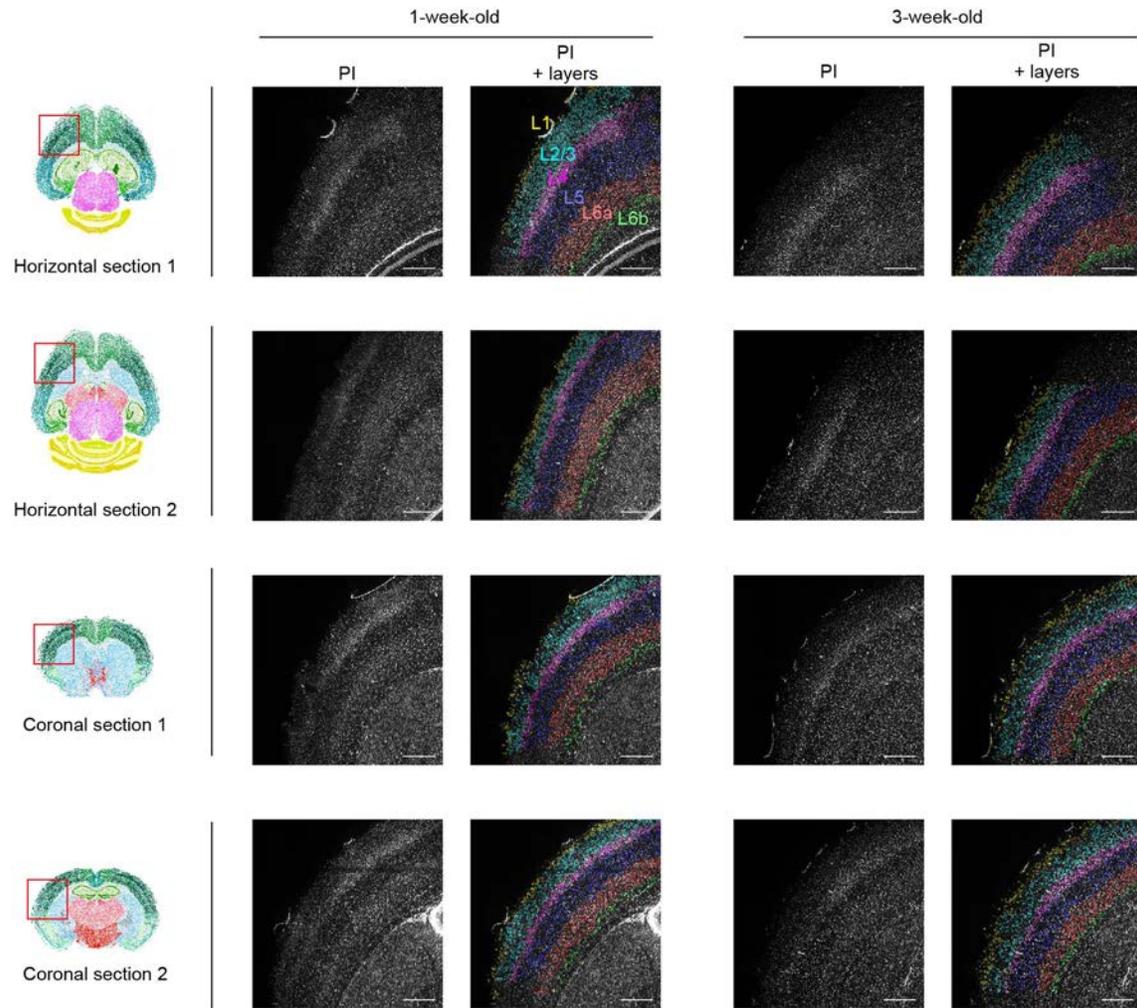
After registration, we detected the nearest neighbor cell as a reference in CUBIC-Atlas to the cell of interest in the obtained images. To reduce computational cost, we performed this search in the locally allocated grids (120 μm × 120 μm × 120 μm) partitioned from the atlas space. We annotated the corresponding region of the nearest neighbor cell in CUBIC-Atlas to a region of the cell of interest (Online Methods).



### Supplementary Figure 18

Whole-brain cell profiling over different developmental stages.

(a) Horizontal images of CUBIC-X brains of different aged mice before and after registration to CUBIC-Atlas. “Before-registration” images are shown on the left. Indicated numbers mean the horizontal slice positions. In the “After-registration” images, green shows the reference brain and magenta shows the registered brains. (b) Normalized cell number over different aged mice ( $n = 3$ ). The average  $\pm$  SD are shown.

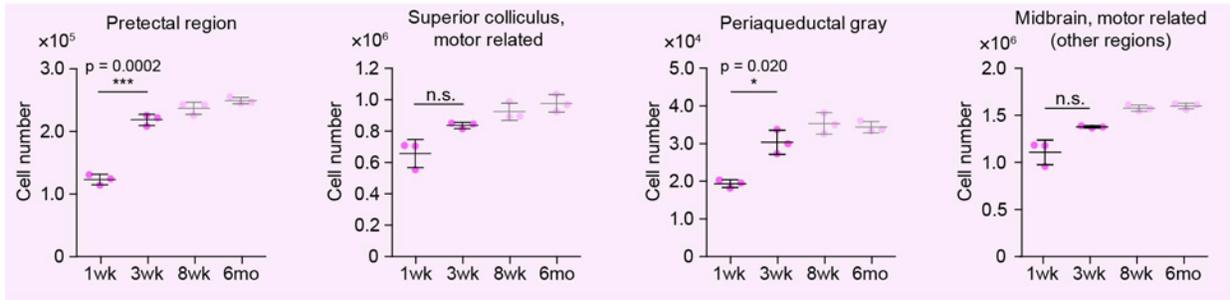


### Supplementary Figure 19

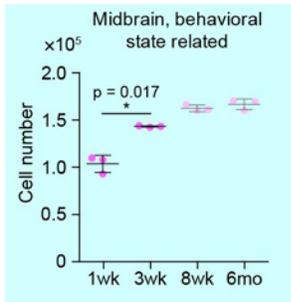
Mapping of 1-week-old and 3-week-old mice brains onto a CUBIC-Atlas.

PI-stained 1-week-old and 3-week-old mice brains were mapped onto and annotated with a CUBIC-Atlas based on an 8-week-old mouse brain. Cerebral cortical layers of somatosensory areas, posterior parietal association areas, and visual areas with annotations are visualized with pseudo colors. On the left, red squares indicate the approximate positions of the magnified brain regions shown in the right. Experiments were repeated three times with independent brains. The representative images are shown. L1, layer 1; L2/3, layer 2/3; L4, layer 4; L5, layer 5; L6a, layer 6a; L6b, layer 6b. Scale bars indicate 1000  $\mu\text{m}$  in the scale of an “after-CUBIC-X” brain.

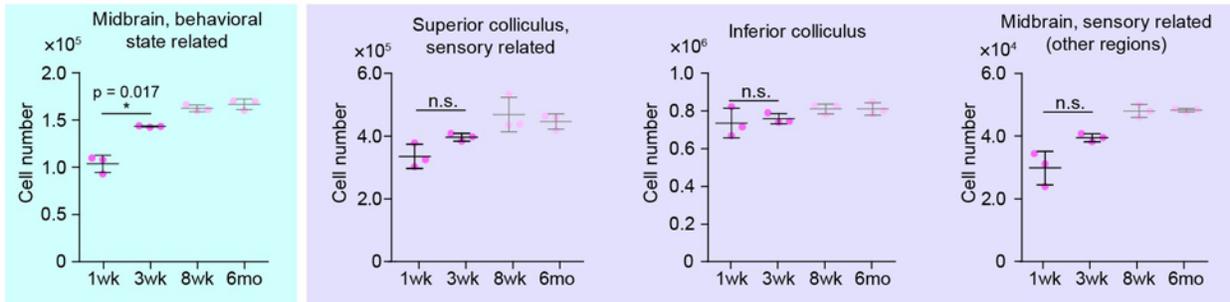
### Midbrain, motor related



### Midbrain, behavioral state related



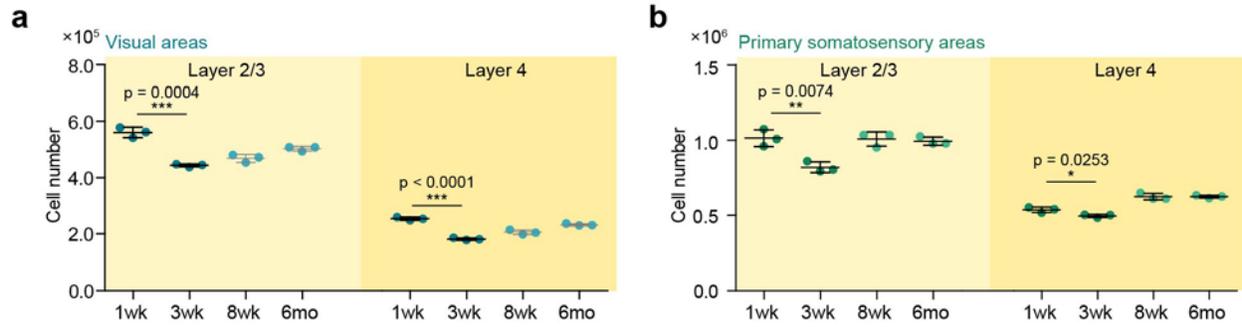
### Midbrain, sensory related



## Supplementary Figure 20

Whole-brain cell profiling over different developmental stages in midbrain.

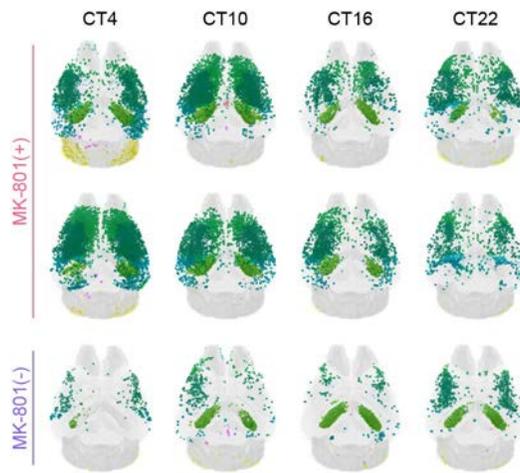
The significant increase of the total cell number in the midbrain ( $n = 3$ ,  $***p < 0.001$ ,  $*p < 0.05$ , independent two-tailed t-test). Background color represents each classification of the midbrain (Pink for motor-related areas, blue for behavioral-state-related areas and purple for sensory-related areas, respectively). The average  $\pm$  SD are shown.



### Supplementary Figure 21

Whole-brain cell profiling over different developmental stages in cerebral cortex.

The significant decrease in the total cell number in visual areas (a) and primary somatosensory areas (b) of the cerebral cortex areas ( $n = 3$ , \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ , independent two-tailed t-test). The average  $\pm$  SD are shown.



### Supplementary Figure 22

Probabilistic mapping of pharmacologically stimulated mouse brains expressing an Arc-dVenus reporter onto a CUBIC-Atlas.

Temporal variation of distribution of hyper-activated cells in the whole brain. Original image data were obtained in Tatsuki et al. (Online Methods), and re-analyzed in this study. Arc-dVenus expressing cells in mice with or without the chronic administration of MK-801 are shown. The color represents an anatomical area of the detected cells. All brain samples used in the experiment are shown except the brain samples shown in Fig. 7b.