CUBIC-Cloud: An Integrative Computational Framework Towards Community-driven Whole-Mouse-Brain Mapping

Tomoyuki Mano^{1,2}, Ken Murata^{3,4}, Kazuhiro Kon⁵, Chika Shimizu², Hiroaki Ono⁵, Shoi Shi^{2,5}, Rikuhiko G. Yamada², Kazunari Miyamichi^{3,4}, Etsuo A. Susaki^{2,5}, Kazushige Touhara^{3,4,6}, and Hiroki R. Ueda^{1,2,5,*}

¹Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-0033, Japan

²Laboratory for Synthetic Biology, RIKEN Center for Biosystems Dynamics Research, Osaka 565-5241, Japan

³Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan

⁴ERATO Touhara Chemosensory Signal Project, JST, The University of Tokyo, Tokyo 113-8657, Japan

⁵Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan

⁶International Research Center for Neurointelligence (WPI-IRCN), UTIAS, The University of Tokyo, Tokyo 113-0033, Japan

*Correspondence: uedah-ty@umin.ac.jp

ABSTRACT

Recent advancements in tissue clearing technologies have offered unparalleled opportunities for researchers to explore the whole mouse brain at cellular resolution. With the expansion of this experimental technique, however, a scalable and easy-to-use computational tool is in demand to effectively analyze and integrate whole-brain mapping datasets. To that end, here we present CUBIC-Cloud, a cloud-based framework to quantify, visualize and integrate whole mouse brain data. CUBIC-Cloud is a fully automated system where users can upload their whole-brain data, run analysis and publish the results. We demonstrate the generality of CUBIC-Cloud by a variety of applications. First, we investigated brain-wide distribution of PV, Sst, ChAT, Th and Iba1 expressing cells. Second, A β plaque deposition in AD model mouse brains were quantified. Third, we reconstructed neuronal activity profile under LPS-induced inflammation by c-Fos immunostaining. Last, we show brain-wide connectivity mapping by pseudo-typed Rabies virus. Together, CUBIC-Cloud provides an integrative platform to advance scalable and collaborative whole-brain mapping.

Keywords: Tissue clearing, Light-sheet fluorescence microscopy (LSFM), Cloud computing, Parvalbumin, Somatostatin, Choline acetylase, Tyrosine hydroxylase, microglia, inflammation, Alzheimer's disease, c-Fos, LPS, Rabies virus, Kiss1, Arcuate nucleus of the hypothalamus

Introduction

Massive and collective observation of complex systems (often referred to as omics approaches) is the driving force of modern biology. It would also be true for studying highly-evolved mammalian brains, where a complex system of intricately connected cells gives rise to intelligent behaviors. In particular, comprehensive approaches to identifying the properties of every single cells *in situ* within this complex system would be pivotal. Recent advancements in tissue clearing technology have brought new breakthroughs in this landscape [1, 2, 3, 4, 5, 6, 7, 8, 9]. Combined with light-sheet fluorescence microscopy (LSFM) [10] and genetical, viral and immunohistochemical labelling techniques, tissue clearing now enables high-speed volumetric imaging of mammalian (most prominently mouse) brains at cellular resolution [7, 11, 12, 13, 14, 15, 16, 17]. Built on top of these technological advancements, we recently reported the construction of CUBIC-Atlas [18], a 3D mouse brain atlas with single cell resolution, where all of the cells in the brain (amounting to approximately 0.1 billion) were digitally analyzed and recorded.

These scientific advancements encourage us to conceive a future where whole-brain mapping projects, which conventionally required institution-scale resources and efforts, can be carried out by individual laboratories, or even by a single researcher [19, 20, 21, 22]. In this regard, the current technological stage can be thought of as parallel to the dawn of genome sequencing technology in early 2000s. In genome science, the importance of the data repository cannot be understated. The emergence of the database to browse and search genomes (such as UCSC Genome Browser [23]) played a critical role in integrating data collected in numerous sites across the globe. Such distributed collaboration prompted a rapidly growing coverage of various organisms and individuals, pioneering the data-driven discoveries of gene functions and new therapeutics. In neuroscience, several large-scale mouse brain datasets have been constructed, such as Allen Mouse Brain Atlas [24] and Brain Architecture Project [25], primarily using serial sectioning tomography methods. However, a common platform has been yet to appear that offers the opportunity for the community to submit and share new data, embracing the tissue clearing and rapid brain scanning techniques. Based on these considerations, we suggest that it is now possible to construct the community-supported mouse brain data repository, borrowing collaborative ideas from genome sciences.

By referring to the previous image analysis pipelines for tissue clearing samples [26, 27, 28, 16, 29], we postulated that the following elements should be considered in composing a framework for whole mouse brain mapping. First, the reference brain (equivalent to the template sequence), to which all brain data are aligned, is necessary. We think that CUBIC-Atlas would play a central role in addressing this challenge. Second, the framework should be constructed around the research community, which allows researchers to submit the data, as well as to browse and search the brains in the previous studies. Third, because of the complexity and the large size of the whole-brain data sets, the framework should be equipped with a toolkit to visualize and quantify the data to assist intuitive understanding. Fourth, these software tools should offer superior accessibility and usability to the end users, without requiring specialized expertise in programming or powerful computer resources.

In this paper, we present a computational framework for single-cell-resolution whole-mouse-brain analysis, named CUBIC-Cloud. We built CUBIC-Cloud on top of the latest cloud computing technologies. After identifying single cells from 3D image stack, users can upload the structure image and detected cell list to the CUBIC-Cloud server. CUBIC-Cloud server automatically aligns individual brains with the CUBIC-Atlas, and construct the user's own mouse brain database. CUBIC-Cloud also offers GUI tools to perform various kinds of quantification tasks. Uploaded brains can be interactively visualized using 3D brain viewer. The scientific results obtained thereby can be easily published in the CUBIC-Cloud's public repository, to allow other researchers to view the data. CUBIC-Cloud is hosted at <https://cubic-cloud.com>.

After describing the software architecture, we extensively demonstrate the capability and generality of CUBIC-Cloud framework by analysing over 50 whole mouse brains, covering four important application domains: (1) investigating the distribution of targeted cell types ("cell-type mapping"), (2) quantifying the pathological state by disease markers, (3) reconstructing the neuronal activity profile by imaging the expression of c-Fos ("activity mapping"), and (4) deciphering brain-wide connectivity using Rabies virus tracers ("circuit mapping"). With these applications, we established a general framework to effectively integrate whole-brain mapping experiments, offering new opportunities for data-driven discoveries in neuroscience.

Results

CUBIC-Cloud whole-brain analysis framework

To design a unified workflow for whole-mouse-brain analysis, we constructed CUBIC-Cloud on top of our previously published tissue clearing protocols and brain mapping strategies [7, 18, 30]. The workflow of whole-brain analysis using CUBIC-Cloud is illustrated in Fig. 1. Briefly, the workflow divides into (1) tissue clearing and image collection, (2) single cell detection, and (3) uploading the data to CUBIC-Cloud, where brain registration, brain-wide quantification and visualization are performed.

In this study, mouse brains were cleared by CUBIC-L and CUBIC-R+ reagents [30] ("step 1" in Fig. 1). In addition, nuclear counter-staining (for brain registration) and, optionally, immunostaining were applied by following the CUBIC-HV protocol [31] (Methods). In most experiments, cleared brains were imaged using macro-zoom LSM with 6.5 μm isotropic XYZ voxel size ("step 2" in Fig. 1, Methods). The transparency of the cleared brain was quantitatively evaluated by imaging fluorescent beads embedded in the tissue, validating homogeneous image quality throughout all brain regions (Supplementary Fig. 2, Methods). Of note, CUBIC-Cloud framework should be general enough to be used with other clearing methods, as long as the equivalent level of tissue transparency and morphology preservation is achieved.

After obtaining the 3D image stack, labeled cells were detected using ilastik software [32, 33] ("step 3" in Fig. 1). ilastik uses supervised machine learning (random forest algorithm) to classify pixels into multiple classes (such as cells and backgrounds), which performed robustly well even when cells from different brain regions presented diverse brightness and morphology. We created a Python script which wrapped ilastik to automate the workflow and add some custom analysis routines (Methods). The accuracy of cell counting is described in the corresponding sections. By this cell counting procedure, raw raster images were converted into a discrete ensemble of cellular points, or point cloud, where positions, volumes and fluorescence intensities of all labeled cells in the brain were recorded. Of note, users can opt to use their own cell detection routine, as long as the output follows the required format by CUBIC-Cloud.

After cell detection, the rest of the analysis were carried out by CUBIC-Cloud. Users can go to the CUBIC-Cloud web site (<https://cubic-cloud.com>) and upload cellular point cloud and corresponding structure image (nuclear staining channel) to the server. CUBIC-Cloud is deployed on Amazon Web Service (AWS) using so-called serverless architecture. Serverless architecture allows the cloud to dynamically scale its capacity based on the computational load, enabling the cloud to handle practically unlimited number of tasks in parallel while minimizing the idling time. The detailed implementation is described in Methods (also see Supplementary Fig. 1). Once the brain data is uploaded to the CUBIC-Cloud server, it is automatically sent to the preprocessing task queue, which is powered by ECS and EC2 on AWS. Preprocessing task first performs brain registration to align submitted brain with the CUBIC-Atlas ("step 4" in Fig. 1). CUBIC-Cloud uses the symmetric diffeomorphic image registration (SyN) algorithm implemented in ANTs library [34]. Using the nuclear staining channel, SyN iteratively optimized the warp field to maximize the normalized cross-correlation (NCC), which measured the similarity between two images (Supplementary Fig. 3i,j, Methods). After registration, cell coordinates are transformed to the CUBIC-Atlas space, and all cells are given anatomical region IDs (following the Allen Brain Atlas CCF v3 [24]).

By repeating the above procedures, users can construct their own brain database in CUBIC-Cloud. As the database grows, users can search brains using tags attached to them, such as cell labels and project names, along with other metadata. To perform quantitative analysis of the brains in the database, CUBIC-Cloud offers a "notebook", a feature that allows users to create various kinds of graphs with graphical user interfaces (GUIs) ("step 6a" in Fig. 1). To give some concrete examples, many of the figure items presented in this paper were created by CUBIC-Cloud's notebook.

CUBIC-Cloud also offers an interactive 3D brain viewer, a feature called "studio" ("step 6b" in Fig. 1). Here, brain data is visualized in point cloud format, where the whole brain is rendered as an ensemble of discrete cellular points, each carrying biological attributes such as protein expression levels. Point cloud format is more efficient to deliver data to the remote clients than sending raw raster data, while carrying the essential biological information. The viewer runs on standard web browsers, and the point clouds are adaptively queried from the server upon the viewer's camera movement, enabling to interact with millions of cells in real time.

The last essential component of CUBIC-Cloud is sharing and publishing. Users can share their whole-brain data with specific users, such as research collaborators, and grant access to the data. Users can opt to publish their data in the CUBIC-Cloud's public repository. Once published, any users can view the brain. The share and publish capability are also supported for the notebooks and studios. Therefore, users can transparently show their analysis results to the research community. To demonstrate this concept, all of the brain data investigated in this study is deposited on CUBIC-Cloud public repository, as well as the notebooks and studios that performed the analysis. Together, CUBIC-Cloud offers cloud-native and integrative software solution for whole-mouse-brain mapping.

Whole-brain analysis of PV, SST, ChAT and Th expressing cells

In the following, we use CUBIC-Cloud in several neuroscience applications to show the generality of the proposed framework. As the first application, we attempted to quantify the whole-brain distribution of distinct cellular subtypes that express the following markers, respectively: parvalbumin (PV), somatostatin (SST), choline acetyltransferase (ChAT), tyrosine hydroxylase (TH), and ionized calcium-binding adapter molecule 1 (Iba1). Whole-brain scale analysis

of PV, SST and ChAT expressing neurons were previously reported using knock-in transgenic mouse line with Cre recombinase and fluorescent proteins [25, 35, 36]. However, no such whole-brain scale analysis have been done so far using immunostaining and tissue clearing. Immunostaining offers several orthogonal advantages over genetical approaches, such as the capability to assess the absolute expression amount, and thus of significant value. Therefore, we used CUBIC-HV protocol [31] to label the whole adult brain tissue with PV, SST, ChAT, TH and Iba1 antibody, respectively (Methods).

The obtained images contained strong and soma-localized signals as well as moderately bright fibrous structures. Here, we attempted to quantify the number of immunoreactive cell bodies, and thus, the ilastik classifier was trained to isolate cell soma and reject fiber-like signals. The accuracy of the cell detection was extensively evaluated (Supplementary Fig. 3c-g). Our cell detection demonstrated 75-95% precision and 60-85% sensitivity for the regions evaluated, which we concluded was accurate enough to perform quantitative analysis.

Whole-brain views of the investigated cell types are shown in Fig. 2a-f (rendered by CUBIC-Cloud's 3D viewer). Here, cells are pseudo-colored, reflecting the fluorescence intensity of the immunostaining. In our analysis, the total number of detected cells were: $(6.1 \pm 0.7) \times 10^5$ (PV, $n = 4$); $(6.7 \pm 0.3) \times 10^5$ (SST, $n = 4$); $(6.5 \pm 0.2) \times 10^4$ (ChAT, $n = 4$); $(6.9 \pm 1.2) \times 10^4$ (TH, $n = 4$); $(2.72 \pm 0.14) \times 10^6$ (Iba1, $n = 7$) (Supplementary Table 2, Supplementary Table 3). In this paper, mean $\pm$ standard deviation is used unless otherwise specified. The relative ratio of each cell type in all brain regions except isocortex is shown in Fig. 2g. Here, the ratio was calculated by dividing the detected cell count by the total number of cells reported in CUBIC-Atlas [18]. We will discuss the brain-wide distribution of these cell types in more detail in Methods.

Within the isocortex, PV neurons were most densely populated in somatosensory and auditory areas, whereas there were relatively sparse populations in medial frontal and lateral association areas (Fig. 2h). Compared to PV, SST neurons were more homogeneously observed in the isocortex, consistent with the Cre-based study by [25]. Laminar density analysis revealed that there were almost no PV cells in layer 1, and that cell density reaches its peak in layer 4 or 5 for both PV and SST (Supplementary Fig. 4a-d), successfully recapitulating the previous observations [37, 38]. There were sparse populations of ChAT neurons, whose ChAT expression levels were very low (Supplementary Fig. 4e,f). These ChAT neurons were most dense in layer 2/3 or 4, corroborating the previous observation [37]. For Th neurons, no significant population above false-positive detection level were found in the isocortex. Iba1 expressing cells were ubiquitously observed in the isocortex, with similar density across layers.

From our cell-type mapping data, certain cell types were clustered in specific areas in the brain, offering the clues to delineate the region boundaries. For instance, we observed that ChAT neurons were densely populated in and around the LDT (for the list of region acronyms, see Supplementary Table 1). From this cluster of ChAT neurons, we defined the polygonal boundary enclosing these cells using alpha shape algorithm (Fig. 2i, Methods). We performed same analysis targeting TH neurons in and around the LC (Fig. 2j). The polygonal boundaries defined by ChAT+ and TH+ neuron cluster, respectively, were neighboring with each other with a small overlap (Fig. 2k). We then evaluated the overlap between polygonal boundaries using the Dice's metric (Methods). The result showed about 70% overlap between ChAT-ChAT and TH-TH pairs (Fig. 2l). Furthermore, the overlap of ChAT-TH pair was 2.84 ± 0.9 %. These results support the accuracy of our registration method with an orthogonal evaluation metric, which is largely independent from information present in the nuclear staining channel. This result thus offers the possibility to accurately delineate the brain regions by collecting more cell-type mapping data in CUBIC-Cloud.

Lastly, to assess whether our analysis pipeline can accurately quantify the expression amount of the proteins, we evaluated the expression of Iba1 under artificially induced inflammation state, because Iba1 expression is known to be correlated with microglial activation [39]. In our experiment, 1 mg/kg of lipopolysaccharides (LPS), a purified extract of the outer membrane of Gram-negative bacteria, was administered to mice via intraperitoneal (i.p.) injection. The brains were sampled twenty-four hours after injection (Methods). In the isocortex, no subregions displayed significant change in the cell density ($p > 0.1$, Welch's t-test, Supplementary Fig. 4i), while the mean expression amount per cell was increased significantly in all subregions ($p < 0.05$, Welch's t-test, Supplementary Fig. 4j). The observation that cortical microglia does not proliferate but increases its Iba1 expression upon LPS administration was reported in [40], and our results confirmed that it is globally true in all cortical areas. Outside the isocortex, most regions showed increases in Iba1 expression amount (Supplementary Fig. 4k), with varying degree of change. For instance, Iba1 expression amount barely changed in cerebellum. On the other hand, we confirmed that increase in Iba1 expression was markedly high in the SFO and IO, which are part of circumventricular organs (CVOs) having highly permeable blood-brain barrier (BBB) [41].

Whole-brain analysis of A β plaque accumulation in AD model mouse brain

We next applied CUBIC-Cloud analysis framework to quantitatively understand the pathological state of the Alzheimer's disease (AD) model mouse. To demonstrate this, the whole brain from an *App*^{NL-G-F/NL-G-F} AD model mouse [42] (9- to 10-months-old) was cleared and stained by anti-A β antibody ($n = 4$, Methods). In LSFM images, A β plaques were observed as dim blobs often accompanying a bright spot at the core. On the other hand, no plaque staining pattern was observed in the control wild-type mouse brain (9- to 10-month-old, $n = 3$, data not shown).

We first quantified the density (number of individual plaques per volume) and the volume ratio (computed as (total plaque volume in the region)/(region volume)). In both metrics, A β plaque amounts were highest in the cerebral cortex

and cerebral nuclei, and relatively lower amount of plaques were observed in the brain stem and cerebellum (Fig. 3a,b). The effective radius of the plaque (computed as $r = \{3V/(4\pi)\}^{1/3}$ where V is the plaque volume) tended to be larger in the isocortex and hippocampus and smaller in cerebellum (Fig. 3c). Within the isocortex, a relatively stronger accumulation of A β were observed in visual and auditory areas, whereas plaques were relatively sparse in medial frontal areas (Fig. 3d,e). Layer-wise abundance of A β plaque showed concave profile, with its peak in layer 4 (Fig. 3f). The whole-brain cartoon heatmap showing the A β volume ratio is shown in Fig. 3g. In the brain stem, the plaque volume ratio was typically 0.5% to 1.0%. Some brain stem regions, however, showed notably larger or smaller amount of A β accumulation. For example, SNr and VMH had relatively higher amount of A β compared to neighboring regions (Fig. 3h,i). ARH, right next to VMH, had almost no A β plaques. We also observed that the regions around the ventricles showed relatively lower amount of plaques, including TRS, DTN, MH, LH and PVT (Fig. 3g, j-l).

Whole-brain analysis of c-Fos expression underling pharmacological sleep induction by LPS

The next important application domain of CUBIC-Cloud is to reconstruct the neuronal activity profile by imaging the protein expressions of immediate early genes (IEGs) such as c-Fos. Such automated analysis would allow comprehensive identification of cellular clusters that underlie an animal's behavioral phenotype [7, 28, 43, 44]. Reciprocally, one could define an animal's phenotype in a bottom up manner based on the activity pattern of neuron ensembles. Here, we focused on the effect of LPS. Phenotypically, LPS induces acute sleep in mice, and we thought that IEG-based activity reconstruction would be suitable to track relatively slow dynamics of wake-sleep cycles.

Administration of 150 μ g/kg LPS to mice via i.p. injection at CT = 14 caused acute immune response accompanied by prolonged sleep duration, as confirmed by SSS measurement (Supplementary Fig. 5a-c, Methods). In a replicate experiment, brains were sampled 2-3 hours after LPS injection, and the whole-brain c-Fos expression profile was analyzed (Methods). The accuracy of the c-Fos cell detection is shown in Supplementary Fig. 3h. Fig. 4a shows the whole-brain 3D rendering of the detected c-Fos expressing cells (also see Supplementary Table 5). We comprehensively searched for the activated or repressed brain regions by both region-wise and voxel-wise statistical analysis (Supplementary Fig. 5d). Our analysis revealed that c-Fos expression in some of the isocortical areas was reduced (Fig. 4b,c), which included motor and somatosensory areas, presumably reflecting the mouse's resting state. We also found that some distinct brain nuclei were activated by LPS. Among those, the most notable regions included the BST, PVH, PVT, CEA, PB, NTS and DMX (Fig. 4d).

In our result, the specific part of the BST (the oval region; ovBST) was strongly activated by LPS (Fig. 4d). Indeed, according to a recent study, ovBST is responsible for the inflammation-induced anorexia, and it receives inputs from CEA and PB [45]. Our result was able to successfully identify elevated c-Fos expressions in these spatially separated yet functionally related neurons.

We also observed heterogeneous c-Fos activation in PVT. In terms of the number of c-Fos+ cells, the increase in number was more pronounced in the posterior PVT (pPVT) than the anterior PVT (aPVT) (Fig. 4e-g). In terms of the expression level, both pPVT and aPVT showed similar level of increase (Fig. 4h,i). Recently, Gao et al. [46] identified two classes of distinct neurons in PVT. Type I neurons, densely located in the pPVT, responds to aversive stimuli. On the other hand, type II neurons, dominantly located in the aPVT, become silent upon aversive stimuli. It is also reported that the type II neurons were active during sleep. In the pPVT, our observation aligns with the insight by Gao, where the activated population was likely Type I neurons. In the aPVT, our result might reflect the mixed response of the type II neurons, where aversive inflammatory stimuli and induced sleep were both present. It should also be noted that Type I neurons in the pPVT project to CEA, ILA and ACB. We indeed observed that ILA and ACB were weakly activated (Supplementary Fig. 5d).

Whole-brain Analysis of Input Cells Projecting to ARH^{Kiss1+} Neurons

As the third application domain of CUBIC-Cloud, we show brain-wide connectivity analysis using pseudo-typed Rabies virus (RV). To demonstrate this application, we focused on a population of neurons that secrete kisspeptin (a neuropeptide encoded by *Kiss1* gene) located in the ARH, hereafter termed as ARH^{Kiss1+}. Those neurons were shown to play an important role in reproduction behavior in mammals by regulating pulsatile release of gonadotrophin-releasing hormone (GnRH) at around 0.3 to 1.0 pulses per hour [47]. Intriguingly, the pulse frequency changes through estrus cycle in females but not in males. As such, we comprehensively investigated the neural inputs to ARH^{Kiss1+} neurons to search the mechanism of pulse generation/modulation on the basis of neural circuitry.

To achieve cell-type specific targeting of virus infection, we used the Cre/loxP system and RV trans-synaptic tracing combined with Cre-dependent AAV vectors [48] (Fig. 5a, Methods). After virus injection, brains (both male and female) were cleared by CUBIC reagents and analyzed by CUBIC-Cloud pipeline (Methods). We first checked the distribution of starter cells (GFP+ and mCherry+) to ensure that the injection was successful and the starter cells were well confined within ARH (Fig. 5b). In the present study, the criteria in selecting successful samples was defined as more than 45% of starter cells were localized in ARH or PVp. (Note that area annotated as PVp in the Allen Brain Atlas belongs to a part of ARH in the Paxinos atlas [49].) With this criteria, out of 20 injections, $n = 3$ and $n = 4$ brains were assessed as successful for male and female, respectively (Fig. 5b).

Fig. 5c shows the whole-brain overview of all input (GFP+ and mCherry-) cells. The accuracy of the GFP+ cell detection is shown in Supplementary Fig. 3b. Our quantitative analysis identified $(3.1 \pm 0.5) \times 10^4$ input cells in the male

1 brain ($n = 3$), the majority of which ($> 85\%$) were located within the hypothalamus (Fig. 5d, Supplementary Table 6). As
2 is shown in Fig. 5e, ARH^{Kiss1+} neurons receive inputs from dozens of discrete structures throughout the forebrain
3 and brainstem, including the striatum (LS), pallidum (BST), thalamus (PVT), hypothalamus (MPO, MPN, AHN, PVH,
4 DMH, VMH and PH), hippocampal formation (HATA and SUB), midbrain (MRN and PAG), and pons (PB). Remarkably,
5 extremely sparse populations, only a few dozens of cells per region, were reproducibly identified (Fig. 5f,g). In terms of
6 cell density, those populations were often equivalent to less than 10 cell/mm³, which could be easily overlooked with
7 slice-based approaches. These sparse populations were not reported in the past literature [50].

8 We next performed statistical analysis comparing the number of input cells between male and female brains
9 (Methods). Overall, binary connectivity differences (i.e. zero in one sex and some finite number in the other sex) were
10 not observed (Supplementary Fig. 6a). Weak differences were suggested in LS, MPO, MPN and AVP (Supplementary
11 Fig. 6b-h, Methods), which are neighboring with each other. The difference was most pronounced in LSr, which is
12 known to inhibit the lordosis behavior during mating interactions [51]. The sexually dimorphic circuit from LSr to PAG is
13 reported, where female brains contain more neurons in LSr that project to PAG [52]. Our result suggest that LSr sends
14 sexually dimorphic projection to ARH^{Kiss1+}. The identities of these populations can be fully characterized in the future
15 studies.

Discussion

In this study, we presented an integrated computational framework for single-cell-resolution whole-brain analysis, named CUBIC-Cloud. Inspired by the data scientific platforms developed in genomics, we postulated that the framework should provide (1) a reference brain atlas and automatic mapping to the reference, (2) open opportunities for the research community to contribute new data, (3) a toolkit to visualize and quantify data, and (4) easy and universal accessibility. As we have shown in this study, CUBIC-Cloud addressed these requirements by designing a new software stack embracing the latest cloud technologies, widely available for researchers in neuroscience (Fig. 1). Users can access CUBIC-Cloud at <https://cubic-cloud.com>.

After describing the software infrastructures offered by CUBIC-Cloud, we extensively evaluated and validated the accuracy and reproducibility of the proposed framework by various applications. First, we quantified the exact number of various cell types, including PV, SST, ChAT, Th and Iba1 expressing cells (Fig. 2). We were surprised by the small variation between individual animals in terms of labelled cell numbers (quantified SD usually less than 10%), which not only highlighted the accuracy of our analysis but also implied the intricate regulation on cell proliferation in the brain. We further showed that CUBIC-Cloud can also be used to quantitatively understand the pathological state of the AD model mouse (Fig. 3). Together, these demonstrations encourage the future application of CUBIC-Cloud to construct the datasets of the disease models and the drug effects, where decreases or increases of particular markers in specific brain regions may correlate with disease progressions.

We have also explored the application of CUBIC-Cloud to reconstruct the whole-brain neuronal activity profile using c-Fos immunostaining. Our whole-brain analysis showed that administration of LPS, which induces sleepiness, repressed the c-Fos expression in some of the cortical areas, while it activated several distinct brain regions (Fig. 4). For example, our analysis revealed that a distinct subpopulation in the ovBST strongly responded to LPS, and that the response in PVT neurons are heterogeneous. CUBIC-Cloud facilitates to share such knowledge through widely accessible data repository, and users will know exactly, with $\leq 100 \mu\text{m}$ spacial resolution in 3D space, where the concerning region is. By iterating such experiments, CUBIC-Cloud opens up the possibility to functionally dissect the mouse brain with much finer detail and specificity.

Furthermore, CUBIC-Cloud was used to comprehensively identify brain-wide neuronal connections. As a proof-of-concept demonstration, input of neuronal populations projecting to ARH^{Kiss1+} neurons were investigated by using Rabies virus tracers (Fig. 5). Our analysis was successful in identifying extremely sparse populations (less than 10 cells/mm³) with high reproducibility. Our result also implied sexually dimorphic projection from LSr to ARH^{Kiss1+}. Again, we emphasize that CUBIC-Cloud can serve as a central hub for researchers to share such anatomical knowledge.

The above analysis results are all openly accessible at CUBIC-Cloud. The raw image data, along with the markers highlighting the detected cells, are available through CATMAID image browser [53], hosted at <http://cubic-atlas.riken.jp>.

CUBIC-Cloud is, to our knowledge, the first cloud-based framework to share and publish whole-mouse-brain datasets. We believe, however, that several technical and administrative advancements need to be made to render it a fully functional data scientific platform for neuroscience. Firstly, sharing and viewing of the raw image data has not been addressed in CUBIC-Cloud. There are several software and repository sites for sharing and viewing large biological image datasets [54, 55, 56]. The challenge would be to define an ontology to effectively tag each data so that easy yet flexible search is facilitated. In addition, the deposited image should preferably be aligned with the reference to allow virtual overlay of multiple brain images. To ensure the persistence of the deposited data, involvement of the public agencies may be critical, as is the case in GenBank or PDB.

Secondly, the current implementation of CUBIC-Cloud asks the user to perform cell detection locally. The cell (object) detection routine, along with the machine learning training, should be integrated in the cloud in the future developments. Indeed, there are already several reports on embracing cloud computing to analyze massive light and electron microscopy images [57, 58, 59, 60]. Interesting bonus of running object detection in the cloud is that the training datasets can also be shared, which will allow to train more general and accurate deep learning models [59]. Furthermore, CUBIC-Cloud is currently aimed at analysing single-cell-resolution image data, where signals of interest are given as dots localized in cell nuclei/soma. Inclusion of object detection routine in the cloud would expand the scope to include fiber and synaptic structures. Technically, such high-resolution image data can be readily obtained by state-of-the-art LSM, with boosted resolution by expansion microscopy [61, 62, 63, 18]. Future software infrastructure developments, including ones discussed here, will pave the new science towards bottom-up and data-driven elucidation of neuronal functions and circuitry.

Acknowledgements

Authors would like to thank K. Wilkins and T. Miyawaki for proofreading the manuscript, T.C. Murakami for discussions on CUBIC-Atlas, M. Kuroda for technical assistance on LSFM imaging, S.I. Kubota for helping LPS experiments, R. Tanaka for supporting sample preparations. The *App^{NL-G-F/NL-G-F}* knock-in mouse strain (RBRC06344) was provided by RIKEN BRC (deposited by T. Saito and T.C. Saido) through the National BioResource Project of the MEXT/AMED, Japan. This work was supported by Grant-in-Aid for JSPS Fellows (JSPS KAKENHI, grant no. 19J13071, to T.M.); ANRI Scholarship (to T.M.); PRESTO (JST, grant no. JPMJPR15F4, to E.A.S.); Grants-in-Aid for Scientific Research (B) (JSPS KAKENHI, grant no. 19H03413 to E.A.S.); Grants-in-Aid for Scientific Research on Innovative Areas (JSPS KAKENHI, grant no. 17H06328 to E.A.S.); ERATO Touhara Chemosensory Signal Project (JST, grant no. JPMJER1202, to K.T.); Grant-in-Aid for Scientific Research (S) (JSPS KAKENHI, grant no. 25221004, to H.R.U.); Brain/MINDS (AMED/MEXT, grant no. JP17DM0207049, to H.R.U.); the Science and Technology Platform Program for Advanced Biological Medicine (AMED/MEXT to H.R.U.); Grants-in-Aid from the Human Frontier Science Program (to H.R.U.); Grants-in-Aid from the Takeda Science Foundation (to E.A.S. and H.R.U.). The graphical user interfaces of CUBIC-Cloud was co-developed by CUBICStars and Tecotec Inc.

Author contributions

T.M. and H.R.U. designed the study. T.M. created and implemented CUBIC-Cloud software stack. T.M. and K.K. performed data analysis. T.M. and E.A.S. performed LSFM imaging. C.S., K.K. and H.O. performed tissue clearing and immunostaining experiments. K.K. performed sleep measurements using SSS. H.O. assisted in the preparation of AD model mice. S.S. and R.G.Y. assisted in the deployment of the cloud server. E.A.S., K.T., K. Murata and K. Miyamichi prepared Cre-mouse, produced virus and performed injection. T.M., S.S. and H.R.U. wrote the manuscript with the inputs from all co-authors.

Declaration of Interests

T.M. and H.R.U. filed a patent application regarding the CUBIC-Cloud software. CUBIC-Cloud web service is provided and maintained by CUBICStars Inc.

Acronym	Full name
AAA	Anterior amygdalar area
ACB	Nucleus accumbens
ADP	Anterodorsal preoptic nucleus
AHN	Anterior hypothalamic nucleus
ARH	Arcuate hypothalamic nucleus
AVP	Anteroventral preoptic nucleus
BST	Bed nuclei of the stria terminalis
CEA	Central amygdalar nucleus
CLI	Central linear nucleus raphe
DMH	Dorsomedial nucleus of the hypothalamus
DMX	Dorsal motor nucleus of the vagus nerve
DTN	Dorsal tegmental nucleus
FL	Flocculus
HATA	Hippocampo-amygdalar transition area
IA	Intercalated amygdalar nucleus
IC	Inferior colliculus
ILA	Infralimbic area
IO	Inferior olivary complex
LC	Locus ceruleus
LDT	Laterodorsal tegmental nucleus
LH	Lateral habenula
LRN	Lateral reticular nucleus
LS	Lateral septal nucleus
LPO	Lateral preoptic area
MEA	Medial amygdalar nucleus
MH	Medial habenula
MRN	Midbrain reticular nucleus
MPN	Medial preoptic nucleus
MPO	Medial preoptic area
NLL	Nucleus of the lateral lemniscus
NOD	Nodulus (X)
NTB	Nucleus of the trapezoid body
NTS	Nucleus of the solitary tract
PAG	Periaqueductal gray
PB	Parabrachial nucleus
PH	Posterior hypothalamic nucleus
PP	Peripeduncular nucleus
PVH	Paraventricular hypothalamic nucleus
PVa	Periventricular hypothalamic nucleus, anterior part
PVi	Periventricular hypothalamic nucleus, intermediate part
PVp	Periventricular hypothalamic nucleus, posterior part
PVpo	Periventricular hypothalamic nucleus, preoptic part
PVT	Paraventricular nucleus of the thalamus
RAmb	Midbrain raphe nuclei
RL	Rostral linear nucleus raphe
RR	Midbrain reticular nucleus, retrorubral area
RT	Reticular nucleus of the thalamus
SFO	Subfornical organ
SNc	Substantia nigra, compact part
SNr	Substantia nigra, reticular part
SO	Supraoptic nucleus
SOC	Superior olivary complex
SUB	Subiculum
TRS	Triangular nucleus of septum
VMH	Ventromedial hypothalamic nucleus
VTA	Ventral tegmental area
ZI	Zona incerta

Table 1: Brain region acronyms used in this paper

Methods

Mice

All experimental procedures and housing conditions were approved by the Animal Care and Use Committee of The University of Tokyo. *App^{NL-G-F/NL-G-F}* mice were provided by RIKEN BioResource Research Center (RBRC No. RBRC06344) [42]. Kiss1-Cre mouse was purchased from The Jackson Laboratory (Kiss1-tm1.1(cre/EGFP)Steil/J, stock no. 017701). In all experiments, male mice were used unless otherwise specified.

To sample brains, animals were anesthetized by an overdose of pentobarbital (> 100 mg/kg), then transcardially perfused with 10 mL of PBS (pH 7.4) and 20 mL of 4% paraformaldehyde (PFA). The brains were dissected, post-fixed in 4% PFA overnight at 4 °C and stored in PBS.

Tissue clearing and whole-brain 3D immunostaining

To clear brain tissues, we used second-generation CUBIC protocols [30, 64]. For delipidation, we used CUBIC-L (10 wt% *N*-butyldiethanolamine, 10 wt% Triton X-100), and for RI matching, either CUBIC-R+(N) (45 wt% Antipyrine, 30 wt% Nicotinamide, 0.5 vol% *N*-butyldiethanolamine) or CUBIC-R+(M) (45 wt% Antipyrine, 30 wt% *N*-methylnicotinamide, 0.5 vol% *N*-butyldiethanolamine) was used. CUBIC-R+(M) was used for samples labelled by fluorescent proteins (FPs) because of its better FP signal retention, and CUBIC-R+(N) was used otherwise. For all brains, nuclear staining (using either SYTOX-G, BOBO-1 or RedDo2) was applied. For whole-mount 3D immunostaining, we followed CUBIC-HV protocol [31]. The details on antibody staining for each experiment can be found in the corresponding sections in Methods. Before imaging, the sample was embedded in a transparent agarose gel so that it could be rigidly mounted on a microscope stage [18, 64].

LSFM imaging

A custom-built macro-zoom LSFM (named GEMINI system) was used to image cleared brains (the details of this microscope can also be found in [31]). For illumination, the microscope was equipped with 488, 532, 594 and 642 nm diode or DPSS lasers (SOLE-6, Omicron). The laser sheet was generated by a cylindrical lens and the sheet thickness was adjustable by a mechanical slit. For detection, the microscope was equipped with 0.63X macro-zoom objective lens (MVPLAPO 0.63X, Olympus) and 0.63-6.3X variable zoom optics (MVX-ZB10, Olympus). After passing a suitable fluorescence filter, the fluorescence signal was captured by sCMOS camera (Zyla 5.5, Andor).

To achieve homogeneous light-sheet thickness throughout the field of view, several rectangular image strips with shifted illumination focus were obtained and digitally stitched together, similar to the approach used in TLS-SPIM [65, 66]. The width of the rectangular strip was matched with the Rayleigh range of the illumination light-sheet. In our setup, the sheet thickness was approximately 10 µm and the rectangular strip width was 1500 µm, which required 6 image strips to cover the entire brain.

For RV samples, the image voxel size was (X,Y,Z) = (8.25, 8.25, 9.0) µm with 1.2X intermediate zoom optics. For other experiments, the voxel size was (X,Y,Z) = (6.45, 6.45, 7.0) µm with 1.6X intermediate zoom optics. It should be noted that the effective resolution should take into account the moderate tissue expansion (~1.5X) caused by CUBIC-R treatment.

For each dye/FP, the following laser and fluorescence filter pair was used: Alexa 594 [Ex: 594 nm, Em: 641/75 nm bandpass (FF02-641/75-32, Semrock)], Cy3 [Ex: 532 nm, Em: 585/40 nm bandpass (FF01-585/40-32, Semrock)], SYTOX-G, BOBO-1 and GFP [Ex: 488 nm, Em: 520/40 nm bandpass (FF01-520/44-32, Semrock)], RedDot2 [Ex: 642nm, Em: 708/75 nm bandpass (FF01-708/75-32, Semrock)], mCherry [Ex: 594 nm, Em: 628/32 nm bandpass (FF01-628/32-32, Semrock)].

Imaging the fluorescent bead embedded in cleared tissue

1.0 µm-diameter green-yellow fluorescent beads (Thermo Fisher, #F8765) were diluted in PBS so that the final bead concentration was 0.9×10^7 particles/ml. This bead-mixed PBS solution was perfused in mice, prior to the PFA perfusion. Because the bead surface was modified with amine, PFA was able to cross-link and fix the beads within the tissue. After tissue clearing, the whole brain was imaged using the macro-zoom LSFM with XYZ voxel resolution of 6.45x6.45x7.0 µm (Supplementary Fig. 2a). Then, single and well-isolated bead particles were manually annotated ($n > 15$ for each brain region) using ITK-SNAP software [67]. Subsequently, the mean spot profiles were computed and fitted with Gaussian using custom Python code. The fitted sigma values from six regions were all within 4.4 to 4.9 µm (lateral) and 5.3 to 6.7 µm (axial) (Supplementary Fig. 2b), validating homogeneous image quality throughout the entire brain. Given the digital sampling frequency (6.5 µm) of the microscope used, this result was nearly the ideal PSF.

Cell detection

The cell detection workflow used in this study is shown in Fig. 1a ("step 3"). As the first step, a machine-learning algorithm was used to classify voxels into labelled cells and other structures. For this step, we used ilastik [32]. To train the classifier, for each label type (such as c-Fos), manual annotation images were prepared. In this study, three classes were defined, which were (1) signals of interest, i.e., cells labelled by FPs or antibodies (2) bright but false signals, such as non-specific binding of antibodies to vascular structures or neurites extending from cell bodies and (3)

background (i.e. void space). Typically, 5,000 to 10,000 voxels were annotated as class 1 per one dataset. To increase the robustness, at least two brains with identical labelling conditions were annotated. Image annotation was performed using ITK-SNAP software [67].

Next, following the ilastik workflow, image feature descriptors were selected to distinguish the three classes defined above. For PV, Sst, ChAT, Th, Iba1, c-Fos, RV-GFP and AAV-mCherry images, the selected descriptors were Gaussian ($\sigma = 0.3, 0.7$ voxel), Gaussian gradient magnitude ($\sigma = 0.7$ voxel), difference of Gaussian ($\sigma = 0.7, 1.0, 1.6, 3.5$ voxel) and Hessian of Gaussian eigenvalues ($\sigma = 1.0, 1.6, 3.5$ voxel). For A β images, difference of Gaussian ($\sigma = 5.0$ voxel) and Hessian of Gaussian eigenvalues ($\sigma = 5.0$ voxel) were additionally included, so that the larger spatial context was taken into account. Then, ilastik software trained the voxel classifier using random forest algorithm. Hyperparameters related to random forest algorithm was automatically configured by ilastik. After the classifier was trained, it was applied to the test dataset to evaluate accuracy. If obvious error was present, manual annotation data was further supplemented, so that the classifier became more robust against such errors. Once the classifier achieves satisfactory accuracy, it could reliably be applied to other images with the same label type.

By applying the voxel classifier trained above to each brain image, a probability image was produced, where the value of each voxel represents the probability of that voxel being class 1 (Supplementary Fig. 3a). The probability value was given in the range [0, 1]. Using this probability image, a custom Python program isolated individual cells in the following way. First, the probability threshold, $P_{th} = 0.7$, was applied to make a binarized image. Then, connected voxels were searched and merged together, to find individual objects. If the identified object volume was larger than a threshold, V_{th} , it was sent to the object separation routine. The object separation routine simply finds local maxima with an exclusion distance $r_{excl} = V_{th}^{1/3}$. V_{th} was heuristically determined to be $V_{th} = 4^3 = 64$ for PV, Sst, c-Fos, Iba1, Rabeis-GFP and AAV-mCherry, and $V_{th} = 5^3 = 125$ for ChAT and Th. For A β plaque segmentation, $V_{th} = \infty$ was used. Hence, there are two free parameters (P_{th} and V_{th}), both of which can be intuitively determined.

Using 20 CPU cores, cell detection of one 3D image stack (typically 2560x2160x1200 voxels) took about 1 hour to complete.

The Python source code and its documentation is available at <https://github.com/DSPsleeporg/ecc>.

Accuracy evaluation of cell detection

Accuracy of the above explained cell detection procedure was extensively evaluated by comparing automated counting results with manual cell counting (Supplementary Fig. 3b-h). Manual cell counting was performed by cropping a small cubic image volume (50 or 75 or 100 voxels, depending on the cell density) from brain images which were not used in machine learning training. Well-trained human annotators ($n = 2$) independently marked all of the cells present in the image and typically yielded 100-200 marked cells. Image annotation was performed using ITK-SNAP software [67]. Cells annotated by both human and algorithm were regarded as true positives. Cells annotated by human but not by algorithm was regarded as false negative. Cells annotated by algorithm but not by human were regarded as false positives. Then, true positive rate (TPR, also called sensitivity) and positive predictive value (PPV, also called precision) were evaluated for each human annotator. To quantify the overall performance, F1 score, defined as $F1 = 2 * (PPV * TPR) / (PPV + TPR)$, was also evaluated. For most of the label types and brain regions, our cell detection algorithm robustly demonstrated good F1 scores, with average score being 0.80 (PV), 0.83 (Sst), 0.88 (ChAT), 0.80 (Th), 0.88 (Iba1), 0.83 (c-Fos) and 0.89 (GFP).

Raw 3D image data, along with markers highlighting all detected cells, are available at <http://cubic-atlas.riken.jp> through an interactive image viewer powered by CATMAID [53].

CUBIC-Atlas

CUBIC-Cloud uses CUBIC-Atlas version 1.1 [18] as the reference brain coordinate, to which all individual brain data were mapped. From the originally published atlas (version 1.0), we added a few minor updates in this study to generate CUBIC-Atlas version 1.1. The updates included (1) slight discontinuity between dorsal and ventral image volume (so-called "theta tile displacement" in Murakami et al. paper) were corrected and (2) region annotation was updated to be compatible with the Allen Brain Atlas CCFv3 (October 2017). CUBIC-Atlas v1.1 can be downloaded from <http://cubic-atlas.riken.jp>.

Brain registration

CUBIC-Cloud uses the symmetric image normalization (SyN) algorithm implemented in ANTs library [34] to run registration between CUBIC-Atlas ("fixed" brain) and individual brain sample ("moving" brain). First, nuclear staining image of the "moving" image was downsampled to a voxel size of 50x50x50 μm . Nuclear staining image of CUBIC-Atlas was downsampled to a voxel size of 80x80x80 μm . Considering the sample's physical expansion by clearing treatment (2.2X for CUBIC-Atlas and 1.5X for CUBIC-R+ treated brains), this downscaling operation resulted in an effective voxel size of about 35x35x35 μm in both images. The registration first computed affine transformation to coarsely align the orientation and size, using mutual information as the optimizer metric. Subsequently, non-linear warping was computed by SyN algorithm, which optimized the warp field by maximizing the normalized cross-correlation (NCC) between the two images under diffeomorphic regularization [34]. Given image $I(\mathbf{x})$ and image $J(\mathbf{x})$, the NCC value

1 between I and J at the voxel position $\mathbf{x}$ is given by

$$NCC(I, J, \mathbf{x}) = \frac{\langle \bar{I}, \bar{J} \rangle}{|\bar{I}| |\bar{J}|} \quad (1)$$

2 where $\langle A, B \rangle$ represents the inner product taken over a local window with radius R centered at position $\mathbf{x}$. $|A|$ is the
3 L2 norm of the vector computed over a local window with radius R . Here, $\bar{I}(\mathbf{x}) = I(\mathbf{x}) - \mu_I(\mathbf{x})$ means the subtraction of
4 the local mean, where local mean $\mu_I(\mathbf{x})$ is computed over a local window with radius R centered at position $\mathbf{x}$. $R = 4$
5 (voxels) was used throughout our brain registration.

6 The representative registration result is visualized in [Supplementary Fig. 3i](#), along with the NCC value heatmap
7 ([Supplementary Fig. 3l](#)). To show the reproducibility of the registration, 20 individual brains were mapped onto
8 CUBIC-Atlas with identical registration parameters. The mean NCC value of each coronal planes were computed and
9 plotted ([Supplementary Fig. 3h](#)). 20 independent curves overlap with each other, meaning that the optimization attempt
10 by the registration reached saturation. NCC value tends to show higher value in the olfactory area and cerebellum,
11 due to the presence of more distinct structural features.

12 **Architecture of CUBIC-Cloud**

13 CUBIC-Cloud's entire application stack is deployed on the cloud computing infrastructures offered by Amazon Web
14 Service (AWS). The cloud is constructed using the serverless architecture [68]. Serverless architecture have zero real
15 instances that are always running; instead, the cloud is composed by connecting microservices, which are dynamically
16 invoked by events. Such cloud design eliminates the cost of idling servers, while allowing to flexibly and instantly scale
17 out the computing power when the traffic or load to the service increases.

18 The schematic illustration of the cloud architecture is shown in [Supplementary Fig. 1](#). When user accesses the
19 web site, the traffic is first handled by CloudFront. CloudFront is responsible for caching static contents, managing
20 SSL/TLS and web application firewall (WAF). Then, static web contents are fetched from S3 bucket and returned to the
21 user. User authentication is handled by Cognito. Once authenticated, users can access the protected API endpoints
22 securely using json web token (JWT). All REST API requests are routed by API Gateway and handled by Lambda.
23 Lambda handlers have access to various back-end resources, including the databases and the data buckets. The
24 metadata of the users, brains, notebooks, and studios are stored in DynamoDB. Large data files (such as images and
25 csv tables) are stored in S3.

26 Once a user uploads the brain data, the upload completion event from S3 triggers a "preprocessing" task in the ECS
27 cluster. Preprocessing includes brain registration, transformation, and data conversion. ECS automatically launches
28 a new EC2 instance, pulls the Docker container from ECR registry, and initiate a new task. The task execution is
29 orchestrated by StepFunctions. Notebook tasks (i.e. generating plots) are similarly orchestrated by StepFunctions,
30 except that the runtime is either Lambda or Fargate, depending on the required memory size of the task.

31 Most of the API handlers are written in Python. The cloud resources are managed by AWS CDK library for Python,
32 and the cloud is deployed using the CloudFormation generated by the CDK library. The user interfaces (UIs) were
33 constructed using HTML/CSS/JavaScript and Vue.js framework.

34 **A web-based whole-brain viewer**

35 CUBIC-Cloud offers a point-cloud based interactive 3D brain viewer, a feature called studio. The viewer is written in
36 JavaScript, and runs in most of the standard web browsers, including Google Chrome and Firefox. It uses WebGL for
37 hardware accelerated 3D rendering. The core of the point cloud rendering engine was adopted from the open-source
38 project, Potree [69]. Following Potree, CloudEye uses a specialized point cloud format where the whole point cloud
39 was divided and stored in a multi-resolution hierarchical structure (octree structure). This hierarchical data structure
40 enabled the adaptive data querying in response to client's viewpoint, in which a portion of the points near the viewer's
41 camera was selectively loaded. The octree-formatted point cloud data was automatically generated as a part of the
42 preprocessing task in the CUBIC-Cloud server. Each point can be attached with several attributes, including the region
43 ID and fluorescent intensity. Points may be colored using these attributes. For example, points may be given gradient
44 colors based on their fluorescence intensity values.

45 Users can navigate and explore the whole-brain data by intuitive mouse interactions. Points can be selectively
46 hidden/displayed based on the region ID. Arbitrary combinations of brains may be overlaid with user-defined colors.
47 Users can also make a slice view, which can be moved or rotated with simple mouse dragging. Users can also grab a
48 point (cell) and query its information, such as fluorescent intensity.

49 **Whole-brain analysis of PV and SST expressing neurons**

50 C57BL/6N wild-type mice brains (8-week-old, $n = 4$) were cleared, stained, imaged and analyzed as described
51 in the corresponding sections in Methods. Brains were stained with PV antibody (Swant, #PV235; 1/50 dilution;
52 anti-mouse IgG1 secondary Fab fragment conjugated with Cy3 (Jackson ImmunoResearch, #115-167-185)), SST
53 antibody (Millipore, #MAB354; 1/10 dilution; anti-rat IgG secondary Fab fragment conjugated with Alexa 594 (Jackson
54 ImmunoResearch, #112-587-008)) and nuclear staining dye (BOBO-1, Thermo Fisher #B3582). The whole-brain
55 summary of detected PV+ and SST+ cells are provided in [Supplementary Table 2](#).

We discussed the PV+ and SST+ cell in the isocortex in the main text. Within the striatum, PV+ cells were almost entirely absent in the LS and anterior, central, intercalated and medial amygdalar nucleus (AAA, CEA, IA and MEA), as observed previously [70]. Distribution in the brain stem faithfully reproduced the previous slice-based immunohistochemical studies [71, 72] (Fig. 2i). In general, the thalamus contained low numbers of PV+ or SST+ cells, except that dense PV+ cells were present in the RT and PP. In the hypothalamus, although PV+ cells were sparse, many nuclei contained medium to high density of SST+ neurons. Within the midbrain, PV+ cells were particularly abundant in the IC and SNr, while SST+ cells were most frequently observed in the RAmb. Within the pons and medulla, the NTB, SOC and NLL contained both PV+ and SST+ cells with relatively high density, while sparsely scattered populations were observed in other areas. In the cerebellum, there were a large number of PV+ neurons in Purkinje layers. Distinct SST+ cell clusters were found in the NOD and FL.

Whole-brain analysis of ChAT expressing neurons

C57BL/6N wild-type mice brains (8-week-old, $n = 4$) were cleared, stained, imaged and analyzed as described in the corresponding sections in Methods. Brains were stained with ChAT antibody (abcam, #ab178850; 1/200 dilution; anti-rabbit IgG secondary Fab fragment conjugated with Alexa 594 (Jackson ImmunoResearch, #111-587-008)) and nuclear staining (SYTOX-G, Thermo Fisher, #S7020). The whole-brain summary of detected ChAT+ cells are provided in [Supplementary Table 2](#).

About half of ChAT+ cells were concentrated in striatum and pallidum (34.5% and 13.8%, respectively). Continuously spreading from these regions, some ChAT+ cells were present in the hypothalamus, including lateral, medial and anteroventral preoptic areas (LPO, MPO, AVP) and SO. ChAT+ neurons were hardly observed in the olfactory area, hippocampus, cortical subplate and thalamus. These observations were in good agreement with the slice-based immunohistochemical study by [73].

Whole-brain analysis of Th expressing neurons

C57BL/6N wild-type mice brains (8-week-old, $n = 4$) were cleared, stained, imaged and analyzed as described in the corresponding sections in Methods. Brains were stained with Th antibody (Santa Cruz Biotechnology, #sc-25269; 1/20 dilution; anti-mouse IgG2a secondary Fab fragment conjugated with Alexa 594 (Jackson ImmunoResearch, #115-587-186)) and nuclear staining (SYTOX-G). The whole-brain summary of detected Th+ cells are provided in [Supplementary Table 2](#).

The majority of the detected Th neurons were localized in well-known dopaminergic cell groups (A8 to A16) and noradrenergic cell groups (A1 to A7) [74]. Dopaminergic cell groups include the RR, SNc, rostral and central linear nucleus raphe (RL and CLI) and VTA, which form the A8, A9 and A10 in midbrain. Within the hypothalamus, Th neurons were clustered in the periventricular hypothalamic nucleus, anterior, posterior, intermediate and propotic parts (PVA, PVP, PVi, PVpo), ARH, ZI and ADP, which form A11-A15 cell groups. Th neurons were numerous in the olfactory area (A16), selectively localized in the glomerular layer. Noradrenergic cell groups formed distinct bands crossing several nuclei in the medulla and pons, which included the LRN, NTS and DMX, which form A1 and A2. In the pons, a particularly high density was observed in and around LC, which forms A6. No significant population of Th+ cells were observed in the isocortex, hippocampus, cortical subplate, striatum and pallidum.

Brain nuclei segmentation using alpha shape algorithm

A densely aggregated ChAT expressing cells were found in the LDT and its neighboring regions. To segment this brain region automatically, a rectangular region containing these cells were manually cropped. Then, to remove the isolated cells that was not part of the continuous body of the nuclei, the following filter was applied: for each cell in the ensemble, the number of neighboring cells within the radius 100 μm were counted, and if the count is less than 2, the cell was removed from the ensemble. After this filtering, the polygonal surface enclosing the points was constructed using alpha shape algorithm with alpha radius parameter 300 μm , implemented in MATLAB. This method was applied to $n = 4$ brains aligned with CUBIC-Atlas, obtaining four independent boundary surfaces. Using the same method, cluster of TH expressing cells around LC were segmented ($n = 4$).

Lastly, the overlap between ChAT-defined boundaries were evaluated by computing the Dice's coefficient, $c = (2 * V_{V1 \cap V2}) / (V_1 + V_2)$ where $V_{V1 \cap V2}$ is the volume of the overlap between the two polygons whose volumes are V_1 and V_2 , respectively. The Dice's coefficient of all possible pairs, $\binom{4}{2} = 6$, were evaluated. The same analysis was applied to TH-defined boundaries. Lastly, the overlaps between ChAT and TH boundaries were evaluated by picking all possible pairs, $4^2 = 16$.

Whole-brain analysis of Iba1 expressing cells

C57BL/6N wild-type mice (8-week-old, $n = 7$ for LPS-administered and control group, respectively) were administered with either 1 mg/kg of LPS or saline via i.p. injection. 24 hours after injection, brains were sampled. Subsequently, brains were cleared, stained, imaged and analyzed as described in the corresponding sections in Methods. Brains were stained with Iba1 antibody (Wako, #013-26471; 1/50 dilution; directly conjugated with red dye) and nuclear staining (SYTOX-G). The whole-brain summary of detected Iba1+ cells are provided in [Supplementary Table 3](#).

Whole-brain analysis of A β plaques

App^{NL-G-F/NL-G-F} mice (9- to 10-month-old, $n = 4$) were cleared, stained, imaged and analyzed as described in the corresponding sections in Methods. Brains were stained with β -Amyloid (6E10) antibody (Biolegend, #93049; 1/100 dilution; anti-mouse IgG1 secondary Fab fragment conjugated with Alexa 594 (Jackson ImmunoResearch, #115-587-185)) and nuclear staining (SYTOX-G).

Whole-brain analysis of c-Fos profile change induced by LPS

To monitor the sleep duration of mice in a non-invasive manner, we used a respiration-based sleep staging method, SSS [75]. C57BL/6N wild-type mice (8-week-old, $n = 4$ each for LPS-administered and control group) were housed in a SSS chamber for 3 days (basal measurement; LD cycle) prior to the injection. On the fourth day at ZT = 14, mice were administered 150 μ g/kg LPS from *Escherichia coli* (Sigma-Aldrich, #L2990) via i.p. injection, while control mice were administered saline (Supplementary Fig. 5d). After injection, mice were housed in the SSS chamber for another 24 hours to monitor sleep.

Brains for whole-brain imaging were collected in a replicate experiment with the identical conditions as above, except that mice were housed in DD cycle and that brains were sampled at CT = 16-17 after i.p. injection ($n = 4$ in total for each group, obtained in 1 batch). Collected samples were cleared, stained, imaged and analyzed as described in corresponding sections in Methods. Brains were stained by c-Fos antibody (CST, #2250S; 1/75 dilution; anti-rabbit IgG secondary Fab fragment conjugated with Alexa 594 (Jackson ImmunoResearch, #111-587-008)) and nuclear staining dye (SYTOX-G).

Rabies virus injection

The following AAV vectors were generated de novo by PENN vector core using the corresponding plasmids. AAV serotype 9 CAG-FLEX-TCb (1.5×10^{13} gp/ml) was made using the plasmid described previously [48]. Here TCb stands for TVA-mCherry expression cassette optimized to increase mCherry brightness. To generate AAV serotype 9 CAG-FLEX-oG (4.5×10^{13} gp/ml), engineered and optimized glycoprotein (oG) [76] sequence was ligated to pAAV-FLEX sequence from pAAV-FLEX-GFP (Addgene).

Preparation of rabies virus was conducted by using the RV Δ G-GFP, B7GG and BHK-EnvA cells as previously described [77]. The EnvA-pseudotyped RV Δ G-GFP+EnvA titer was estimated to be 1.0×10^9 infectious particles/ml based on serial dilutions of the virus stock followed by infection of the HEK293-TVA800 cell line (a gift from Dr. Edward Callaway at Salk Institute).

For trans-synaptic tracing using rabies virus, about 20 nL of mixture of AAV9 CAG-FLEX-TCb and CAG-FLEX-oG (diluted to 1.5×10^{12} gp/ml each) was injected into the ARH of Kiss1-Cre mice. The first AAV transduced a TVA receptor (fused with mCherry) for EnvA. The second AAV transduced RV glycoprotein (oG) playing a predominant role in the trans-synaptic transport of RV. The injection coordinate was P1.1, L0.2, V5.9 (distance in mm from the Bregma for the posterior [P], and lateral left [L] positions and from the brain surface for the ventral [V] position). Three weeks later, 30 nL of Rabies Δ G-GFP+EnvA was injected into the same brain region to initiate trans-synaptic tracing. Because there is no cognate receptor for EnvA in the mouse brain, RV Δ G+EnvA only infects TVA-expressing cells. oG expression from the second AAV complements the RV Δ G, allowing retrograde monosynaptic tracing from Cre-expressing cells. Seven days later, brains were sampled for CUBIC treatment.

Whole-brain analysis of RV-injected brains

After AAV and RV injection, Kiss1-Cre mice (13- or 14-week-old at the time of brain sampling) were cleared, stained, imaged and analyzed as described in the corresponding sections in Methods. Brains were stained with nuclear staining (RedDot2, Biotium, #40061).

As a negative control experiment, AAV and RV injection was performed using BALB/c wild-type mice brain ($n = 3$). After clearing, the whole-brain image was obtained by LSM. No GFP or mCherry signals were observed by manual inspection, confirming the absence of Cre-independent leakage of AAV vectors and specificity of the virus delivery.

Starter cells were searched by identifying dual positive (mCherry+ and GFP+) cells. For each channel, cell counting was independently performed, and the center of the mass of the detected cell was obtained. For each mCherry+ cells, if a GFP+ cell was present within a distance of 24 μ m, the cell was counted as starter. Note that because the cleared tissue was expanded by a factor of ~ 1.5 , this was roughly 16 μ m in untreated tissue. Occasionally, slight voxel shift (typically no more than 4 voxels) occurred between GFP and mCherry channels, which was presumably caused by slight misalignment between the 488 nm and 594 nm illumination laser or re-focusing of the microscope. To correct this, small 3D volumes with distinct features (typically (X,Y,Z) = (50,50,20) voxel volume, $n = 4$ or $n = 3$) from mCherry and GFP channels were cropped, and the voxel shift was computed by registering two images using ANTs, where transformation was restricted to only translation. Then, the cell coordinates were corrected by the mean of the computed shift.

To carry out statistical analysis of input cell numbers between male and female brains, we used the normalized cell count, $n_{\text{norm},i}$, where i represents the ID of the brain region. If we let the raw cell number of each brain region be $n_{\text{raw},i}$, $n_{\text{norm},i}$ is simply expressed as $n_{\text{norm},i} = n_{\text{raw},i} / (\sum_i n_{\text{raw},i})$.

Data availability

Raw 3D image data analyzed in this study are available at <http://cubic-atlas.riken.jp> through an interactive image viewer powered by CATMAID [53]. The analyzed whole-brain data in point cloud format is deposited on <http://cubic-atlas.riken.jp>, as well as CUBIC-Cloud's public repository. The additional data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

Cell detection program used in this study is available at GitHub code repository (<https://github.com/DSPsleeporg/ecc>). CUBIC-Cloud computing service is accessible at <https://cubic-cloud.com>.

CUBIC-Cloud is free for use for viewing data. Requests to upload brains or create notebooks are handled for a fee, to cover and compensate the maintenance cost of the cloud server.

References

- [1] Hans-Ulrich Dodt, Ulrich Leischner, Anja Schierloh, Nina Jährling, Christoph Peter Mauch, Katrin Deininger, Jan Michael Deussing, Matthias Eder, Walter Zieglgänsberger, and Klaus Becker. Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nature Methods*, 4(4):331–336, apr 2007.
- [2] Hiroshi Hama, Hiroshi Kurokawa, Hiroyuki Kawano, Ryoko Ando, Tomomi Shimogori, Hisayori Noda, Kiyoko Fukami, Asako Sakaue-Sawano, and Atsushi Miyawaki. Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. *Nature Neuroscience*, 14(11):1481–1488, nov 2011.
- [3] Ali Ertürk, Klaus Becker, Nina Jährling, Christoph P Mauch, Caroline D Hojer, Jackson G Egen, Farida Hellal, Frank Bradke, Morgan Sheng, and Hans-Ulrich Dodt. Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nature Protocols*, 7(11):1983–1995, nov 2012.
- [4] Meng-Tsen Ke, Satoshi Fujimoto, and Takeshi Imai. SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nature Neuroscience*, 16(8):1154–1161, aug 2013.
- [5] Kwanghun Chung, Jenelle Wallace, Sung-Yon Kim, Sandhiya Kalyanasundaram, Aaron S. Andalman, Thomas J. Davidson, Julie J. Mirzabekov, Kelly A. Zalocusky, Joanna Mattis, Aleksandra K. Denisin, Sally Pak, Hannah Bernstein, Charu Ramakrishnan, Logan Grose, Viviana Gradinaru, and Karl Deisseroth. Structural and molecular interrogation of intact biological systems. *Nature*, 497(7449):332–337, may 2013.
- [6] Bin Yang, Jennifer B. Treweek, Rajan P. Kulkarni, Benjamin E. Deverman, Chun Kan Chen, Eric Lubeck, Sheel Shah, Long Cai, and Viviana Gradinaru. Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell*, 158(4):945–958, 2014.
- [7] Etsuo A. Susaki, Kazuki Tainaka, Dimitri Perrin, Fumiaki Kishino, Takehiro Tawara, Tomonobu M. Watanabe, Chihiro Yokoyama, Hirotaka Onoe, Megumi Eguchi, Shun Yamaguchi, Takaya Abe, Hiroshi Kiyonari, Yoshihiro Shimizu, Atsushi Miyawaki, Hideo Yokota, and Hiroki R. Ueda. Whole-Brain Imaging with Single-Cell Resolution Using Chemical Cocktails and Computational Analysis. *Cell*, 157(3):726–739, apr 2014.
- [8] Kazuki Tainaka, Shimpei I. Kubota, Takeru Q. Suyama, Etsuo A. Susaki, Dimitri Perrin, Maki Ukai-Tadenuma, Hideki Ukai, and Hiroki R. Ueda. Whole-body imaging with single-cell resolution by tissue decolorization. *Cell*, 159(4):911–24, nov 2014.
- [9] Morgane Belle, David Godefroy, Gérard Couly, Samuel A Malone, Francis Collier, Paolo Giacobini, and Alain Chédotal. Tridimensional Visualization and Analysis of Early Human Development. *Cell*, 169(1):161–173.e12, aug 2017.
- [10] Rory M Power and Jan Huiskens. A guide to light-sheet fluorescence microscopy for multiscale imaging. *Nature Methods*, 14(4):360–373, 2017.
- [11] Nicolas Renier, Zhuohao Wu, David J. Simon, Jing Yang, Pablo Ariel, and Marc Tessier-Lavigne. iDISCO: A Simple, Rapid Method to Immunolabel Large Tissue Samples for Volume Imaging. *Cell*, 159(4):896–910, nov 2014.
- [12] Evan Murray, Jae Hun Cho, Daniel Goodwin, Taeyun Ku, Justin Swaney, Sung-Yon Kim, Heejin Choi, Young-Gyun Park, Jeong-Yoon Park, Austin Hubbert, Margaret McCue, Sara Vassallo, Naveed Bakh, Matthew P. Frosch, Van J. Wedeen, H. Sebastian Seung, and Kwanghun Chung. Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell*, 163(6):1500–1514, dec 2015.
- [13] Hiroshi Hama, Hiroyuki Hioki, Kana Namiki, Tetsushi Hoshida, Hiroshi Kurokawa, Fumiyoshi Ishidate, Takeshi Kaneko, Takumi Akagi, Takashi Saito, Takaomi Saido, and Atsushi Miyawaki. ScaleS: an optical clearing palette for biological imaging. *Nature Neuroscience*, 18(10):1518–1529, oct 2015.
- [14] Ken Y. Chan, Min J. Jang, Bryan B. Yoo, Alon Greenbaum, Namita Ravi, Wei Li Wu, Luis Sánchez-Guardado, Carlos Lois, Sarkis K. Mazmanian, Benjamin E. Deverman, and Viviana Gradinaru. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nature Neuroscience*, 20(8):1172–1179, 2017.
- [15] Shimpei I. Kubota, Kei Takahashi, Jun Nishida, Yasuyuki Morishita, Shogo Ehata, Kazuki Tainaka, Kohei Miyazono, and Hiroki R. Ueda. Whole-Body Profiling of Cancer Metastasis with Single-Cell Resolution. *Cell Reports*, 20(1):236–250, jul 2017.

- [16] Ruiyao Cai, Chenchen Pan, Alireza Ghasemigharagoz, Mihail Ivilinov Todorov, Benjamin Förstera, Shan Zhao, Harsharan S. Bhatia, Arnaldo Parra-Damas, Leander Mrowka, Delphine Theodorou, Markus Rempfler, Anna L. R. Xavier, Benjamin T. Kress, Corinne Benakis, Hanno Steinke, Sabine Liebscher, Ingo Bechmann, Arthur Liesz, Bjoern Menze, Martin Kerschensteiner, Maiken Nedergaard, and Ali Ertürk. Panoptic imaging of transparent mice reveals whole-body neuronal projections and skull–meninges connections. *Nature Neuroscience*, 22(2):317–327, feb 2019.
- [17] Christoph Kirst, Sophie Skriabine, Alba Vieites-Prado, Thomas Topilko, Paul Bertin, Gaspard Gerschenfeld, Florine Verny, Piotr Topilko, Nicolas Michalski, Marc Tessier-Lavigne, and Nicolas Renier. Mapping the Fine-Scale Organization and Plasticity of the Brain Vasculature. *Cell*, 180(4):780–795.e25, 2020.
- [18] Tatsuya C Murakami, Tomoyuki Mano, Shu Saikawa, Shuhei A Horiguchi, Daichi Shigeta, Kousuke Baba, Hiroshi Sekiya, Yoshihiro Shimizu, Kenji F Tanaka, Hiroshi Kiyonari, Masamitsu Iino, Hideki Mochizuki, Kazuki Tainaka, and Hiroki R Ueda. A three-dimensional single-cell-resolution whole-brain atlas using CUBIC-X expansion microscopy and tissue clearing. *Nature Neuroscience*, 21(4):625–637, apr 2018.
- [19] Kazuki Tainaka, Akihiro Kuno, Shimpei I. Kubota, Tatzya Murakami, and Hiroki R. Ueda. Chemical Principles in Tissue Clearing and Staining Protocols for Whole-Body Cell Profiling. *Annual Review of Cell and Developmental Biology*, 32(1):713–741, oct 2016.
- [20] Etsuo A. Susaki and Hiroki R. Ueda. Whole-body and Whole-Organ Clearing and Imaging Techniques with Single-Cell Resolution: Toward Organism-Level Systems Biology in Mammals. *Cell Chemical Biology*, 23(1):137–157, jan 2016.
- [21] Tomoyuki Mano, Alexandre Albanese, Hans-Ulrich Dodt, Ali Erturk, Viviana Gradinaru, Jennifer B Treweek, Atsushi Miyawaki, Kwanghun Chung, and Hiroki R Ueda. Whole-Brain Analysis of Cells and Circuits by Tissue Clearing and Light-Sheet Microscopy. *The Journal of Neuroscience*, 38(44):9330–9337, oct 2018.
- [22] Viviana Gradinaru, Jennifer Treweek, Kristin Overton, and Karl Deisseroth. Hydrogel-Tissue Chemistry: Principles and Applications. *Annual review of biophysics*, 47:355–376, 2018.
- [23] Donna Karolchik, Angie S Hinrichs, and W James Kent. The UCSC Genome Browser. *Current protocols in bioinformatics*, Chapter 1:Unit1.4, dec 2009.
- [24] Quanxin Wang, Song-Lin Ding, Yang Li, Josh Royall, David Feng, Phil Lesnar, Nile Graddis, Maitham Naeemi, Benjamin Facer, Anh Ho, Tim Dolbeare, Brandon Blanchard, Nick Dee, Wayne Wakeman, Karla E. Hirokawa, Aaron Szafer, Susan M. Sunkin, Seung Wook Oh, Amy Bernard, John W. Phillips, Michael Hawrylycz, Christof Koch, Hongkui Zeng, Julie A. Harris, and Lydia Ng. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*, 181(4):936–953.e20, may 2020.
- [25] Yongsoo Kim, Guangyu Robert Yang, Kith Pradhan, Kannan Umadevi Venkataraju, Mihail Bota, Luis Carlos García del Molino, Greg Fitzgerald, Keerthi Ram, Miao He, Jesse Maurica Levine, Partha Mitra, Z. Josh Huang, Xiao-Jing Wang, and Pavel Osten. Brain-wide Maps Reveal Stereotyped Cell-Type-Based Cortical Architecture and Subcortical Sexual Dimorphism. *Cell*, 171(2):456–469.e22, oct 2017.
- [26] P. Frasconi, L. Silvestri, P. Soda, R. Cortini, F. S. Pavone, and G. Iannello. Large-scale automated identification of mouse brain cells in confocal light sheet microscopy images. *Bioinformatics*, 30:i587–i593, 2014.
- [27] Etsuo A Susaki, Kazuki Tainaka, Dimitri Perrin, Hiroko Yukinaga, Akihiro Kuno, and Hiroki R Ueda. Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nature Protocols*, 10(11):1709–1727, oct 2015.
- [28] Nicolas Renier, Eliza L. Adams, Christoph Kirst, Zhuhao Wu, Ricardo Azevedo, Johannes Kohl, Anita E. Autry, Lolahon Kadiri, Kannan Umadevi Venkataraju, Yu Zhou, Victoria X. Wang, Cheuk Y. Tang, Olav Olsen, Catherine Dulac, Pavel Osten, and Marc Tessier-Lavigne. Mapping of Brain Activity by Automated Volume Analysis of Immediate Early Genes. *Cell*, 165(7):1789–1802, jun 2016.
- [29] Daniel Fürth, Thomas Vaissière, Ourania Tzortzi, Yang Xuan, Antje Märtin, Iakovos Lazaridis, Giada Spigolon, Gilberto Fisone, Raju Tomer, Karl Deisseroth, Marie Carlén, Courtney A. Miller, Gavin Rumbaugh, and Konstantinos Meletis. An interactive framework for whole-brain maps at cellular resolution. *Nature Neuroscience*, 21(1):139–149, jan 2018.
- [30] Kazuki Tainaka, Tatsuya C. Murakami, Etsuo A. Susaki, Chika Shimizu, Rie Saito, Kei Takahashi, Akiko Hayashi-Takagi, Hiroshi Sekiya, Yasunobu Arima, Satoshi Nojima, Masako Ikemura, Tetsuo Ushiku, Yoshihiro Shimizu, Masaaki Murakami, Kenji F. Tanaka, Masamitsu Iino, Haruo Kasai, Toshikuni Sasaoka, Kazuto Kobayashi, Kohei Miyazono, Eiichi Morii, Tadashi Isa, Masashi Fukayama, Akiyoshi Kakita, and Hiroki R. Ueda. Chemical Landscape for Tissue Clearing Based on Hydrophilic Reagents. *Cell Reports*, 24(8):2196–2210.e9, aug 2018.

- [31] Etsuo A. Susaki, Chika Shimizu, Akihiro Kuno, Kazuki Tainaka, Xiang Li, Kengo Nishi, Ken Morishima, Hiroaki Ono, Koji L. Ode, Yuki Saeki, Kazunari Miyamichi, Kaoru Isa, Chihiro Yokoyama, Hiroki Kitaura, Masako Ikemura, Tetsuo Ushiku, Yoshihiro Shimizu, Takashi Saito, Takaomi C. Saido, Masashi Fukayama, Hirotaka Onoe, Kazushige Touhara, Tadashi Isa, Akiyoshi Kakita, Mitsuhiro Shibayama, and Hiroki R. Ueda. Versatile whole-organ/body staining and imaging based on electrolyte-gel properties of biological tissues. *Nature Communications*, 11(1):1982, 2020.
- [32] Christoph Sommer, Christoph Straehle, Ullrich Kothe, and Fred A. Hamprecht. Ilastik: Interactive learning and segmentation toolkit. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 230–233. IEEE, mar 2011.
- [33] Stuart Berg, Dominik Kutra, Thorben Kroeger, Christoph N. Straehle, Bernhard X. Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, Kemal Eren, Jaime I Cervantes, Buote Xu, Fynn Beuttenmueller, Adrian Wolny, Chong Zhang, Ullrich Koethe, Fred A. Hamprecht, and Anna Kreshuk. ilastik: interactive machine learning for (bio)image analysis. *Nature Methods*, 4522(May):3–5, sep 2019.
- [34] B B Avants, C L Epstein, M Grossman, and J C Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, feb 2008.
- [35] Chen Zhang, Cheng Yan, Miao Ren, Anan Li, Tingwei Quan, Hui Gong, and Jing Yuan. A platform for stereological quantitative analysis of the brain-wide distribution of type-specific neurons. *Scientific Reports*, 7(1):14334, dec 2017.
- [36] Xiangning Li, Bin Yu, Qingtao Sun, Yalun Zhang, Miao Ren, Xiaoyan Zhang, Anan Li, Jing Yuan, Linda Madisen, Qingming Luo, Hongkui Zeng, Hui Gong, and Zilong Qiu. Generation of a whole-brain atlas for the cholinergic system and mesoscopic projectome analysis of basal forebrain cholinergic neurons. *Proceedings of the National Academy of Sciences*, 115(2):415–420, jan 2018.
- [37] Yuri Gonchar, Quanxin Wang, and Andreas Burkhalter. Multiple distinct subtypes of GABAergic neurons in mouse visual cortex identified by triple immunostaining. *Frontiers in neuroanatomy*, 1(March):3, 2007.
- [38] Xiangmin Xu, Keith D. Roby, and Edward M. Callaway. Immunochemical characterization of inhibitory mouse cortical neurons: Three chemically distinct classes of inhibitory cells. *The Journal of Comparative Neurology*, 518(3):389–404, feb 2010.
- [39] Daisuke Ito, Y Imai, K Ohsawa, K Nakajima, Y Fukuuchi, and S Kohsaka. Microglia-specific localisation of a novel calcium binding protein, Iba1. *Brain research. Molecular brain research*, 57(1):1–9, jun 1998.
- [40] Zhihong Chen, Walid Jalabi, Karl B. Shpargel, Kenneth T. Farabaugh, Ranjan Dutta, Xinghua Yin, G. J. Kidd, Cornelia C. Bergmann, Stephen A. Stohlman, and Bruce D. Trapp. Lipopolysaccharide-Induced Microglial Activation and Neuroprotection against Experimental Brain Injury Is Independent of Hematogenous TLR4. *Journal of Neuroscience*, 32(34):11706–11715, aug 2012.
- [41] Eriko Furube, Shintaro Kawai, Haruna Inagaki, Shohei Takagi, and Seiji Miyata. Brain Region-dependent Heterogeneity and Dose-dependent Difference in Transient Microglia Population Increase during Lipopolysaccharide-induced Inflammation. *Scientific Reports*, 8(1):1–15, 2018.
- [42] Takashi Saito, Yukio Matsuba, Naomi Mihira, Jiro Takano, Per Nilsson, Shigeyoshi Itohara, Nobuhisa Iwata, and Takaomi C. Saido. Single App knock-in mouse models of Alzheimer’s disease. *Nature Neuroscience*, 17(5):661–663, may 2014.
- [43] Fumiya Tatsuki, Genshiro A. Sunagawa, Shoi Shi, Etsuo A. Susaki, Hiroko Yukinaga, Dimitri Perrin, Kenta Sumiyama, Maki Ukai-Tadenuma, Hiroshi Fujishima, Rei-ichiro Ohno, Daisuke Tone, Koji L. Ode, Katsuhiko Matsumoto, and Hiroki R. Ueda. Involvement of Ca²⁺-Dependent Hyperpolarization in Sleep Duration in Mammals. *Neuron*, 90(1):70–85, apr 2016.
- [44] Casper Bo Gravesen Salinas, Tess Tsai Hsiu Lu, Sanaz Gabery, Kasper Marstal, Tomas Alanentalo, Aaron Jeffrey Mercer, Anda Cornea, Knut Conradsen, Jacob Hecksher-Sørensen, Anders Bjørholm Dahl, Lotte Bjerre Knudsen, and Anna Secher. Integrated Brain Atlas for Unbiased Mapping of Nervous System Effects Following Liraglutide Treatment. *Scientific Reports*, 8(1):1–12, 2018.
- [45] Yong Wang, JungMin Kim, Matthew B. Schmit, Tiffany S. Cho, Caohui Fang, and Haijiang Cai. A bed nucleus of stria terminalis microcircuit regulating inflammation-associated modulation of feeding. *Nature Communications*, 10(1):2769, dec 2019.

- [46] Claire Gao, Yan Leng, Jun Ma, Victoria Rooke, Shakira Rodriguez-Gonzalez, Charu Ramakrishnan, Karl Deisseroth, and Mario A. Penzo. Two genetically, anatomically and functionally distinct cell types segregate across anteroposterior axis of paraventricular thalamus. *Nature Neuroscience*, 23(2):217–228, feb 2020.
- [47] Allan E. Herbison. The Gonadotropin-Releasing Hormone Pulse Generator. *Endocrinology*, 159(11):3723–3736, nov 2018.
- [48] Kazunari Miyamichi, Yael Shlomei-Fuchs, Marvin Shu, Brandon C. Weissbourd, Liqun Luo, and Adi Mizrahi. Dissecting Local Circuits: Parvalbumin Interneurons Underlie Broad Feedback Control of Olfactory Bulb Output. *Neuron*, 80(5):1232–1245, dec 2013.
- [49] George Paxinos and Keith Franklin. *Paxinos and Franklin's the Mouse Brain in Stereotaxic Coordinates 4th Edition*. Academic Press, 2012.
- [50] Shel-Hwa Yeo, Victoria Kyle, Clemence Blouet, Susan Jones, and William Henry Colledge. Mapping neuronal inputs to Kiss1 neurons in the arcuate nucleus of the mouse. *PLOS ONE*, 14(3):e0213927, mar 2019.
- [51] Shinji Tsukahara, Moeko Kanaya, and Korehito Yamanouchi. Neuroanatomy and sex differences of the lordosis-inhibiting system in the lateral septum. *Frontiers in Neuroscience*, 8(SEP):1–13, sep 2014.
- [52] Shinji Tsukahara and Korehito Yamanouchi. Sex Difference in Septal Neurons Projecting Axons to Midbrain Central Gray in Rats: A Combined Double Retrograde Tracing and ER-Immunohistochemical Study. *Endocrinology*, 143(1):285–294, jan 2002.
- [53] Stephan Saalfeld, Albert Cardona, Volker Hartenstein, and Pavel Tomančák. CATMAID: Collaborative annotation toolkit for massive amounts of image data. *Bioinformatics*, 25(15):1984–1986, 2009.
- [54] Eleanor Williams, Josh Moore, Simon W. Li, Gabriella Rustici, Aleksandra Tarkowska, Anatole Chessel, Simone Leo, Bálint Antal, Richard K. Ferguson, Ugis Sarkans, Alvis Brazma, Rafael E. Carazo Salas, and Jason R. Swedlow. Image Data Resource: A bioimage data integration and publication platform. *Nature Methods*, 14(8):775–781, 2017.
- [55] Dean Kleissas, Robert Hider, Derek Pryor, Timothy Gion, Priya Manavalan, Jordan Matelsky, Alex Baden, Kunal Lillaney, Randal Burns, Denise D'Angelo, William Gray Roncal, and Brock Wester. The Block Object Storage Service (bossDB): A Cloud-Native Approach for Petascale Neuroscience Discovery. *bioRxiv*, pages 1–9, 2017.
- [56] Joshua T Vogelstein, Eric Perlman, Benjamin Falk, Alex Baden, William Gray Roncal, Vikram Chandrashekhar, Forrest Collman, Sharmishta Seshamani, Jesse L. Patsolic, Kunal Lillaney, Michael Kazhdan, Robert Hider, Derek Pryor, Jordan Matelsky, Timothy Gion, Priya Manavalan, Brock Wester, Mark Chevillet, Eric T. Trautman, Khaled Khairy, Eric Bridgeford, Dean M. Kleissas, Daniel J. Tward, Ailey K. Crow, Brian Hsueh, Matthew A. Wright, Michael I. Miller, Stephen J. Smith, R. Jacob Vogelstein, Karl Deisseroth, and Randal Burns. A community-developed open-source computational ecosystem for big neuro data. *Nature Methods*, 15(11):846–847, nov 2018.
- [57] Matthias G. Haberl, Christopher Churas, Lucas Tindall, Daniela Boassa, Sébastien Phan, Eric A. Bushong, Matthew Madany, Raffi Akay, Thomas J. Deerinck, Steven T. Peltier, and Mark H. Ellisman. CDeep3M—Plug-and-Play cloud-based deep learning for image segmentation. *Nature Methods*, 15(9):677–680, sep 2018.
- [58] Thorsten Falk, Dominic Mai, Robert Bensch, Özgün Çiçek, Ahmed Abdulkadir, Yassine Marrakchi, Anton Böhm, Jan Deubner, Zoe Jäckel, Katharina Seiwald, Alexander Dovzhenko, Olaf Tietz, Cristina Dal Bosco, Sean Walsh, Deniz Saltukoglu, Tuan Leng Tay, Marco Prinz, Klaus Palme, Matias Simons, Ilka Diester, Thomas Brox, and Olaf Ronneberger. U-Net: deep learning for cell counting, detection, and morphometry. *Nature Methods*, 16(1):67–70, jan 2019.
- [59] Dylan Bannon, Erick Moen, Morgan Schwartz, Enrico Borba, Sunny Cui, Kevin Huang, Isabella Camplisson, Nora Koe, Daniel Kyme, Takamasa Kudo, Brian Chang, Edward Pao, Erik Osterman, William Graf, and David Van Valen. Dynamic allocation of computational resources for deep learning-enabled cellular image analysis with Kubernetes. *bioRxiv*, page 505032, 2019.
- [60] Jingpeng Wu, William M. Silversmith, Kisuk Lee, and H. Sebastian Seung. Chunkflow: Distributed Hybrid Cloud Processing of Large 3D Images by Convolutional Nets. *Frontiers in Neural Circuits*, 13:1–9, apr 2019.
- [61] F. Chen, P. W. Tillberg, and E. S. Boyden. Expansion microscopy. *Science*, 347(6221):543–548, jan 2015.

- [62] Taeyun Ku, Justin Swaney, Jeong-Yoon Park, Alexandre Albanese, Evan Murray, Jae Hun Cho, Young-Gyun Park, Vamsi Mangena, Jiawei Chen, and Kwanghun Chung. Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nature Biotechnology*, 34(9):973–981, sep 2016.
- [63] Raju Tomer, Li Ye, Brian Hsueh, and Karl Deisseroth. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *Nature Protocols*, 9(7):1682–1697, jun 2014.
- [64] Katsuhiko Matsumoto, Tomoki T. Mitani, Shuhei A. Horiguchi, Junichi Kaneshiro, Tatsuya C. Murakami, Tomoyuki Mano, Hiroshi Fujishima, Ayumu Konno, Tomonobu M. Watanabe, Hirokazu Hirai, and Hiroki R. Ueda. Advanced CUBIC tissue clearing for whole-organ cell profiling. *Nature Protocols*, 14(12):3506–3537, dec 2019.
- [65] Liang Gao. Extend the field of view of selective plan illumination microscopy by tiling the excitation light sheet. *Optics Express*, 23(5):6102, 2015.
- [66] Qinyi Fu, Benjamin L. Martin, David Q. Matus, and Liang Gao. Imaging multicellular specimens with real-time optimized tiling light-sheet selective plane illumination microscopy. *Nature Communications*, 7:11088, 2016.
- [67] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3):1116–1128, jul 2006.
- [68] Gojko Adzic and Robert Chatley. Serverless computing: economic and architectural impact. In *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering - ESEC/FSE 2017*, pages 884–889, New York, New York, USA, 2017. ACM Press.
- [69] Markus Schuetz. Potree: Rendering Large Point Clouds in Web Browsers. *DIPLOMA THESIS, Vienna University of Technology*, 0(0):0, 2016.
- [70] A.J. McDonald and R.L. Betette. Parvalbumin-containing neurons in the rat basolateral amygdala: morphology and co-localization of Calbindin-D28k. *Neuroscience*, 102(2):413–425, jan 2001.
- [71] M R Celio. Calbindin D-28k and parvalbumin in the rat nervous system. *Neuroscience*, 35(2):375–475, jan 1990.
- [72] O. Johansson, T. Hökfelt, and R.P. Elde. Immunohistochemical distribution of somatostatin-like immunoreactivity in the central nervous system of the adult rat. *Neuroscience*, 13(2):265–IN2, oct 1984.
- [73] David M. Armstrong, Clifford B. Saper, Allan I. Levey, Bruce H. Wainer, and Robert D. Terry. Distribution of cholinergic neurons in rat brain: Demonstrated by the immunocytochemical localization of choline acetyltransferase. *Journal of Comparative Neurology*, 216(1):53–68, 1983.
- [74] Annica Dahlström and Kjell Fuxe. EVIDENCE FOR THE EXISTENCE OF MONOAMINE-CONTAINING NEURONS IN THE CENTRAL NERVOUS SYSTEM. I. DEMONSTRATION OF MONOAMINES IN THE CELL BODIES OF BRAIN STEM NEURONS. *Acta physiologica Scandinavica. Supplementum*, 39(6):SUPPL 232:1–55, sep 1964.
- [75] Genshiro A. Sunagawa, Kenta Sumiyama, Maki Ukai-Tadenuma, Dimitri Perrin, Hiroshi Fujishima, Hideki Ukai, Osamu Nishimura, Shoi Shi, Rei-ichiro Ohno, Ryohei Narumi, Yoshihiro Shimizu, Daisuke Tone, Koji L. Ode, Shigehiro Kuraku, and Hiroki R. Ueda. Mammalian Reverse Genetics without Crossing Reveals Nr3a as a Short-Sleeper Gene. *Cell Reports*, 14(3):662–677, jan 2016.
- [76] Euseok J. Kim, Matthew W. Jacobs, Tony Ito-Cole, and Edward M. Callaway. Improved Monosynaptic Neural Circuit Tracing Using Engineered Rabies Virus Glycoproteins. *Cell Reports*, 15(4):692–699, apr 2016.
- [77] Fumitaka Osakada and Edward M. Callaway. Design and generation of recombinant rabies virus vectors. *Nature Protocols*, 8(8):1583–1601, aug 2013.
- [78] Daqing Wang, Xiaobing He, Zhe Zhao, Qiru Feng, Rui Lin, Yue Sun, Ting Ding, Fuqiang Xu, Minmin Luo, and Cheng Zhan. Whole-brain mapping of the direct inputs and axonal projections of POMC and AgRP neurons. *Frontiers in Neuroanatomy*, 9(March):1–17, mar 2015.

Fig. 1 | CUBIC-Cloud: A Cloud-based Computational Framework for Whole Mouse Brain Analysis

- 1 The overview of whole-brain analysis pipeline by CUBIC-Cloud. In this study, mouse brains were cleared by CUBIC-L and CUBIC-R+ reagents
- 2 and 3D-stained by CUBIC-HV protocol. Cleared brains were imaged using macro-zoom LSMF. From the obtained image stacks, single cells were
- 3 isolated using ilastik software, converting the raw raster image into an ensemble of discrete cells (i.e. point cloud). Users then upload the point cloud
- 4 cells and structure image to CUBIC-Cloud. In the cloud, brain registration is automatically performed to align individual brains with the reference
- 5 brain. Thereby, user's own brain database is constructed. Then, user can perform various kinds of brain-wide quantification using "notebook".
- 6 CUBIC-Cloud also offers an interactive 3D whole-brain viewer ("studio"). Lastly, CUBIC-Cloud lets the user share and publish their point-cloud
- 7 whole-brain data, as well as notebooks and studios, to allow broad access from the researchers.

Fig. 2 | Whole-brain Analysis of PV, Sst, ChAT and Th Expressing Cells

- 1 Whole-brain distribution of PV, SST, ChAT, TH and Iba1 expressing cells were investigated by applying 3D immunostaining and by using CUBIC-Cloud
2 analysis framework.
- 3 **a-f**, Whole-brain views of labelled cells. Each point (i.e. single cell) was assigned a pseudo-color based on its fluorescence intensity.
- 4 **g**, Relative population heatmap of all brain regions outside the isocortex. The number of each cell type was normalized by the total number of cells

- 1 in each regions, derived from CUBIC-Atlas ($n = 4$ for PV, SST, ChAT and TH; $n = 7$ for Iba1).
- 2 **h**, Density of PV and Sst expressing cells in the isocortex. Data are shown as mean $\pm$ STD ($n = 4$).
- 3 **i**, From the cluster of ChAT expressing cells, the boundary surface including LDT was extracted ($n = 4$).
- 4 **j**, From the cluster of TH expressing cells, the boundary surface including LC was extracted ($n = 4$).
- 5 **k**, Merge of the two boundary surfaces (yellow = ChAT, magenta = TH, cyan = overlapping region).
- 6 **l**, The volume overlaps of the boudary surfaces.
- 7 Brain region acronyms follow the ontology defined by the Allen Brain Atlas.

Fig. 3 | Whole-brain Analysis of Aβ Plaques Accumulation in AD Model Mouse Brain

- 1 Using *App*^{NL-G-F/NL-G-F} AD model mouse brain (9- to 10-months-old), brain-wide accumulation of Aβ plaques were quantified by applying whole-mount 3D immunostaining and by using CUBIC-Cloud analysis framework.
- 2
- 3 **a**, Density of Aβ plaques (number of plaques/mm³) in major brain divisions ($n = 4$). Data are shown as mean $\pm$ STD ($n = 4$).
- 4 **b**, Volume ratio of Aβ plaque in major brain divisions ($n = 4$), computed as (total plaque volume in the region)/(region volume). Data are shown as mean $\pm$ STD ($n = 4$).
- 5
- 6 **c**, Distribution of effective radius of Aβ plaques in the isocortex, hippocampus (HPF), striatum (STR), midbrain (MB) and cerebellum (CB) ($n = 4$).
- 7 **d,e**, The volume ratio of Aβ plaques in the isocortex ($n = 4$). In **e**, data are shown as mean $\pm$ STD.
- 8 **f**, Layer-wise average of **d**. Data are shown as mean $\pm$ STD ($n = 4$).
- 9 **g**, Cartoon heatmap showing the Aβ plaque volume ratio in each brain region ($n = 4$).
- 10 **h-l**, Raw 6E10 immunostaining images around SNr **h**, VMH **i**, TRS and MEPO **j**, LDT and DTN **k** and MH, LH and PVT **l**.

Fig. 4 | Whole-brain Analysis of c-Fos Expression Level Changes by LPS Administration

- 1 LPS acutely induces sleep in mice. Brain-wide neural activity change induced by LPS was quantified by applying whole-mount 3D immunostaining
2 of c-Fos and by using CUBIC-Cloud analysis framework.
3 **a**, Whole-brain views of all cFos+ cells, showing saline (upper) and LPS (lower) administered brains. Each point (i.e. single cell) was assigned a
4 pseudo-color based on its fluorescence intensity.

1 **b**, Magnified 3D view of **a**, where the left isocortex was selectively displayed. Orientation arrows stand for R (right), D (dorsal) and P (posterior).
2 **c**, P-value heatmap showing the isocortex regions whose c-Fos+ cell density was significantly affected by LPS ($n = 4$ each). P-value was computed
3 by comparing the c-Fos+ cell count. The color lookup table is log scaled (base 10), where red color represents the regions that were activated (i.e.
4 more c-Fos+ cells) by MK-801, and blue represents the repressed regions. Regions with no statistical significance ($p > 0.05$) were assigned a gray
5 color.
6 **d**, Distinct brain regions activated by LPS ($n = 4$ each). The top row shows the voxel-wise p-value map. Color lookup table follows that of **c**. The
7 second and third rows are the raw c-Fos images of saline- and LPS-administered group, respectively. The forth row shows the number of c-Fos+
8 cells of the identified regions.
9 **e**, Plot of c-Fos+ cells in PVT. Cells are pseudo-colored with their intensity values. Pink (blue) dots are from LPS (saline) administered brains,
10 respectively.
11 **f**, The number of c-Fos+ cells in the PVT in 10 divisions along the anterior-posterior (AP) axis.
12 **g**, The number of c-Fos+ cells in the anterior and posterior half of the PVT. The boundary between anterior and posterior region was set at the
13 center of the PVT along AP axis.
14 **h**, The c-Fos expression levels per cell in the PVT 10 divisions along the anterior-posterior (AP) axis.
15 **i**, The c-Fos expression levels per cell in the anterior and posterior half of the PVT.
16 * $p < 0.05$, ** $p < 0.01$, Welch's t-test. Brain region acronyms follow the ontology defined by the Allen Brain Atlas.

Fig. 5 | Whole-brain Analysis of Input Cell Populations Projecting to ARH^{Kiss1+} Neurons

- a, Virus injection scheme. AAV carrying mCherry, TVA receptor and optimized glycoprotein (oG) was injected to ARH of Kiss1-Cre transgenic mouse, followed by injection of modified Rabies virus carrying GFP. Cells expressing both mCherry and GFP are the starter cells.
- b, Quantification of starter cell localization. The ratio was computed by dividing the cell count in each region by the total number of starter cells. The total number of starter cells each sample is shown on the right end of the heatmap.
- c, Whole-brain view of all input cells.
- d, Total cell count and the distribution of input cells. Only male brains were considered here.
- e, Cell density heatmap of all brain regions (excluding the isocortex and cerebellum, where virtually no input cells were detected). The mean of male and female brains are shown.
- f, The plot shows extremely sparse input cell populations in previously unidentified brain regions ($n = 3$). Only male brains were considered here.
- g, Raw GFP (black) and nuclear staining (RedDot2, purple) images showing the regions identified in f. Macro view (top) and zoomed-in view (bottom) are shown.
- Brain region acronyms follow the ontology defined by the Allen Brain Atlas.

Supplementary Figure 1 | Architecture of CUBIC-Cloud

1 Schematic diagram showing the architecture of CUBIC-Cloud (see Methods).

Supplementary Figure 2 | Imaging 1.0 μm -diameter Fluorescent Beads Embedded in Cleared Tissue

- 1.0 μm -diameter fluorescent beads embedded in cleared tissue were imaged using macro-zoom LSM (1X magnification, 6.5 μm voxel size) from the ventral side (see Methods).
- a**, Representative sagittal slice (virtually reconstructed from horizontal-major 3D image stack) showing the fluorescent beads embedded in tissue.
- b**, Bead spot profiles measured in six brain regions. Lateral and axial profiles were fitted with Gaussian, respectively, and the fitted curves are shown along with the raw data points (see Methods). The fitted sigma values (with 95% confidence interval) of the Gaussian are also shown. The number of particles used to average are shown in the graph.
- Brain region acronyms follow the ontology defined by Allen Brain Atlas.

Supplementary Figure 3 | Validation of Cell Counting and Registration Methods

- a**, The cell detection workflow overview (see Methods).
- b-h**, Accuracy evaluation of the cell counting algorithm. True positive rate (TPR), positive predictive value (PPV) and F1 score were evaluated for seven label types in more than five brain regions. Ground truth was prepared by two independent annotators (see Methods).
- i**, Normalized cross-correlation (NCC) value between two brains after registration. Mean NCC value of each coronal slice are plotted. 20 brains from different mice were independently registered onto CUBIC-Atlas. Individual profiles (thin lines with light colors) as well as the mean (thick green line) are shown.
- j**, Representative brain registration result. CUBIC-Atlas (cyan) and registered brain (magenta) are overlaid.
- k**, Voxel-wise NCC value map computed for the images shown in **j**.
- Brain region acronyms follow the ontology defined by Allen Brain Atlas.

Supplementary Figure 4 | Whole-brain analysis of PV, SST, ChAT, TH and Iba1 Expressing Cells

- a, Heatmap showing PV expressing cell density.
- b, Layer-wise average density of PV expressing cells.
- c, Heatmap showing SST expressing cell density.
- d, Layer-wise average density of SST expressing cells.

- 1 **e**, Heatmap showing ChAT expressing cell density.
- 2 **f**, Layer-wise average density of ChAT expressing cells.
- 3 **g**, Heatmap showing Iba1 expressing cell density.
- 4 **h**, Layer-wise average density of Iba1 expressing cells.
- 5 **i**, Mean Iba1+ cell density in the isocortex. Saline (gray) and LPS (blue) groups are compared. Data are shown as mean $\pm$ STD ($n = 7$ for each
- 6 group).
- 7 **i**, Mean Iba1 expression level per cell in the isocortex.
- 8 **j**, Heatmap showing mean Iba1 expression level per cell in all brain regions outside the isocortex, comparing saline and LPS groups.
- 9 Brain region acronyms follow the ontology defined by the Allen Brain Atlas.

Supplementary Figure 5 | Whole-brain analysis of c-Fos expression under LPS administration

- 1 **d**, Voxel-wise p-value heatmap showing the affected regions by LPS. P-values of individual voxels were computed by c-Fos+ cell count between
- 2 saline- and LPS-administered groups. The color lookup table is log scaled (base 10), where red color represents the regions that were activated
- 3 (i.e. more c-Fos+ cells) by LPS, and blue represents the repressed regions. Voxels with no significance ($p > 0.05$) were uncolored. Background is
- 4 nuclear staining image of CUBIC-Atlas for navigation. Step between consecutive slices is 0.34 mm.

Supplementary Figure 6 | Additional Data of Rabies Virus Tracer Experiment

a, P-value heatmap where the number of input cells were compared between male and female brains. The color lookup table is log scaled (base 10), where red color represents the regions where more input cells were found in female brains, and blue represents the inverse. Regions with no statistical significance ($p > 0.05$) were assigned a gray color.

b, Raw GFP (black) and nuclear staining (RedDot2, purple) images around lateral septal nucleus (LS) and medial preoptic area (MPO). The images are digitally reconstructed sagittal sections. Maximum intensity project (MIP) spanning 300 μm thickness.

c-h, The plot shows the normalized input cell count in regions where sexual dimorphisms were suggested. Of note, using RV injection and slice-based observation, [78] investigated the input cell population of pro-opiomelanocortin (POMC) neurons and agouti-related peptide (AgRP) neurons in the ARH, another dominant cell types in the ARH. The brain areas containing the input cells to ARH^{Kiss1+} neurons largely overlapped with those of POMC neurons and AgRP neurons. In some areas, however, interesting differences were observed. For example, in ventral tegmental nucleus (VTA) and nucleus incertus (NI), no input cells were detected for ARH^{Kiss1+}, while some input cells were reported to exist for POMC and AgRP. On the other hand, basomedial amygdalar nucleus (BMA), posterior amygdalar nucleus (PA), septofimbrial nucleus (SF), central amygdalar nucleus (CEA), substantia innominate (SI), triangular nucleus of septum (TRS), posterior intralaminar thalamic nucleus (PIL), midbrain reticular nucleus (MRN) and parabrachial nucleus (PB) contained small number of input cells to ARH^{Kiss1+} neurons (Fig. 5e,f,g), while they were not reported for POMC and AgRP neurons. This absence of input cells may reflect the actual biological differences, or it may reflect the superior sensitivity of our experimental methods to detect sparse populations. Further investigations are needed to draw conclusions.

* $p < 0.05$, ** $p < 0.01$; Welch's t-test.

Brain region acronyms follow the ontology defined by the Allen Brain Atlas.

Supplementary Table 2: Number of PV, SST, ChAT and Th expressing cells of adult mouse brains

- 1 The columns G to R list the number of PV+ cells ($n = 4$, data of each brain, the mean and standard deviations are shown) of all brain regions defined
- 2 in Allen Brain Atlas (CCFv3, version 2017), along with the total fluorescence intensity. The columns S to AD list the results for SST+ cells ($n = 4$).
- 3 The columns AE to AP list the results for ChAT+ cells ($n = 4$). The columns AQ to BB list the results for Th+ cells ($n = 4$).

Supplementary Table 3: Number and expression level of Iba1 expressing cells of LPS- or saline-administered brains

- 4 The columns G to X list the number of Iba1+ cells of saline-administered group ($n = 7$, data of each brain, the mean and standard deviations are
- 5 shown) of all brain regions defined in Allen Brain Atlas (CCFv3, version 2017), along with the total fluorescence intensity. The columns Y to AP list
- 6 the results for LPS-administered group ($n = 7$).

Supplementary Table 4: Number and volume of A β plaques of *App*^{NL-G-F/NL-G-F} mouse brain

- 7 The columns list the number of A β plaques ($n = 4$, data of each brain, the mean and standard deviations are shown) of all brain regions defined in
- 8 Allen Brain Atlas (CCFv3, version 2017), along with the total plaque volume and fluorescence intensity.

Supplementary Table 5: Number and expression level of c-Fos expressing cells of LPS- or saline-administered brains

- 9 The columns G to R list the number of c-Fos+ cells of saline-administered group ($n = 4$, data of each brain, the mean and standard deviations are
- 10 shown) of all brain regions defined in Allen Brain Atlas (CCFv3, version 2017), along with the total fluorescence intensity. The columns S to AD list
- 11 the results for LPS-administered group ($n = 4$).

Supplementary Table 6: Number of input cells projecting to ARH^{Kiss1+} neurons

- 12 The columns list the number of input cells (i.e. GFP+ and mCherry- cells) ($n = 3$ for male, $n = 4$ for female, data of each brain, the mean and
- 13 standard deviations are shown) of all brain regions defined in Allen Brain Atlas (CCFv3, version 2017), along with the total GFP intensity.

Supplementary Movie 1: Quick guide to CUBIC-Cloud

- 14
- 15 The movie shows the general usage of the CUBIC-Cloud.