Complete single neuron reconstruction reveals morphological diversity in molecularly defined claustral and cortical neuron types

Yun Wang[1], Peng Xie[1], Hui Gong[2,3], Zhi Zhou[1], Xiuli Kuang[4], Yimin Wang[5,6], An-an Li[2,3], Yaoyao Li[4], Lijuan Liu[6], Matthew B. Veldman[7], Tanya L. Daigle[1], Karla E. Hirokawa[1], Lei Qu[8,6], Phil Lesnar[1], Shengdian Jiang[6], Yang Yu[1], Wayne Wakeman[1], Shaoqun Zeng[2], Xiangning Li[2,3], Jing Yuan[2,3], Thuc Nghi Nguyen[1], Rachael Larsen[1], Sara Kebede[1], Yuanyuan Song[6], Lulu Yin[6], Sujun Zhao[6], Aaron Feiner[1], Elise Shen[1], Chris Hill[1], Quanxin Wang[1], Stephanie Mok[1], Susan M. Sunkin[1], Z. Josh Huang[9], Luke Esposito[1], Zizhen Yao[1], Michael J. Hawrylycz[1], Bosiljka Tasic[1], Lydia Ng[1], Staci A. Sorensen[1], X. William Yang[7], Julie A. Harris[1], Christof Koch[1], Qingming Luo[2], Hanchuan Peng[1], Hongkui Zeng[1,]*

[1]Allen Institute for Brain Science, Seattle, WA 98109, USA.

[2]Britton Chance Center for Biomedical Photonics, Wuhan National Lab for Optoelectronics, Huazhong University of Science and Technology, Wuhan, Hubei, China.

[3]HUST-Suzhou Institute for Brainsmatics, JITRI Institute for Brainsmatics, Suzhou, Jiangsu, China.

[4]School of Optometry and Ophthalmology, Wenzhou Medical University, Wenzhou, Zhejiang, China.

[5]School of Computer Engineering and Science, Shanghai University, Shanghai, China.

[6]Institute for Brain and Intelligence, Southeast University, Nanjing, Jiangsu, China.

[7]Center for Neurobehavioral Genetics, Jane and Terry Semel Institute for Neuroscience and Human Behavior; Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA.

[8]Key Laboratory of Intelligent Computation & Signal Processing, Ministry of Education, Anhui University, Hefei, Anhui, China.

[9]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

*Lead Contact and Correspondence: hongkuiz@alleninstitute.org

SUMMARY

Ever since the seminal findings of Ramon y Cajal, dendritic and axonal morphology has been recognized as a defining feature of neuronal types and their connectivity. Yet our knowledge about the diversity of neuronal morphology, in particular its distant axonal projections, is still extremely limited. To systematically obtain single neuron full morphology on a brain-wide scale in mice, we established a pipeline that encompasses five major components: sparse labeling, whole-brain imaging, reconstruction, registration, and classification. We achieved sparse, robust and consistent fluorescent labeling of a wide range of neuronal types across the mouse brain in an efficient way by combining transgenic or viral Cre delivery with novel transgenic reporter lines, and generated a large set of high-resolution whole-brain fluorescent imaging datasets containing thousands of reconstructable neurons using the fluorescence micro-optical sectioning tomography (fMOST) system. We developed a set of software tools based on the visualization and analysis suite, Vaa3D, for large-volume image data processing and computation-assisted morphological reconstruction. In a proof-of-principle case, we reconstructed full morphologies of 96 neurons from the claustrum and cortex that belong to a single transcriptomically-defined neuronal subclass. We developed a data-driven clustering approach to classify them into multiple morphological and projection types, suggesting that these neurons work in a targeted and coordinated manner to process cortical information. Imaging data and the new computational reconstruction tools are publicly available to enable community-based efforts towards large-scale full morphology reconstruction of neurons throughout the entire mouse brain.

INTRODUCTION

Like discovering the periodic table of elements in chemistry, and a taxonomy of living species in biology, obtaining an atlas of cell types is fundamental to neuroscience, essential for the understanding of brain function and its dysfunction in diseases. Brain cells, especially neurons, exhibit tremendous diversity across molecular, morphological, physiological, connectional and functional levels. A proper form of classification and creation of a cell type atlas needs to consider and integrate all these distinct cellular properties, obtained from thousands to millions of single cells in a systematic, consistent and comprehensive manner (Zeng and Sanes, 2017). Recent advances in high-throughput single cell RNA-sequencing have enabled systematic classification of cell types at the transcriptomic level (Saunders et al., 2018; Tasic et al., 2016; Tasic et al., 2018; Zeisel et al., 2018; Zeisel et al., 2015). This approach captures the essence of major cell types with known anatomical and functional properties, and reveals many new potential cell types. Systematic classification of cortical neurons at a combined morpho-electrical level has also been carried out (Jiang et al., 2015; Markram et al., 2015), and new technologies have been developed to correlate morpho-electrical and transcriptomic cell types (Cadwell et al., 2016; Fuzik et al., 2016).

Ramon y Cajal's foundational work revealed the extraordinary diversity of dendritic and axonal structures of neurons across brain regions and species using the Golgi technique, leading to the concept of neuronal types (Ramón y Cajal, 1909). Dendritic and axonal morphologies have long been held as the most defining features of neuronal types. Brain-wide inter-areal connectivity has been mapped extensively using bulk injections of anterograde and retrograde tracers to label populations of projection neurons (Gamanut et al., 2018; Oh et al., 2014; Zingg et al., 2014; Harris et al., bioRxiv preprint https://www.biorxiv.org/content/10.1101/292961v2, 2018). However, it remains largely unknown how population-level projection patterns are reflected at the single neuron level. Retrograde tracing studies suggest that individual neurons within a brain region often have heterogeneous axonal projection patterns, likely carrying diverse messages in the behaving brain (Minciacchi et al., 1985; Zingg et al., 2018). Thus, characterizing single neuron axonal projections through reconstruction of complete morphologies will provide critical information about how neural signals are organized and transmitted to their target regions. This knowledge is a critical component of any complete description of neuronal cell types.

Despite its importance, information about axonal morphologies is currently lacking for most projection neuron types in mammals, whose axons often cover large distances and are severed in studies based on *ex vivo* brain slices. Efforts have been made to fully label single neurons with small molecules or fluorescent proteins through *in vivo* whole-cell patching, *in vivo* electroporation (Han et al., 2018; Li et al., 2017), or sparse viral labeling with sindbis virus (Aransay et al., 2015; Ghosh et al., 2011) or adeno-associated virus (AAV) (Economo et al., 2016; Lin et al., 2018a; Lin et al., 2018b). This is conventionally followed by serial sectioning and imaging of each section, and manual reconstruction of the labeled neurons across many brain sections. Although relatively few such studies exist to date due to the highly laborious process, these studies reveal unique and novel features of specific projection neuron types that likely have important functional implications (Economo et al., 2018; Kita and Kita, 2012). The recent development of high-throughput and high-resolution fluorescent imaging platforms, such as fMOST (Gong et al., 2016) and MouseLight (Economo et al., 2016), coupled with more efficient sparse viral labeling strategies, are now enabling large-scale generation of neuronal morphology datasets. These studies also reveal a continued need for further improvements in tools for generating very sparse and strong labeling of single neurons at a brain-wide scale, as well as computational tools to expedite the laborious reconstruction process.

For our case study, we chose to reconstruct full morphologies of claustrum projection neurons and their close relatives in the cortex. The claustrum is a small, elongated structure in the cortical subplate, located medial to and beneath lateral cortical areas (*e.g.*, agranular insular cortex), just outside the striatum. It has widespread reciprocal connections with nearly all isocortical and retrohippocampal areas, with particularly strong and bilateral connections with the prefrontal cortex (including anterior cingulate, prelimbic, infralimbic, orbitofrontal and agranular insular areas) and the posterior lateral cortex (including entorhinal, perirhinal, and lateral higher visual areas) (Wang et al., 2017; Zingg et al., 2018). Based largely on connectional evidence, it was postulated that the claustrum could play important roles in multimodal integration of stimulus information into a single conscious percept (Crick and Koch, 2005), amplification of cortical oscillations (Smythies et al., 2012, 2014), salience and novelty detection (Remedios et al., 2010, 2014), selective attention (Goll et al., 2015; Mathur, 2014), and regulation of vigilance states. A triptych of *in vivo* and *in vitro* studies in mice support the role of the claustrum in the inhibitory control of relevant cortical excitatory neurons during attentional control (Atlan et al., 2018; Jackson et al., 2018; Narikiyo et al, bioRxiv preprint https://www.biorxiv.org/content/10.1101/286773v1, 2018). Recently, a patch-clamp electrophysiological study identified two spiny projection neuron subtypes and three aspiny interneuron subtypes in the claustrum, and found that visual cortex, parietal association cortex, and anterior cingulate cortex receive input from the two types of projection neurons in differing proportions (White and Mathur, 2018). Gene expression studies suggest that claustrum projection neurons bear molecular resemblance to endopiriform nucleus neurons and a group of neurons present in multiple lateral cortical areas, sharing a number of common marker genes (Smith et al., 2018). To better understand the role of the claustrum in regulating cortical function, it is necessary to study the morphology and projection pattern of individual claustrum neurons, and understand how they specifically and cooperatively cover the cortical mantle and their relationship with molecular profiles.

We established a systematic pipeline that allows us to label, image, reconstruct and classify single neurons on a brain-wide scale using complete morphology data. Since this remains a highly challenging task for any single group, we make all the imaging data and computational reconstruction tools publicly available. We aim to enable a community-based effort to generate a sufficiently large number, *e.g.*, tens to hundreds of thousands or millions, of full neuronal morphology reconstructions throughout the mouse brain, to facilitate meaningful cell type classification using soma locations, dendrites, and local and long-distance axon morphological features and gain a comprehensive understanding of cell type diversity. Here, we demonstrate proof of principle for this pipeline by reconstructing and classifying full morphologies of 96 neurons from the claustrum and cortex. Even with this relatively small set of full reconstructions, we uncovered surprising and striking morphological diversity within the relatively homogeneous molecularly-defined claustral and cortical cell subclass. This intriguing result lends support for the need to approach the cell type classification problem from different perspectives, reconciling and integrating multi-modal information for a unified understanding of cell type.

RESULTS

Sparse, robust and consistent neuronal labeling

Previous genetic approaches to achieve sparse or single neuron labeling include viral delivery (*e.g.*, using sindbis virus or AAV) and *in vivo* electroporation (Aransay et al., 2015; Economo et al., 2016; Han et al., 2018; Li et al., 2017; Lin et al., 2018a; Lin et al., 2018b). Both approaches have substantial cell-to-cell and animal-to-animal variations, and are usually restricted to few brain regions. To achieve more efficient, widespread and consistently sparse yet strong labeling, we utilized TIGRE2.0 transgenic reporter lines shown to exhibit viral-like transgene expression levels (Daigle et al., 2018), coupling them with Cre expression from either Cre driver lines or viral Cre delivery. Two general strategies were employed. The first was to use the GFP expressing Ai139 or Ai140 reporter line in conjunction with sparse Cre-mediated recombination (**Fig. 1A**). We used CreERT2 driver lines, and titrated the level of CreERT2-mediated recombination using low-dose tamoxifen induction (**Table S1**). We found optimal doses for sparse labeling in each case by screening for expression using serial two photon tomography (STPT) (**Fig. 1C-E**). Alternatively, to achieve brain-wide sparse labeling, we delivered pSyn-Cre via retroorbital injection of diluted AAV with the PHP.eB serotype (Chan et al., 2017) (**Fig. 1C**). We further tested a combination of TIGRE2.0 (Ai140) and TIGRE1.0 (Ai82) reporter lines, using the tTA2 from a single TIGRE2.0 allele to drive two copies of TRE promoter driven GFP expression cassettes (**Fig. 1A**). We found that this strategy generates even higher level of GFP expression, well suited for fMOST imaging (see below).

The second strategy was to use a new TIGRE2.0 line, TIGRE-MORF (also named Ai166; Veldman et al, in submission), which is sparsely activated in conjunction with Cre delivery (**Fig. 1B**). TIGRE-MORF/Ai166 expresses the MORF gene, which is composed of a farnesylated EGFP (GFPf) preceded by a mononucleotide repeat of 22 guanines (G_{22}-GFPf; Veldman et al, in submission). The GFPf transgene is not translated at the baseline due to the out-of-frame of the G_{22} repeat relative to the open reading frame of GFPf, which lacks its own translation start codon. However, during DNA replication or repair, rare events of stochastic frameshift of the mononucleotide repeat could result in correction of the translation frame (*e.g.*, G_{22} to G_{21}) and expression of the GFPf protein in a small subset of the progeny cells. In our companion study (Veldman et al, in submission), TIGRE-MORF and three other MORF mouse lines all exhibit a labeling frequency of 1-5% when they are crossed to different Cre mouse lines. Even with this frequency, we found that combining TIGRE-MORF/Ai166 with many Cre driver lines densely expressing the Cre transgene did not produce sufficient sparsity to readily untangle the axonal ramifications, whereas combining it with Cre lines that are already relatively sparse to begin with, or with CreERT2 lines with intermediate dosing level of tamoxifen (**Table S1**), did lead to extremely sparse labeling that is well suited for reconstruction of even very elaborate axonal arborizations of many neuronal types (**Fig. 1F-L**). The use of membrane associated GFPf also enabled robust labeling of very thin axon fibers. Leaky background expression of GFP reported in other TIGRE2.0 lines (Daigle et al., 2018) is not present in TIGRE-MORF/Ai166 mice due to the strict dependency of translational frameshift for the expression of GFPf reporter, making TIGRE-MORF/Ai166 (below simplified as Ai166) an ideal reporter line for sparse and strong labeling of various neuronal types across the brain.

fMOST imaging

We acquired whole brain images with sufficient resolution (~0.3 x 0.3 x 1 μm XYZ) for reconstructing fine-caliber axons, using fluorescence micro-optical sectioning tomography (fMOST), a high-throughput, high-resolution, brain-wide fluorescent imaging platform. In this approach, a GFP-labeled brain is first embedded in resin. The resin-embedded GFP fluorescence can be recovered through chemical reactivation (Xiong et al., 2014) provided by adding Na_2CO_3 in the imaging water bath. Thus, a wide-field block-face imaging system can be employed to maximize imaging speed. Following imaging of the entire block-face, the top 1-μm tissue is sliced off by a diamond knife, exposing the next face of the block for imaging. For the entire mouse brain, a 15-20 TB dataset containing ~10,000 coronal planes of 0.2-0.3 μm X-Y resolution and 1 μm Z sampling rate is generated.

So far we have generated >46 high-quality fMOST datasets for a number of cortical, thalamic, claustral, striatal, cholinergic, noradrenergic and serotonergic neuronal types (**Table S1**). It is worth noting that this approach can be extended to any other cell types for which appropriate Cre-dependent labeling methods are available, and we continue to test new Cre lines and other sparse Cre delivery approaches to generate novel fMOST datasets. **Figure 2** shows representative images acquired using fMOST for a wide variety of neuronal types and their dendritic and axonal arborizations, visualized using 100-μm maximum intensity projection (MIP) images (i.e., projected from 100 consecutive image planes) (for more examples see **Fig. S1**). In the cortex, we imaged neurons suitable for single cell reconstructions from different excitatory projection classes (Tasic et al., 2018). For example, Cux2-CreERT2;Ai166 labeled the cortical L2/3/4 intratelencephalic (IT) subclasses of excitatory neurons (**Fig. 2A** and **Movie S1**). Plxnd1-CreER;Ai166 labeled the cortical L2/3 and L5 IT subclasses, as well as striatal medium spiny neurons (**Fig. 2C** and **Movie S2**). Fezf2-CreER;Ai166 labeled the cortical L5 pyramidal tract (PT) subclass (**Fig. 2B** and **Movie S3**). Nxph4-T2A-CreERT2;Ai166 labeled the cortical L6b subplate neurons (**Fig. S1A**). We also labeled and imaged cells from the cortical *Pvalb*+ subclass of inhibitory interneurons, in addition to a subset of L5 PT excitatory neurons, using Pvalb-T2A-CreERT2;Ai166 (**Fig. 2G** and **Movie S4**), and the *Sst*+ subclass of interneurons using Sst-Cre;Ai166 (**Fig. S1B**). Gnb4-IRES2-CreERT2;Ai140;Ai82 labeled the *Car3*+ IT subclass of claustral and cortical excitatory neurons (**Fig. 2D** and **Movie S5**, also see below). Several cell types containing neuromodulators were also labeled and imaged, including noradrenergic neurons in the locus ceruleus using Dbh-Cre_KH212;Ai166 (**Fig. 2E**), and serotonergic neurons in the dorsal raphe and brainstem using Slc6a4-CreERT2;Ai166 (**Fig. 2F** and **Movie S6**). Tnnt1-IRES2-CreERT2;Ai140;Ai82 labeled thalamic excitatory projection neurons as well as striatal medium spiny neurons (**Fig. 2H** and **Movie S7**). Vipr2-IRES2-Cre-neo;Ai166 also labeled thalamic excitatory projection neurons but with an enrichment in the visual thalamic nucleus, the dorsal lateral geniculate nucleus (LGd), as well as retinal ganglion cells and cortical chandelier cells (**Fig. 2I** and **Movie S8**). Many of these Cre lines also sparsely label other populations of neurons in other parts of the brain, which are not described in detail here.

Each of these brains contains ~100-1,000 labeled neurons (**Table S1**). Thus, overall, tens of thousands of neurons could be reconstructed from these and additional datasets in the coming years. These datasets are already, or will be made, publicly available under the BRAIN Initiative Cell Census effort (https://biccn.org/) as a unique resource for the community.

Even without reconstruction, it is apparent that these neurons display a remarkable array of dendritic and axonal morphologies. Traditionally, it has been a well-known technical challenge to fully visualize very fine axon fibers. But in each example case shown here, the axonal labeling appears to be complete as judged by the visibility of a terminal bouton at the end of each axon fiber (see arrowheads in **Fig. 2**). We assess each image series and cell to make sure the labeling fills all the way to an identifiable end. If it doesn't, we don't consider it complete. Specifically, in these sparsely labeled brains, cortical IT and PT neurons not only have main long-range projections but also local axonal branches that are well segregated and clearly identifiable, enabling truly complete reconstruction of the entire local and long-range, cortical and subcortical axonal arborization (**Fig. 2A-C**). L5 PT neurons form the 'driving' type of synapses in the thalamus (Guo et al., 2017; Sherman, 2016), which can be seen as enlarged and intensely fluorescent boutons (**Fig. 2G**). L6b subplate neurons extend their axons upwards into layer 1 (Fig. **S1A**). The axons of thalamic projection neurons form dense or dispersed clusters in the cortex (**Fig. 2H-I**). On the other hand, claustral, noradrenergic and serotonergic neurons have widely dispersed, thin axons that are still labeled well (**Fig. 2D-F**). One can also clearly see individual axons in substantia nigra from striatal medium spiny neurons (**Fig. 2C,H**), individual axon terminal clusters in superior colliculus likely from retinal ganglion cells (**Fig. 2I**), as well as the very dense and fine local axonal branches of a variety of cortical and striatal interneurons (*e.g.*, basket cells, chandelier cells, and Martinotti cells) (**Fig. 2G,I** and **S1B**). The consistency and high quality of brain-wide labeling and imaging demonstrates the wide applicability of our approach in studying morphologies of diverse neuronal types throughout the brain.

Pipeline for image data processing, morphology reconstruction and registration

We established a standardized image data processing and whole mouse brain morphology reconstruction process (**Fig. 3**), utilizing Vaa3D, an open-source, cross-platform visualization and analysis system (Peng et al., 2010). Each fMOST dataset is first converted to a multi-level navigatable dataset using the Vaa3D-TeraFly program (Bria et al., 2016), which allows smooth handling of terabyte-scale datasets. Morphology reconstruction is then carried out on the TeraFly files within Vaa3D. A series of tools were developed within Vaa3D to facilitate semi-automated and manual reconstruction, as well as brain registration and analysis. In particular, a virtual reality (VR) environment created within Vaa3D, named TeraVR, significantly enhanced the manual reconstructor's ability to see the 3D relationships among intertwined axonal segments, and thus improved the precision and efficiency of morphology reconstruction (Wang et al., bioRxiv preprint https://doi.org/10.1101/621011, 2018). To produce an accurate reconstruction of a neuron, TeraVR enabled several annotators working collaboratively on the same neuron to discuss around uncertain structures. After completion of the quality control (QC) check and manual correction, Auto-Refinement fitted the tracing to the center of fluorescent signals and filled any pixel gaps as the last step of reconstruction. The final reconstructed morphology was completed as a single tree without breaks, loops, multiple branches from the same point, etc.

In parallel, each fMOST dataset was registered to the 3D Allen mouse brain reference atlas, the Common Coordinate Framework (CCFv3, http://atlas.brain-map.org/), using a process specifically designed for fMOST datasets (**Fig. S2**) to handle the challenges of brain shrinkage and deformation after fixative perfusion and resin embedding, and stripe artifacts due to diamond knife cutting. Following the registration of the whole-brain fluorescent image dataset, all the individual reconstructions from the brain were also registered to the CCF. The registration enables digital anatomical delineation and spatial quantification of each reconstructed morphology and its compartments (*e.g.*, soma, dendrites, axon arbors). Co-registration of multiple brains into the common 3D space allows analysis and comparison of morphological features of neurons across different brains.

Reconstruction and classification of *Gnb4*+ claustral and cortical neurons

As a proof of principle, we first studied the morphological diversity of neurons labeled in the Gnb4-IRES2-CreERT2 line. *Gnb4* is a marker gene that selectively labels neurons in the claustrum, endopiriform nucleus, and a population of neurons in the deep layers of lateral cortical areas (Wang et al., 2017) (**Fig. S3**). Previous population-level axonal projection mapping showed that claustrum neurons predominantly target, and have extensive reciprocal connections with, neocortical and allocortical areas, with particularly strong interconnections with prefrontal and retrohippocampal cortical areas (Wang et al., 2017). We first confirmed that Gnb4-IRES-Cre labeled cells in the claustrum have this same widespread cortical projection pattern, including preferential targeting to prefrontal and medial cortical areas, using bulk labeling with the Cre-dependent AAV2/1-pCAG-FLEX-GFP tracer (**Fig. 4A-B**). *Gnb4*+ cortical neurons were mapped exclusively to the previously described (Tasic et al., 2018) L6 IT Car3 subclass of cortical excitatory neurons by single-cell transcriptomics (see next section below). Similar projection experiments with the same AAV tracer injected in lateral cortical areas (primary or secondary somatosensory cortex) in Gnb4-IRES2-CreERT2 mice, which labeled axons from deep L6 neurons, also showed intracortical projections, but with a more restricted, distinct set of targets compared to claustrum projections (**Fig. 4C-E**). Note that these 5 selected datasets all had small, spatially specific, injection sites that were located very close to each other. These small bulk injections demonstrate very distinct projection patterns between claustral and cortical *Gnb4*+ neurons.

To sparsely label claustral and cortical projection neurons, we generated 4 fMOST datasets using the Gnb4-IRES2-CreERT2 line crossed with Ai139 (#236174), Ai140 (#17109), and Ai82;Ai140 (#17781 and 17782) (**Fig. 5A** and **Table S1**). From these 4 brains, we reconstructed 34 claustrum (CLA) projection neurons and 62 cortical neurons (**Table S2**). For most cells, distinguishing CLA and cortical neurons was relatively easy using soma locations. We assigned CLA cell identity based on their soma locations within the claustrum region, whereas the *Gnb4*+ pyramidal cells (PCs) were found in various cortical regions, largely all in deep layer 6 ("L6PCs"). Following registration to the CCF, the location and spatial distribution of each reconstructed neuron and its dendritic and axonal arbors could be visualized in the 3D reference atlas space. The somata of these neurons spanned the entire 3.2-mm anteroposterior extent of claustrum and a much larger range of lateral cortical areas (**Fig. 5A**). The full axonal morphology of each reconstructed neuron is shown in **Figure S4** for CLA cells and **Figure S5** for L6PCs. It is important to note that all these neurons projected exclusively into the cortex. None had axon projections into the striatum, demonstrating a major difference between these *Gnb4*+ claustral and cortical neurons and other types of corticocortical IT neurons which also have axon collaterals projecting to the striatum (Harris and Shepherd, 2015).

The dendrites of CLA cells and L6PCs were visually and quantitatively distinguishable from each other based on several features (**Fig. 5B**). The dendrites of CLA cells were very flat, with more extended processes along the anteroposterior axis and the thinnest dimension in the mediolateral axis, in keeping with the shape of claustrum and embedded in the axon bundle of CLA cells traversing through the claustrum. The dendrites of L6PCs were more extended compared with the CLA dendrites, although they were also relatively flattened at a plane parallel to the white matter. They rarely had a typical apical dendrite extending vertically towards pia, rather, their "apical" dendrite as defined by branching at the highest order was often obliquely or horizontally oriented, parallel to the white matter. In addition, the L6PCs often had a big basal dendrite, resembling the bitufted PCs in layer 6a, a pyramidal cell type having claustral projections (Wang et al., 2018; Zhang and Deschenes, 1997). Quantitative analyses of dendritic features confirmed that L6PCs had significantly greater depth and depth-to-width ratio, but smaller number of branches, compared to CLA cells (**Fig. 5B** and **Table S3**).

The axons of CLA cells and L6PCs also had distinct features both globally and locally (**Fig. 5C, S4, S5** and **Table S3**). Quantitative analyses of the entire axonal arbors showed that CLA cells had significantly greater scale (width and height), Euclidean distance and path distance to soma. The axons of CLA cells emerged from the somata with no or only a couple of local collaterals, and typically bifurcated proximally into a few major, thick axon stems which often projected in opposite directions (forward and backward) and traveled through the claustrum in parallel to the white matter. On the other hand, L6PCs had significantly more elaborate local axon branches forming a cluster around the soma, and projected to one or multiple target cortical areas. The differences in both dendritic and axonal morphologies suggest that the claustral and cortical *Gnb4*+ neurons are distinct.

We first classified cells based on expert manual annotation of axonal projection patterns (**Fig. 6**). The CLA neurons can be classified into a total of 4 types, including 2 major types and 2 subtypes within each major type (**Fig. 6A**). CLA type I cells (CLA_I) preferentially project to midline cortical areas. CLA type II cells (CLA_II) have distal, lateral cortical projections, avoiding the midline cortical areas. Within each major type, neurons were grouped into an ipsilateral-only projecting subtype (CLA_I-ipsi and CLA_II-ipsi) and a bilateral projecting (to both ipsilateral and contralateral hemispheres) subtype (CLA_I-bi and CLA_II-bi). Most of the reconstructed CLA cells belonged to the CLA_I types. CLA_II-bi appeared to be rare among different CLA types.

We also observed that projection patterns were topographically organized from anterior to posterior claustrum within individual types (**Fig. 6C**). The somata of CLA_I-ipsi cells were distributed through the full length of the claustrum, while those of other types were located mostly at the anterior and middle parts of the claustrum. The CLA_I-ipsi cells located in the anterior claustrum typically extended their axons from anterior to posterior midline cortical areas. The CLA_I-ipsi cells located in the middle claustrum typically formed a tighter axon cluster only in the anterior half of midline cortical areas, with an intense innervation to the prelimbic (PL) and frontal pole (FRP) areas. They also more frequently projected backward to entorhinal (ENT) and perirhinal (PERI) areas forming an axon cluster there. Remarkably, individual axons of the CLA_I-ipsi cells located in the posterior claustrum were found to wrap around the entire ipsilateral cortex (not forming a loop though), and hence named "crown of thorns" neurons. Those cells often formed axon clusters in ENT, PL and visual cortical areas in addition to a cluster in the midline of cortex. CLA_I-bi cells projected bilaterally forming an axon cluster on each side of the midline cortical areas. Typically, the two axon clusters on two sides were asymmetrical, predominant in the ipsilateral hemisphere. Those CLA_I-bi cells located relatively anteriorly often projected to ENT, PL and somatosensory cortical areas. Those from more middle locations had asymmetrical axon clusters along midline of both hemispheres but limited projections to other cortical areas.

Similarly, the cortical L6PCs were also manually classified into two types, ipsilateral-only projecting (PC_ipsi) and bilateral projecting (PC_bi) (**Fig. 6B**). The PC_ipsi type usually had their axonal projections in different cortical areas (within the same hemisphere) from their soma location, whereas the PC_bi type usually had their axonal projections in the homotypic cortical area on the contralateral hemisphere, forming two largely symmetrical axon clusters on both sides. For the latter type, occasionally additional projections were formed in other cortical areas in the ipsilateral hemisphere. Quantitative analysis showed that PC_ipsi cells had significantly greater numbers of axon branching nodes, ends, segments, and total length. Thus, compared to PC_bi, PC_ipsi had more complex axonal arborizations. Even though we only manually divided L6PCs into 2 morphological types, we observed an even greater degree of topographic diversity of axon projection patterns from anterior to posterior cortical areas for both ipsilateral and bilateral projecting types (**Fig. 6D**).

We next established a data-driven clustering-based morphology classification approach. Clustering analysis (see **Methods**) was performed using four feature sets: projection pattern (**Table S4**), soma location (**Table S2**), axon morphology and dendrite morphology (**Table S3**). A total of 14 morphological clusters were identified, containing 30 CLA and 52 L6PC neurons (**Fig. 7A,B** and **Fig. S6-S7**). A total of 14 cells were excluded by outlier removal because they were each morphologically unique to themselves due to the small sample size, see **Methods**. We found that the segregation of the clusters was mainly driven by their projection targets (**Fig. 7B** and **S6**). These clusters generally corresponded well with the manually classified morphological types (**Fig. 7A-C**), and further enabled subdivision of the neurons based on their specific projection targets (**Fig. 7D-F**). We calculated the total number of projection targets (ipsilateral and contralateral targets counted separately) contacted by each neuron using two different thresholds to label a region as "targeted", a minimum of 1,000 µm of axon length (same as the threshold used in a previous study, (Han et al., 2018)) or a minimum of 1 axon terminal (a more stringent threshold to minimize contribution from passing fibers) (**Fig. 7E**). The median total number of projection targets is 29 or 21 for CLA neurons and 18 or 12 for L6PCs, in either case substantially greater than that reported previously for primary visual cortex L2/3 IT neurons (Han et al., 2018).

L6PCs were assigned to a larger number of clusters. Clusters 7, 10, 11, 12 and 13 belonged to the PC-ipsilateral group (**Fig. 7F**). Clusters 5, 6 and 14 belonged to the PC-bilateral group. The clusters were arranged topographically from anterior to posterior cortex based on both soma location and projection target specificity, in other words, each cluster contained a group of neurons that were located close to each other and projected to similar cortical target areas. This is consistent with the above observation from manual classification even though only two morphological types were defined manually. Thus, both manual and computational approaches identify extensive diversity of L6PC neuronal morphologies and projections.

Clusters 1, 2, 3 and 8 belonged to the CLA-ipsilateral group, in which cluster 1 was lateral-projecting and clusters 2, 3 and 8 were midline-projecting (**Fig. 7F**). These computationally derived clusters also revealed topographical projection patterns consistent with observations made in the above manual classification. Cluster 4 contained CLA-bilateral midline-projecting cells. Cluster 9 contained a mixture of CLA cells and L6PCs, as they showed highly similar axon projection features and soma locations. This cluster had preferential projection to the FRP and ORBl areas on the ipsilateral side.

Taken together, the morphological features of fully reconstructed claustral neurons and L6PCs suggest that long-range axonal projections vary according to soma locations, indicative of a topographic organization of structural connectivity networks in claustrum and cortex.

Classification of Car3 (*Gnb4*+) subclass of claustral and cortical neurons by single-cell RNA-sequencing

We previously established a standardized single-cell RNA-sequencing (scRNA-seq) pipeline using the SMART-Seq v4 method to profile cells isolated from various brain regions (Tasic et al., 2018). Here, we analyzed 1062 cells isolated from claustrum and overlying cortical areas of Gnb4-IRES2-CreERT2;Ai140 and Gnb4-IRES2-CreERT2;Ai82;Ai140 mice by fluorescence activated cell sorting (FACS). We performed a large-scale co-clustering analysis by combining the scRNA-seq data from these cells and ~74,000 other cortical and hippocampal cells using our established analysis procedure and criteria (Tasic et al., 2018). Out of ~290 clusters, we found one major branch at the subclass level that contains nearly all the claustral and cortical neurons isolated from the Gnb4-IRES2-CreERT2;Ai140 and Gnb4-IRES2-CreERT2;Ai82;Ai140 mice (**Fig. 8A**). This subclass expresses a unique marker gene *Car3* and contains all the cortical neurons previously identified from our VISp-ALM single-cell transcriptomic study that belonged to the L6 IT Car3 subclass of glutamatergic excitatory neurons. In addition, this subclass also contains neurons from several other lateral cortical areas including agranular insular cortex (AI), supplementary somatosensory cortex (SSs) and temporal cortex (TEa) that were isolated from various Cre lines including Slc7a7-IRES2-Cre, Cux2-CreERT2 and Esr2-IRES2-Cre. Since this subclass contains both cortical and claustral neurons, we now rename it as the Car3 subclass (dropping the specific reference to cortical L6 and IT in (Tasic et al., 2018)).

This Car3 subclass branch (total 1,997 cells) consists of one major cluster that contains 1,206 cells from both claustrum and cortex, plus six additional minor clusters (15-236 cells each) (**Fig. 8A**). As shown in the tSNE plot (**Fig. 8A** lower right panel), the claustral and cortical neurons are intermingled together in the major cluster, whereas one of the minor clusters, Car3 TEa-PERI-ECT_GU Hgf, contains mostly cells from TEa. The co-clustering of claustral and cortical *Car3*+ *Gnb4*+ neurons suggests that they are highly related to each other, possibly reflecting a common developmental origin.

In an attempt to link molecular identities with the morphological diversity described above, we performed Retro-seq (Tasic et al., 2018) on cells isolated from claustrum and several cortical areas (*i.e.*, SSs, VISp, AI and TEa) that were labeled by retrograde tracers injected into medial (e.g., anterior cingulate area, ACA) or lateral cortical areas (e.g., lateral orbital area, ORBl, or posterolateral visual area, VISpl). Cells were isolated from both hemispheres – ipsilateral or contralateral to the injection site. These Retro-seq cells were included in the above large-scale clustering analysis together with all other cells.

To try to better resolve subtle differences between cells with different projectional and morphological properties and thus uncover potential molecular correlates of the morphological diversity, we then re-clustered the Car3 subclass cells (1,997 cells total, 298 of which from Retro-seq, all cells contributing to this clustering analysis are described in **Table S5**) with a more relaxed statistical criterion (see **Methods**). This approach resulted in 19 clusters that belonged to 5 main branches (**Fig. 8B**). As seen in the cell distribution across different areas below the dendrogram and in the tSNE plots (**Fig. 8B,C**), these clusters were quite continuous. There was no clear separation between CLA and cortical Car3 neurons. There was even a certain degree of segregation driven by Cre driver lines (*i.e.*, cells isolated from the same regions but different Cre lines were segregated by Cre lines, **Fig. 8B** rightmost panel in comparison with **Fig. 8C** leftmost panel). Thus, we intentionally reached a level of cluster resolution high enough to even detect technical variations in order to search for gene expression correlates of morphological diversity.

We then focused on the cluster distribution of Car3 Retro-seq cells. We found that, even under such relaxed clustering conditions, CLA and cortical cells projecting to various cortical areas, ACA, ORBl or VISpl, were located closely together with no clear segregation (**Fig. 8C** middle panel). Furthermore, even though there were only a few Car3 Retro-seq cells (14 cells) isolated from the contralateral side of the tracer injection sites, they were completely intermingled with Car3 Retro-seq cells isolated from the ipsilateral side (284 cells) of the injection sites (**Fig. 8C** right panel), suggesting that ipsilateral- and bilateral-projecting cells were also molecularly indistinguishable using the current criteria.

Thus, our study reveals striking morphological diversity within the molecularly defined cortical and claustral Car3 (and *Gnb4*+) cell class/type.

DISCUSSION

To fully understand the morphological and projectional specificity of neurons across the brain, it is generally expected that a large number, likely in the range of hundreds of thousands to millions, of neurons will need to be examined. Approaches such as MAPseq (Kebschull et al., 2016) can be used to quickly survey projection specificity at the regional level for many neurons in a high throughput manner. However, many essential details can only be obtained through full morphological reconstruction. Collecting such ground truth data will provide an invaluable opportunity to uncover principles of neuronal diversity and brain circuit organization, and inform us how functional studies should be conducted. To do this, a systematic and reproducible technology platform needs to be established. Here we report our effort in setting up such a platform that can be applied to potentially any neuronal type in any part of the mouse brain.

Our labeling strategy using stable and universal transgenic reporter mouse lines coupled with a variety of sparse Cre delivery methods has several advantages. First, we showed that the TIGRE2.0-based transgenic reporter lines, especially Ai166 which expresses a farnesylated GFP, produce very bright GFP labeling of axon fibers under fMOST imaging, revealing numerous terminal boutons, an essential requirement for obtaining truly complete, full morphologies. Second, this strategy enables sparse labeling across multiple regions within the same brain, improving efficiency compared to other methods (*e.g.*, *in vivo* electroporation or stereotaxic virus injection). Third, the labeling is highly consistent from cell to cell, cell type to cell type, region to region, and brain to brain, reducing variability often seen in other methods and creating a truly reproducible platform. Finally, sparse Cre recombination can be achieved through the use of transgenic Cre or CreERT2 driver lines labeling any neuronal type, or low-dose Cre viral vectors delivered through either local or systemic (*e.g.*, retroorbital) injections. In the future, sparse intersectional reporter lines like Ai166 can be developed, and their use in combination with dual driver lines can further enhance cell type targeting specificity and/or flexibility.

Development of novel and accessible software tools are also essential for reconstruction efficiency. The enhanced Vaa3D-based reconstruction toolkit streamlines large-volume image data processing and computation-assisted manual reconstruction. We have developed a pipeline of tools to facilitate this process. Given the generally rapid progress in algorithm and software development, abundant opportunities also exist to develop even more tools to expedite the reconstruction process. The registration of the fMOST whole-brain datasets to the CCF allows quantification of projection strength in each target region across the entire brain for each neuron, and subsequent clustering analysis to identify the similarities and differences between neurons and to group them into types. The high signal intensity and low background in the fMOST fluorescent datasets allows the generation of high-quality reconstructions. In the future, with the accumulation of an increasingly larger set of reconstructed full morphologies, these can be used as training datasets to develop machine learning-based automatic reconstruction algorithms that have the potential to dramatically increase the throughput of reconstruction.

We are making all the fMOST imaging data publicly available through the BRAIN Cell Data Center (BCDC), and Vaa3D-based computational reconstruction tools are open-source. Our hope is that this will enable a community-based effort to collectively generate a sufficiently large number of full neuronal morphology reconstructions throughout the mouse brain, to facilitate a comprehensive understanding of the diversity and specificity of local and long-range connectivity at the single cell level across different neural circuits.

In a proof-of-principle case, we reconstructed the full morphologies of 96 neurons from the claustrum and cortex that belong to a single transcriptomically defined neuronal subclass, the Car3 subclass. We first found that both the dendritic morphology and the axonal projection target fields are substantially different between claustral and cortical Car3/Gnb4 neurons. Claustral neurons have significantly flatter dendritic arborizations than cortical Car3 neurons. Both types of neurons can be further divided into ipsilateral projecting only and bilateral projecting subtypes according to their axonal targets. Claustral neuron axons of the ipsilateral types travel over long distances mainly targeting distal cortical regions such as prefrontal and retrohippocampal cortical areas, whereas cortical Car3 neurons mainly target cortical regions around soma and project to other regions usually closer than CLA axons project. Claustral neuron axons of the bilateral types typically form asymmetrical clusters between the two hemispheres, while Car3 neurons typically form symmetrical clusters on both sides. We found that claustral neurons exhibit further diversity based on each neuron's unique axonal projection pattern. They can be grouped into medially projecting and laterally projecting subtypes. These two subtypes are intermingled together at the anterior and middle parts of the claustrum. But those CLA cells reconstructed from the posterior claustrum all belong to the medially projecting subtype, which we called "crown of thorns" neurons because their axons wrap around the entire ipsilateral cortex. There is also topographic variation of projection target fields along the anteroposterior axis among CLA. Collectively they cover almost the entire cortical surface, with particularly dense axonal projections into the medial and lateral prefrontal cortex and retrohippocampal cortex. Overall, we suggest that they work in a targeted and coordinated manner to affect cortical function. The cortical Car3 neurons exhibit an even greater degree of topographical specificity, with both source and target region-specific projection patterns.

In contrast to the morphological diversity, we observed very modest difference in single-cell transcriptomics among claustral neurons and cortical Car3 neurons coming from different cortical areas, even under more relaxed clustering conditions. The large diversity of axonal morphologies and projection patterns observed from less than 100 fully reconstructed neurons from a transcriptomically relatively homogeneous cell population is striking. It underscores the necessity of scaling up the full neuronal morphology characterization effort to thousands to millions of neurons across all neural circuits and pathways to gain a true understanding of the extent of such diversity. Such knowledge is foundational for the understanding of brain connectivity and function. A critical bottleneck in this effort is the slow process of morphology reconstruction, a problem we believe is solvable using advanced computational approaches on high-quality imaging data.

The apparent lack of correlation between transcriptomic and morphological profiles in the Car3 subclass of neurons at the current stage is also intriguing. It is possible that the current unsupervised clustering approach is insufficient to uncover the genes specifically relevant to morphology out of the overall gene expression variations. Alternatively, it is also possible that morphological/connectional specificity is established during circuit development and its gene expression correlate might also exist only at that time. In either case, the result emphasizes the importance to perform single cell characterization in multiple modalities and take an integrated approach to describe and classify cell types in an unbiased and comprehensive manner. For example, in this case, the cell type classification needs to incorporate both types of information that has been independently obtained. One classification scheme could be constructed with the major branch (first order) being the Car3 molecular subclass, the second order branches being the claustral and cortical divide, and the terminal leaves being the morphological subtypes with different projection patterns. More morphology reconstructions of these neurons will be needed to consolidate the terminal leaves, which might become clearer discrete types or a continuum. In the future, it will be important to develop methods that allow full morphology reconstruction and gene expression profiling to be conducted in the same cell, and apply them to the study of single cells in both adult stage and during brain development, so that potential molecular correlates of morphological/connectional features can be identified. This and other approaches together will ultimately lead to an integrated understanding of the extraordinary cellular diversity in the brain that underlies brain function.

ACKNOWLEDGMENTS

We are grateful to the In Vivo Sciences, Molecular Biology, Histology, and Imaging teams at the Allen Institute for their technical support. This work was funded by the Allen Institute for Brain Science, NIH grants MH105982 and EY023173 to H.Z., NIH grant MH114830 to H.Z. and Q.L., NIH grants MH106008 and MH117079 to X.W.Y., NSFC grant 61721092 to Q.L., NSFC grant 61871411 to L.Q., and by the Tiny Blue Dot Foundation to C.K. The data conversion, neuron annotation and processing tools TeraVR and some of the analyses including brain registration were funded by the Southeast University – Allen Institute Joint Center and Southeast University. The Wenzhou Medical University reconstruction team received funding from the EPFL - Blue Brain Project. The authors wish to thank the Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support.

AUTHOR CONTRIBUTIONS

H.Z. conceptualized the study. M.B.V., T.L.D., B.T. and X.W.Y. generated the TIGRE-MORF (Ai166) mouse line. Z.J.H. provided Fezf2-CreER and Plxnd1-CreER mouse lines. K.E.H., R.L.

T.L.D., B.T. and J.A.H. contributed to the generation and characterization of specific transgenic mouse lines. H.G., A.L., S.Z., X.L., J.Y. and Q.L. conducted fMOST imaging. W.W., S.J., Y.Y. and C.H. handled the imaging data. L.Q., L.N. and H.P. developed methods for registration of fMOST datasets to CCF. Z.Z., S.J., Y.Y., Yimin W. and H.P. developed software tools for data conversion and morphology reconstruction. Yun W., X.K., Y.L., L.L., P.L., Y.S., L.Y., S.Z., A.F. and E.S. performed manual morphology reconstruction. Yun W., P.X., J.A.H. and H.P. performed manual or computational classification of morphological types. Z.Z., S.K. and S.A.S. assisted with morphological analysis. K.E.H., Q.X. and J.A.H. conducted anterograde AAV tracing. T.N.N. performed retrograde tracing. Z.Y., T.N.N. and B.T. conducted scRNA-seq data generation and analysis. S.M. and S.M.S. provided project management. L.E., M.J.H., B.T., L.N., S.A.S., J.A.H., H.G., Q.L., H.P., H.Z. and C.K. provided scientific management. H.Z. wrote the manuscript in consultation with all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Aransay, A., Rodriguez-Lopez, C., Garcia-Amado, M., Clasca, F., and Prensa, L. (2015). Long-range projection neurons of the mouse ventral tegmental area: a single-cell axon tracing analysis. Frontiers in neuroanatomy 9, 59.

Atlan, G., Terem, A., Peretz-Rivlin, N., Sehrawat, K., Gonzales, B.J., Pozner, G., Tasaka, G.I., Goll, Y., Refaeli, R., Zviran, O., et al. (2018). The Claustrum Supports Resilience to Distraction. Curr Biol 28, 2752-2762 e2757.

Bria, A., Iannello, G., Onofri, L., and Peng, H. (2016). TeraFly: real-time three-dimensional visualization and annotation of terabytes of multidimensional volumetric images. Nature methods 13, 192-194.

Cadwell, C.R., Palasantza, A., Jiang, X., Berens, P., Deng, Q., Yilmaz, M., Reimer, J., Shen, S., Bethge, M., Tolias, K.F., et al. (2016). Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. Nature biotechnology 34, 199-203.

Chan, K.Y., Jang, M.J., Yoo, B.B., Greenbaum, A., Ravi, N., Wu, W.L., Sanchez-Guardado, L., Lois, C., Mazmanian, S.K., Deverman, B.E., et al. (2017). Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. Nature neuroscience 20, 1172-1179.

Chatterjee, S., Sullivan, H.A., MacLennan, B.J., Xu, R., Hou, Y., Lavin, T.K., Lea, N.E., Michalski, J.E., Babcock, K.R., Dietrich, S., et al. (2018). Nontoxic, double-deletion-mutant rabies viral vectors for retrograde targeting of projection neurons. Nature neuroscience 21, 638-646.

Crick, F.C., and Koch, C. (2005). What is the function of the claustrum? Philosophical transactions of the Royal Society of London Series B, Biological sciences 360, 1271-1279.

Daigle, T.L., Madisen, L., Hage, T.A., Valley, M.T., Knoblich, U., Larsen, R.S., Takeno, M.M., Huang, L., Gu, H., Larsen, R., et al. (2018). A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. Cell 174, 465-480 e422.

Economo, M.N., Clack, N.G., Lavis, L.D., Gerfen, C.R., Svoboda, K., Myers, E.W., and Chandrashekar, J. (2016). A platform for brain-wide imaging and reconstruction of individual neurons. eLife 5, e10566.

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, L.T., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79-84.

Fuzik, J., Zeisel, A., Mate, Z., Calvigioni, D., Yanagawa, Y., Szabo, G., Linnarsson, S., and Harkany, T. (2016). Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. Nature biotechnology 34, 175-183.

Gamanut, R., Kennedy, H., Toroczkai, Z., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., and Burkhalter, A. (2018). The Mouse Cortical Connectome, Characterized by an Ultra-Dense Cortical Graph, Maintains Specificity by Distinct Connectivity Profiles. Neuron 97, 698-715 e610.

Gang, Y., Zhou, H., Jia, Y., Liu, L., Liu, X., Rao, G., Li, L., Wang, X., Lv, X., Xiong, H., et al. (2017). Embedding and Chemical Reactivation of Green Fluorescent Protein in the Whole Mouse Brain for Optical Micro-Imaging. Frontiers in neuroscience 11, 121.

Ghosh, S., Larson, S.D., Hefzi, H., Marnoy, Z., Cutforth, T., Dokka, K., and Baldwin, K.K. (2011). Sensory maps in the olfactory cortex defined by long-range viral tracing of single neurons. Nature 472, 217-220.

Goll, Y., Atlan, G., and Citri, A. (2015). Attention: the claustrum. Trends in neurosciences 38, 486-495.

Gong, H., Xu, D., Yuan, J., Li, X., Guo, C., Peng, J., Li, Y., Schwarz, L.A., Li, A., Hu, B., et al. (2016). High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. Nature communications 7, 12142.

Guo, Z.V., Inagaki, H.K., Daie, K., Druckmann, S., Gerfen, C.R., and Svoboda, K. (2017). Maintenance of persistent activity in a frontal thalamocortical loop. Nature 545, 181-186.

Han, Y., Kebschull, J.M., Campbell, R.A.A., Cowan, D., Imhof, F., Zador, A.M., and Mrsic-Flogel, T.D. (2018). The logic of single-cell projections from visual cortex. Nature 556, 51-56.

Harris, K.D., and Shepherd, G.M. (2015). The neocortical circuit: themes and variations. Nature neuroscience 18, 170-181.

Hnasko, T.S., Perez, F.A., Scouras, A.D., Stoll, E.A., Gale, S.D., Luquet, S., Phillips, P.E., Kremer, E.J., and Palmiter, R.D. (2006). Cre recombinase-mediated restoration of nigrostriatal dopamine in dopamine-deficient mice reverses hypophagia and bradykinesia. Proceedings of the National Academy of Sciences of the United States of America 103, 8858-8863.

Jackson, J., Karnani, M.M., Zemelman, B.V., Burdakov, D., and Lee, A.K. (2018). Inhibitory Control of Prefrontal Cortex by the Claustrum. Neuron 99, 1029-1039 e1024.

Jiang, X., Shen, S., Cadwell, C.R., Berens, P., Sinz, F., Ecker, A.S., Patel, S., and Tolias, A.S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. Science (New York, NY 350, aac9462.

Kebschull, J.M., Garcia da Silva, P., Reid, A.P., Peikon, I.D., Albeanu, D.F., and Zador, A.M. (2016). High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. Neuron 91, 975-987.

Kita, T., and Kita, H. (2012). The subthalamic nucleus is one of multiple innervation sites for long-range corticofugal axons: a single-axon tracing study in the rat. J Neurosci 32, 5990-5999.

Li, A., Gong, H., Zhang, B., Wang, Q., Yan, C., Wu, J., Liu, Q., Zeng, S., and Luo, Q. (2010). Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain. Science (New York, NY 330, 1404-1408.

Li, L., Ouellette, B., Stoy, W.A., Garren, E.J., Daigle, T.L., Forest, C.R., Koch, C., and Zeng, H. (2017). A robot for high yield electrophysiology and morphology of single neurons in vivo. Nature communications 8, 15604.

Lin, H.M., Kuang, J.X., Sun, P., Li, N., Lv, X., and Zhang, Y.H. (2018a). Reconstruction of Intratelencephalic Neurons in the Mouse Secondary Motor Cortex Reveals the Diverse Projection Patterns of Single Neurons. Frontiers in neuroanatomy 12, 86.

Lin, R., Wang, R., Yuan, J., Feng, Q., Zhou, Y., Zeng, S., Ren, M., Jiang, S., Ni, H., Zhou, C., et al. (2018b). Cell-type-specific and projection-specific brain-wide reconstruction of single neurons. Nature methods 15, 1033-1036.

Markram, H., Muller, E., Ramaswamy, S., Reimann, M.W., Abdellah, M., Sanchez, C.A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., et al. (2015). Reconstruction and Simulation of Neocortical Microcircuitry. Cell 163, 456-492.

Martersteck, E.M., Hirokawa, K.E., Evarts, M., Bernard, A., Duan, X., Li, Y., Ng, L., Oh, S.W., Ouellette, B., Royall, J.J., et al. (2017). Diverse Central Projection Patterns of Retinal Ganglion Cells. Cell reports 18, 2058-2072.

Mathur, B.N. (2014). The claustrum in review. Frontiers in systems neuroscience 8, 48.

Minciacchi, D., Molinari, M., Bentivoglio, M., and Macchi, G. (1985). The organization of the ipsi- and contralateral claustrocortical system in rat with notes on the bilateral claustrocortical projections in cat. Neuroscience 16, 557-576.

Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al. (2014). A mesoscale connectome of the mouse brain. Nature 508, 207-214.

Peng, H., Chung, P., Long, F., Qu, L., Jenett, A., Seeds, A.M., Myers, E.W., and Simpson, J.H. (2011). BrainAligner: 3D registration atlases of Drosophila brains. Nature methods 8, 493-500.

Peng, H., Ruan, Z., Long, F., Simpson, J.H., and Myers, E.W. (2010). V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. Nature biotechnology 28, 348-353.

Ragan, T., Kadiri, L.R., Venkataraju, K.U., Bahlmann, K., Sutin, J., Taranda, J., Arganda-Carreras, I., Kim, Y., Seung, H.S., and Osten, P. (2012). Serial two-photon tomography for automated ex vivo mouse brain imaging. Nature methods 9, 255-258.

Ramón y Cajal, S. (1909). Histologie Du Système Nerveux de L'homme & Des Vertébrés. (Paris: Maloine [Translated by N. Swanson and L.W. Swanson, Oxford University Press, 1995]).

Remedios, R., Logothetis, N.K., and Kayser, C. (2010). Unimodal responses prevail within the multisensory claustrum. J Neurosci 30, 12902-12907.

Remedios, R., Logothetis, N.K., and Kayser, C. (2014). A role of the claustrum in auditory scene analysis by reflecting sensory change. Frontiers in systems neuroscience 8, 44.

Saunders, A., Macosko, E.Z., Wysoker, A., Goldman, M., Krienen, F.M., de Rivera, H., Bien, E., Baum, M., Bortolin, L., Wang, S., et al. (2018). Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. Cell 174, 1015-1030 e1016.

Scorcioni, R., Polavaram, S., and Ascoli, G.A. (2008). L-Measure: a web-accessible tool for the analysis, comparison and search of digital reconstructions of neuronal morphologies. Nature protocols 3, 866-876.

Sherman, S.M. (2016). Thalamus plays a central role in ongoing cortical functioning. Nature neuroscience 19, 533-541.

Smith, J.B., Alloway, K.D., Hof, P.R., Orman, R., Reser, D.H., Watakabe, A., and Watson, G.D.R. (2018). The relationship between the claustrum and endopiriform nucleus: A perspective towards consensus on cross-species homology. The Journal of comparative neurology.

Smythies, J., Edelstein, L., and Ramachandran, V. (2012). Hypotheses relating to the function of the claustrum. Frontiers in integrative neuroscience 6, 53.

Smythies, J., Edelstein, L., and Ramachandran, V. (2014). Hypotheses relating to the function of the claustrum II: does the claustrum use frequency codes? Frontiers in integrative neuroscience 8, 7.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nature neuroscience 19, 335-346.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature 563, 72-78.

Tervo, D.G., Hwang, B.Y., Viswanathan, S., Gaj, T., Lavzin, M., Ritola, K.D., Lindo, S., Michael, S., Kuleshova, E., Ojala, D., et al. (2016). A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. Neuron 92, 372-382.

Wang, Q., Ng, L., Harris, J.A., Feng, D., Li, Y., Royall, J.J., Oh, S.W., Bernard, A., Sunkin, S.M., Koch, C., et al. (2017). Organization of the connections between claustrum and cortex in the mouse. The Journal of comparative neurology 525, 1317-1346.

Wang, Y., Ye, M., Kuang, X., Li, Y., and Hu, S. (2018). A simplified morphological classification scheme for pyramidal cells in six layers of primary somatosensory cortex of juvenile rats. IBRO Rep 5, 74-90.

White, M.G., and Mathur, B.N. (2018). Claustrum circuit components for top-down input processing and cortical broadcast. Brain structure & function 223, 3945-3958.

Xiong, H., Zhou, Z., Zhu, M., Lv, X., Li, A., Li, S., Li, L., Yang, T., Wang, S., Yang, Z., et al. (2014). Chemical reactivation of quenched fluorescent protein molecules enables resin-embedded fluorescence microimaging. Nature communications 5, 3992.

Yardeni, T., Eckhaus, M., Morris, H.D., Huizing, M., and Hoogstraten-Miller, S. (2011). Retro-orbital injections in mice. Lab Anim (NY) 40, 155-160.

Zeisel, A., Hochgerner, H., Lonnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Haring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular Architecture of the Mouse Nervous System. Cell 174, 999-1014 e1022.

Zeisel, A., Munoz-Manchado, A.B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science (New York, NY 347, 1138-1142.

Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. Nature reviews 18, 530-546.

Zhang, Z.W., and Deschenes, M. (1997). Intracortical axonal projections of lamina VI cells of the primary somatosensory cortex in the rat: a single-cell labeling study. J Neurosci 17, 6365-6379.

Zingg, B., Dong, H.W., Tao, H.W., and Zhang, L.I. (2018). Input-output organization of the mouse claustrum. The Journal of comparative neurology 526, 2428-2443.

Zingg, B., Hintiryan, H., Gou, L., Song, M.Y., Bay, M., Bienkowski, M.S., Foster, N.N., Yamashita, S., Bowman, I., Toga, A.W., et al. (2014). Neural networks of the mouse neocortex. Cell 156, 1096-1111.

Figure Legends

Figure 1. Sparse, robust and consistent neuronal labeling by combining TIGRE2.0 transgenic reporter lines with sparse Cre delivery.

(**A**) Schematic diagram showing the combination of CreERT2 transgenic driver line or Cre-expressing AAV with the GFP-expressing TIGRE2.0 reporter line Ai139 or Ai140. Very low dose tamoxifen induction of CreERT2 or very low-titer AAV-Cre delivery results in activation of the reporter in a spatially sparse manner. Transgenic reporter expression of GFP is robust and consistent across different cells. An optional addition is to cross in the GFP-expressing TIGRE1.0 reporter line Ai82, so that the tTA2 from Ai139 or Ai140 will activate the expression of GFP from two alleles – Ai139/Ai140 and Ai82, further increasing the level of GFP within Cre+ cells.

(**B**) Schematic diagram showing the combination of Cre or CreERT2 transgenic driver line or Cre-expressing AAV with the GFP-expressing sparse reporter line TIGRE-MORF/Ai166. Due to the intrinsic sparse expression of MORF (G_{22}-GFPf), some conventional Cre lines, moderate doses of tamoxifen induction of CreERT2, or moderate titers of AAV-Cre delivery can result in sparse labeling.

(**C-L**) Representative TissueCyte images showing sparse and strong labeling of various types of neurons using the above approach. (C) A Synapsin I promoter-driven CreERT2-expressing AAV serotyped with PHP.eB was delivered at a dilution of 1:1000 by retroorbital injection into an Ai139 mouse, followed by a 1-day tamoxifen induction one week post injection, resulting in random sparse labeling of neurons throughout the brain. (D) In a Tnnt1-IRES2-CreERT2;Ai140 brain, low-dose tamoxifen induction results in sparse labeling of thalamic projection neurons (left panel) with their axon terminal clusters in cortex clearly visible (right panel). (E) In a Gnb4-IRES2-CreERT2;Ai140;Ai82 brain, low-dose tamoxifen induction results in sparse labeling of *Gnb4*+ claustral and cortical neurons with their widely dispersed axon fibers clearly visible. (F) Cortical L6b neurons in a Ctgf-T2A-dgCre;Ai166 brain. (G) Cortical L2/3/4 neurons in a Cux2-CreERT2;Ai166 brain. (H) Cortical L5 PT neurons in a Fezf2-CreER;Ai166 brain. (I) Serotonergic neurons in dorsal raphe (DR) in a Slc6a4-CreERT2_EZ13;Ai166 brain. (J) Interneurons in cortex and cerebellum in a Pvalb-T2A-CreERT2;Ai166 brain. (K) *Th*+ cortical interneurons in a Th-Cre;Ai166 brain. (L) Projection neurons in LGd and other thalamic nuclei in a Vipr2-IRES2-Cre-neo;Ai166 brain. Third panel, axon projections from LGd neurons are seen in primary visual cortex (VISp). Fourth panel, axon projections likely from retinal ganglion cells are seen in superior colliculus (SC). Tamoxifen doses for CreERT2-containing mice are shown in **Table S1**.

Figure 2. Sparse, robust and consistent visualization of the dendritic and axonal arborizations of a wide range of neuronal types by fMOST imaging.

Images shown are 100-µm maximum intensity projection (MIP) images (*i.e.*, projected from 100 consecutive 1-µm image planes). Arrowheads indicate observed terminal boutons at the end of the axon segments. Tamoxifen doses are shown in **Table S1**.

(**A**) Cortical L2/3 IT neurons and their extensive local axon collaterals clearly labeled in a Cux2-CreERT2;Ai166 brain.

(**B**) Cortical L5 PT neurons and their sparse local axon collaterals seen in a Fezf2-CreER;Ai166 brain.

(**C**) Cortical L5 IT neurons and their local axon collaterals seen in a Plxnd1-CreER;Ai166 brain. Striatal medium spiny neurons (STR MSN) are also sparsely labeled, and their individual axons are clearly seen in substantia nigra (SN).

(**D**) *Gnb4*+ claustral (CLA) and cortical (L6PC) neurons with their widely dispersed axon fibers seen in a Gnb4-IRES2-CreERT2;Ai140;Ai82 brain.

(**E**) Noradrenergic neurons labeled in the locus ceruleus (LC), and their long-range axon fibers seen in cortex (CTX) and hypothalamus (HY) in a Dbh-Cre_KH212;Ai166 brain.

(**F**) Serotonergic neurons labeled in the dorsal raphe (DR), and their long-range axon fibers seen in hippocampus (HIP) and cortex (CTX) in a Slc6a4-CreERT2_EZ13;Ai166 brain.

(**G**) Cortical inhibitory basket cells (BC) and translaminar basket cells (t-BC), as well as L5 PT excitatory neurons, seen in a Pvalb-T2A-CreERT2;Ai166 brain. The L5 PT neurons form driving-type axon clusters with large boutons in the thalamus (TH).

(**H**) Thalamic projection neurons (TH PN) with their dense axon terminal clusters in cortex seen in a Tnnt1-IRES2-CreERT2;Ai82;Ai140 brain. Some STR MSNs are also labeled and they form intense axon clusters in SN.

(**I**) In a Vipr2-IRES2-Cre-neo;Ai166 brain, axon clusters from projection neurons in visual thalamic nuclei are seen in CTX, axon clusters likely from retinal ganglion cells are seen in superior colliculus (SC), and a cortical chandelier cell (ChC) is also fully labeled with its characteristic axonal branches.

Figure 3. The workflow of neuron visualization, reconstruction, mapping to Common Coordinate Framework (CCF) and analysis.

A complete fMOST image dataset is first converted to TereFly file format by TeraConverter, the data formatting tool in TeraFly (Bria et al., 2016). Then annotators work in the TeraVR annotation system (Wang et al., bioRxiv preprint https://doi.org/10.1101/621011, 2018) to reconstruct the full morphology of each neuron. In parallel, the whole brain image dataset is registered to CCF using BrainAligner (Peng et al., 2011), first using RLM, the Reliable-Landmark-Matching module of BrainAligner, then LQM, the Little-Quick-Warp module of BrainAligner. Following registration of the image dataset, all the reconstructed morphologies from the same brain are also registered for subsequent visualization and quantitative analysis.

Figure 4. Anterograde projection mapping from *Gnb4*+ neurons in claustrum or lateral cortex.

(**A-E**) AAV2/1-pCAG-FLEX-GFP tracer was injected into the claustrum (A-B), SSs (C-D) or SSp (E) in Gnb4-IRES2-Cre or Gnb4-IRES2-CreERT2 mice. Brains were imaged by the TissueCyte STPT system. First panel in each row shows the top-down view of segmented GFP-labeled axon projections in the cortex. Second panel shows the injection site. Third panel shows the fine axon fibers in respective target cortical areas. Fourth panel shows the segmented image of the third panel to visualize and quantify the axon fibers. Full STPT image datasets are available at the Allen Mouse Brain Connectivity Atlas web portal (http://connectivity.brain-map.org/) with the following experiment IDs: A, 514505957; B, 485902743; C, 553446684; D, 581327676; E, 656688345.

Figure 5. Soma locations and quantitative comparison of dendritic and axonal morphology features of *Gnb4*+ claustral and cortical neurons.

(**A**) Soma locations of all reconstructed neurons in the four Gnb4-IRES2-CreERT2 brains. Soma locations are registered and flipped to one hemisphere, shown in three views and grouped by the brain ID's. Color scheme: L6PC in blue, CLA in orange. Same below.

(**B**) Dendritic features that distinguish CLA and L6PC neurons. P-values (shown underneath each feature label) for all violin plots were calculated by Mann-Whitney tests (same below). Top panel: distribution of dendritic morphological features grouped by neuron types. Dendrites were centered by the soma and rotated by PCA (principle component analysis) so the width/height/depth match the longest to shortest dimensions. Lower left panel: 2D display of dendrites ordered by the depth. Lower right panel: Gaussian kernel density distribution of dendrites combined by neuron types. Red arrows indicate the 5th and 95th percentile of each direction.

(**C**) Axonal features that distinguish CLA and L6PC neurons. Left panels: differential global axon morphological features grouped by neurons types. Width/height/depth match the anterior-posterior/top-bottom/left-right directions of the mouse brain. Middle panels: Gaussian kernel density distribution of axon skeletons combined by neuron types. Right panels: differential local axon morphological features grouped by neuron types. Bottom: 2D display of local axons ordered by the width.

Figure 6. Axonal morphologies and manual classification of *Gnb4*+ claustral and cortical neurons.

(**A**) CLA neurons (n = 34) from 4 Gnb4 brains were manually classified into four morphological types, each shown in top-down, coronal and sagittal views with all neurons for each type (their somata and dendrites are labeled in blue and their axons labeled in different colors). CLA cells of each type were merged together according to their soma locations after registration to CCF. CLA_I-ipsi cells (n = 19) are characterized by their axonal projections within the ipsilateral hemisphere and predominantly along the midline cortical regions, which have somata distributed at the full length of the claustrum. CLA_I-bi cells (n = 9) are characterized by the bilateral axonal projections predominantly along the midline cortical regions of the ipsilateral side. CLA_II-ipsi cells (n = 5) are characterized by their axonal projections within the ipsilateral hemisphere and virtually avoiding the midline cortical regions. CLA_II-bi cell (n = 1) is characterized by its axonal projections to both ipsi- and contralateral hemispheres and virtually avoiding the midline region on both sides.

(**B**) L6PCs (n = 62) from 4 Gnb4 brains were manually classified into two morphological types, shown here in the same way as CLA cells after registration to CCF. L6PC_ipsi cells (n = 36) are characterized by their axonal projections within the ipsilateral hemisphere; L6PC_bi cells (n = 26) have axonal projections to both hemispheres.

(**C**) CLA neurons have topographically distinct projection patterns based on the anterior-posterior positions of their somata. CLA_I-ipsi: the panels from left to right contain 5, 5 and 9 cells, respectively. CLA_I-bi: the panels from left to right contain 3 and 6 cells, respectively.

(**D**) L6PC neurons have topographically distinct projection patterns based on the anterior-posterior positions of their somata. L6PC_ipsi: the panels from left to right contain 10, 6 and 20 cells, respectively. L6PC_bi: the panels from left to right contain 16, 2 and 8 cells, respectively.

Figure 7. Computational clustering-based classification of *Gnb4*+ claustral and cortical neurons.

(**A**) Integrated co-clustering matrix by averaging the co-clustering matrices of four feature sets: projection pattern, soma location, axon morphology and dendrite morphology. Side bars indicate manually assigned types with color codes shown below the matrix.

(**B**) Dendrogram based on the co-clustering matrix. Threshold for cluster calls is shown as the dashed line. Each cluster is annotated by the brain regions where somata (black) and axon clusters (red) reside. Regions were selected to represent >50% of cluster members.

(**C**) Left panel: 2D UMAP of co-clustering matrix. Each dot represents a neuron colored by manually assigned types and numbered by computational cluster ID's. Right panel: Sankey plot for the correspondence between manual and computational types. Width of connecting lines represents number of cells.

(**D**) Heatmap for representative target brain regions of each neuron. Columns represent single cells sorted by cluster assignments and rows represent target brain regions.

(**E**) Total number of cortical targets contacted by each neuron grouped by clusters. Ipsilateral and contralateral targets are counted separately. A minimum of 1 axon terminal (top panel) or 1,000 µm of axon length (bottom panel) is used as the threshold to label a region as "targeted".

(**F**) Top-down views of neurons in each cluster. Neurons are flipped all to the left hemisphere for comparison of axon projection patterns. Stars indicate soma locations and are flipped to the right hemisphere for visualization purpose.

Figure 8. Single-cell RNA-seq characterization of Car3 subclass of claustral and cortical neurons.

(**A**) Dendrogram of the co-clustering analysis of SMART-Seq v4 data from ~75,000 cortical, hippocampal and claustral cells reveals a distinct branch of Car3 subclass (shown in the dashed box, upper left panel), which contains nearly all the neurons isolated from claustrum from the Gnb4-IRES2-CreERT2;Ai140 mice (lower left panel). Middle panel: The Car3 subclass consists of one major cluster that contains cells from both claustrum and cortex, plus six additional minor clusters. Dot plot below the dendrogram shows the number of cells from each cortical region or claustrum contributing to each cluster. In the tSNE plots (right panels), the claustral and cortical neurons are intermingled together in the major cluster, whereas one of the minor clusters, Car3 TEa-PERI-ECT_GU Hgf, contains mostly cells from TEa.

(**B**) Re-clustering analysis of SMART-Seq v4 data from 1,997 Car3 subclass cells from (A) using a more relaxed clustering criterion. This includes 298 Retro-seq cells labeled by retrograde tracers injected into ACA, VISpl, ORBl or MOp and isolated from claustrum or various cortical areas (*i.e.*, SSs, VISp, AI and TEa) on both ipsilateral and contralateral sides of the tracer injection. Dot plot below the new dendrogram of 19 clusters shows the number of cells from each cortical area or claustrum contributing to each cluster. The tSNE plots (right panels) show distribution of cells color-coded by clusters or Cre lines.

(**C**) Correlation of transcriptomic clusters with regions where all 1,997 Car3 cells were isolated from (left), and retrograde tracer injection sites (middle) and hemispheres (right) the 298 Retro-seq cells came from using tSNE plots.

Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Hongkui Zeng (hongkuiz@alleninstitute.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal care and use

Both male and female transgenic mice ≥ P56 were utilized for all experiments. All animals were housed 3-5 per cage and maintained on a 12-hour light/dark cycle, in a humidity- and temperature-controlled room with water and food available *ad libitum*. All experimental procedures related to the use of mice were conducted with approved protocols in accordance with NIH guidelines, and were approved by the Institutional Animal Care and Use Committee (IACUC) of the Allen Institute for Brain Science.

METHOD DETAILS

For the acronyms and full names of all brain regions mentioned, see the CCFv3 ontology tab of Table S4.

Transgenic mice

All transgenic crosses are listed in **Table S1**. Data for systematic characterization of the expression pattern of each transgenic mouse line can be found in the AIBS Transgenic Characterization database (http://connectivity.brain-map.org/transgenic/search/basic).

Induction of CreERT2 driver lines was done by administration via oral gavage (PO) of tamoxifen (50 mg/ml in corn oil) at original (0.2 mg/g body weight) or reduced dose for one day in an adult mouse. The dosage for mice age P7-P15 is 0.04 ml. Mice can be used for experiments at 2 or more weeks after tamoxifen dosing. Specific dose of tamoxifen to induce sparse labeling in each CreERT2 driver line is shown in **Table S1**.

Brain-wide delivery of low-dose Cre-expressing AAV-PHP.eB virus was achieved by retro-orbital injection using a previously described technique (Yardeni et al., 2011). The PHP.eB variant of AAV can cross the blood brain barrier with a tropism towards CNS cells, allowing vascular delivery of the virus to extend throughout the brain (Chan et al., 2017). A viral load of approximately 1×10^8 particles was diluted into 50 µL sterile PBS and injected into the retro-orbital sinus of anesthetized mice using a 31G insulin syringe. In Cre driver virus experiments, mice were perfused 21 days following injection. In CreERT2 driver virus experiments, mice were perfused 21 days following a single dose of tamoxifen induction administered 5-7 days after AAV injection, either at the original or reduced dose as described in the CreERT2 driver line experiments above.

TissueCyte STPT imaging

Imaging by serial two-photon (STP) tomography (TissueCyte 1000, TissueVision Inc. Somerville, MA) has been described in earlier published studies (Martersteck et al., 2017; Oh et al., 2014; Ragan et al., 2012).

Mice were deeply anesthetized with 5% isoflurane and intracardially perfused with 10 ml of saline (0.9% NaCl) followed by 50 ml of freshly prepared 4% paraformaldehyde (PFA) at a flow rate of 9 ml/min. Brains were dissected and post-fixed in 4% PFA at room temperature for 3–6 h and then overnight at 4 °C. Brains were rinsed briefly with PBS and stored in PBS with 0.1% sodium azide until imaging.

Prior to imaging, the brain was embedded in a 4.5% oxidized (10 mM NaIO4) agarose solution in a grid-lined embedding mold to standardize its placement in an aligned coordinate space. The agarose block was then left at room temperature for 20 min to allow solidification. Covalent interactions between brain tissue and agarose were promoted by placing the solidified block in 0.5% sodium borohydride in 0.5 M sodium borate buffer (pH 9.0) overnight at 4 °C. The agarose block was then mounted on a 1 × 3 glass slide using Loctite 404 glue and prepared immediately for serial imaging.

Image acquisition was accomplished using TissueCyte 1000 systems (TissueVision, Cambridge, MA) coupled with Mai Tai HP DeepSee lasers (Spectra Physics, Santa Clara, CA). The mounted specimen was fixed through a magnet to the metal plate in the centre of the cutting bath filled with degassed, room-temperature PBS with 0.1% sodium azide. A new blade was used for each brain on the vibratome and aligned to be parallel to the leading edge of the specimen block. Brains were imaged from the caudal end. The specimen was illuminated with 925 nm wavelength light through a Zeiss 320 water immersion objective (NA = 1.0), with 250 mW light power at objective. The two-photon images for red, green and blue channels were taken at 75 µm below the cutting surface. To scan a full tissue section, individual tile images were acquired, and the entire stage was moved between each tile. After an entire section was imaged, the x and y stages moved the specimen to the vibratome, which cut a 100-µm section, and returned the specimen to the objective for imaging of the next plane. The blade vibrated at 60 Hz and the stage moved towards the blade at 0.5 mm per sec during cutting. Images from 140 sections were collected to cover the full range of mouse brain at an x-y resolution of 0.35 µm per pixel. Upon completion of imaging, sections were retrieved from the cutting bath and stored in PBS with 0.1% sodium azide at 4°C.

fMOST imaging

All tissue preparation has been described previously (Gang et al., 2017). Following fixation, each intact brain was rinsed three times (6 h for two washes and 12 h for the third wash) at 4°C in a 0.01 M PBS solution (Sigma-Aldrich Inc., St. Louis, US). Then the brain was subsequently dehydrated via immersion in a graded series of ethanol mixtures (50%, 70%, and 95% (vol/vol) ethanol solutions in distilled water) and the absolute ethanol solution three times for 2 h each at 4°C. After dehydration, the whole brain was impregnated with Lowicryl HM20 Resin Kits (Electron Microscopy Sciences, cat.no. 14340) by sequential immersions in 50, 75, 100 and 100% embedding medium in ethanol, 2 h each for the first three solutions and 72 h for the final solution. Finally, each whole brain was embedded in a gelatin capsule that had been filled with

HM20 and polymerized at 50°C for 24 h.

The whole brain imaging is realized using a fluorescence microscopic optical sectioning tomography (fMOST) system. The basic structure of the imaging system is the combination of a wide-field upright epi-fluorescence microscopy with a mechanic sectioning system. This system runs in a wide-field block-face mode but updated with a new principle to get better image contrast and speed and thus enables high throughput imaging of the fluorescence protein labeled sample (manuscript in preparation). Each time we do a block-face fluorescence imaging across the whole coronal plane (X-Y axes), then remove the top layer (Z axis) by a diamond knife, and then expose next layer, and image again. The thickness of each layer is 1.0 micron. In each layer imaging, we used a strip scanning (X axis) model combined with a montage in Y axis to cover the whole coronal plane (Li et al., 2010). The fluorescence, collected using a microscope objective, passes a bandpass filter and is recorded with a TDI-CCD camera. We repeat these procedures across the whole sample volume to get the required dataset.

The objective used is 40X WI with numerical aperture (NA) 0.8 to provide a designed optical resolution (at 520 nm) of 0.35 μm in XY axes. The imaging gives a sample voxel of 0.35 x 0.35 x 1.0 μm to provide proper resolution to trace the neural process. The voxel size can be varied upon difference objective. Other imaging parameters for GFP imaging include an excitation wavelength of 488 nm, and emission filter with passing band 510-550 nm.

Full neuronal morphology reconstruction system

We developed Vaa3D, an open-source, cross-platform visualization and analysis system, for the tasks of reconstructing massive neuronal morphologies. To efficiently and effectively deal with the whole-mouse brain imaging data, we incorporated several enabling modules into Vaa3D, such as TeraFly and TeraVR. TeraFly supports visualization and annotation of multidimensional imaging data with virtually unlimited scales. The user can flexibly choose to work at a specific region of interest (ROI) with desired level of detail (LoD). The out-of-core data management of TeraFly allows the software to smoothly deal with terabyte-scale of data even on a portable workstation with normal RAM size. Driven by virtual reality (VR) technologies, TeraVR is an annotation tool for immersive neuron reconstruction that has been proved to be critical for achieving precision and efficiency in morphology data production. It creates stereo visualization for image volumes and reconstructions and offers an intuitive interface for the user to interact with such data. TeraVR excels at handling various challenging yet constantly encountered data situations during whole-brain reconstruction, such as the noisy, complicated, and weakly labeled signals. The complete morphologies generated by using TeraVR are essential for a number of downstream applications including cell profiling and categorization. Both TeraFly and TeraVR are seamlessly integrated in Vaa3D and can be used combinedly and flexibly.

Full morphology reconstruction

At the beginning stage, we used Neurolucida 360 (NL360, MBF Bioscience, USA) to trace neurons. Since NL360 had a limitation to handle whole brain image data set, a portion of the imaging data (200-300 images corresponding to a 200-300 μm coronal brain section) was loaded at a time to reconstruct neurons. In order to trace a single neuron such as a CLA cell at whole brain level, a serial image stacks (>30 coronal brain sections) were made with a software called

Biolucida Converter (MBF Bioscience, USA). Later, TeraVR was developed and integrated with Vaa3D, which not only allowed us to handle whole brain image data set at once, but also significantly improved the precision and efficiency in morphology reconstruction of single neurons at whole brain level. For instance, the 3D relationship among intertwined axonal segments could be visualized at various angles in the virtual space at a resolution not reached before. Thus, many missing branches were recovered and mis-connected segments were corrected for those cells reconstructed with NL360. At the later stage, neurons were traced using Vaa3D-TeraVR directly.

The *Gnb4*+ claustral and L6PC neurons are particularly challenging to reconstruct, as each neuron's axon projection covers a wide area and axons from different neurons are often intermingled together. We often encountered two axon segments crossing or touching each other. In most cases, it was not difficult to identify one to continue the tracing segment by comparing the similarity between the parent and extending segments in their structural features including thickness and tortuosity, bouton shape and size and density, and labeling condition, etc. In some cases, however, two or multiple similar-looking segments were intertwined or bundled together so closely that could not be separated even using Tera-VR. For these special cases, an extending segment and its branches were kept if it was continued reaching all ends, or given up if it was eventually connected to a main axonal stem that lead to another soma. A neuron was considered fully reconstructed in the case that it was completed with all ends that typically had very well labeled, enlarged boutons. Those neurons that were labeled too faintly to be confidently traced were abandoned and excluded from this study. Finally, a QC-checking procedure was performed by an experienced annotator using TeraVR by double checking the entire reconstruction of a neuron, at a high magnification paying special attention to the proximal axonal part or a main axonal trunk of an axon cluster where axonal collaterals often emerged and branches were more frequently missed due to the local image environment being composed of crowded high contrasting structures. Auto-refinement fits the tracing to the center of fluorescent signals as the last step. The final reconstruction is a single tree without breaks, loops, multiple branches from a single point, etc.

Registration to CCF

We performed 3D registration from fMOST images (subject) to average mouse brain template of CCFv3 (target). 1) fMOST images were first down-sampled by 64x64x16 (X, Y, Z) to roughly match the size of the target brain. 2) 2D stripe-removal was then performed using frequency notch filters. 3) A dozen or so matching landmark pairs between subject and target were manually added to ensure correct affine transformation that approximately aligned the orientation and scales. 4) Affine transformation was applied to minimize the sum of squared difference (SSD) of intensity between target and subject images. 5) Intensity was normalized by matching the local average intensity of subject image to that of target image. 6) A candidate list of landmarks across CCF space was generated by grid search (grid size=16 pixels). Then the BrainAligner software searched corresponding landmarks in the subject image and performed local alignment.

Quantitative morphology data analysis

Pre-processing of SWC files:

SWC files were processed and examined with Vaa3D plugins to ensure topological correctness: sorted single tree with root node as soma. Terminal branches < 10 microns were pruned to remove artifacts. SWC files were resampled with a step size of 5 microns.

Quantification of axon projection patterns:

To analyze the distribution and amount of axon in brain-wide targets following registration to the CCFv3, we used a manually curated set of 316 non-overlapping structures at a mid-ontology level that are most closely matched in size or division (named "summary structures", second tab in **Table S4**). Ipsi- and contra-lateral sides of brain regions were calculated separately. Brain regions with non-zero levels in less than 3 neurons were excluded from further analysis.

Morphological features:

Axonal and dendritic morphological features, defined according to L-measurement (Scorcioni et al., 2008), were calculated using Vaa3D plugin "global_neuron_feature". Selected features include:

(Axon global)

'Overall Width', 'Overall Height', 'Overall Depth', 'Total Length', Euclidean Distance', 'Max Path Distance', 'Number of Branches'.

(Axon local)

'Total Length', 'Number of Branches'.

(Dendrite)

'Overall Width', 'Overall Height', 'Overall Depth', 'Total Length', 'Max Euclidean Distance', 'Max Path Distance', 'Number of Branches', 'Max Branch Order'.

Local axons were defined as axon arbors within 200 microns from the somata. Local axons and dendrites were rotated based on principle component analysis (PCA) so dimensions were aligned with the largest to smallest spans. Then shifting was performed to localize somata at the origin of coordinates.

Steps of clustering-based cell type classification

Data normalization: Morphological features were normalized by the mean and standard variation in a feature-wise manner. Projection pattern features were normalized by the total length per 100 µm in a sample-wise manner and scaled by logarithm. Soma locations were flipped to the same hemisphere.

Similarity metrics: For each feature set, we first calculated the Euclidean distance matrix. Then a ranked K-nearest neighbor (KNN) matrix was created. We then applied the Shared Nearest Neighbor (SNN) approach to measure the similarity between each pair of samples x_i and x_j. The SNN metric was defined as the maximum average rank among their common neighbors:

$$S(x_i, x_j) = \max_{v \in NN(x_i) \cap NN(x_j)} \{k - \frac{1}{2}[rank_{NN(x_i)}(v) + \frac{1}{2}(rank_{NN(x_j)}(v)]\}$$

Similarity scores were set as 0 for pairs with non-overlapping KNN sets and a weighted SNN graph is created.

Co-clustering analysis: Co-clustering matrix for each feature set was calculated by iterative random sampling. During each iteration, 95% of samples were randomly selected to create an SNN graph. We then applied the “Fast-greedy” community detection algorithm using python package “python-igraph” for clustering assignment. For each pair of samples, the co-clustering score was defined as the times of co-clustering normalized by the iterations of co-occurring. Resampling was performed 5,000 times to reach saturation. The overall co-clustering matrix is a weighted average of the four feature sets. Agglomerative clustering was performed to the co-clustering matrix to get clusters.

Outlier removal: Outliers were detected by comparing the Euclidean distance between a sample and the other samples with the same cluster identity. We used overall within-cluster distance as the background distribution. Samples with significantly higher (one-sided Mann-Whitney test) within-cluster distance were filtered out as outliers. Agglomerative clustering was performed for the remaining co-clustering matrix. This process iterated until no new outlier could be detected.

Characterization of cell types: For each feature set, we performed two-sided Mann-Whitney tests: claustrum vs. cortical neurons; each cluster vs. other clusters. P-values were adjusted by Bonferroni correction.

Anterograde tracing and retrograde labeling

For anterograde projection mapping, we injected AAV2/1-pCAG-FLEX-EGFP-WPRE-pA (Oh et al., 2014) into CLA, SSs or SSp of Gnb4-IRES2-Cre or Gnb4-IRES2-CreERT2 mice at P37-P65 respectively. Stereotaxic injection procedures were performed as previous described (Oh et al., 2014) and stereotaxic coordinates used for each experiment can be found in the data portal. For the Gnb4-IRES2-CreERT2 mice, tamoxifen induction was conducted 1 week post injection at full dose (0.2 mg/g body weight) for 5 consecutive days. Mice survived 3 weeks (or 4 weeks for the tamoxifen-induced mice) post injection, and brains were perfused and collected for TissueCyte imaging.

For retrograde labeling, we injected AAV2-retro-EF1a-dTomato (Tervo et al., 2016) into ORBl, ACA, MOp, or VISpl of Gnb4-IRES2-CreERT2 and Cux2-CreERT2 mice crossed to Ai140 or Ai148. We additionally injected RVΔGL-Cre (Chatterjee et al., 2018) or CAV2-Cre (Hnasko et al., 2006) into MOp, SSs, or VISpm of Ai14 mice. We FACS-sorted and collected RFP+ or RFP+/GFP+ cells from CLA and surrounding cortical regions. Stereotaxic injection procedures were performed as previous described (Oh et al., 2014). Mice were injected at P40 or older, with 16-31 days survival post-injection.

Single-cell RNA-sequencing

Cells from transgenic mice or transgenic mice injected with retrograde tracers were collected by microdissection of AI, CLA, ENTl, MOp, ORB, PL-ILA, SSp, SSs, TEa-PERI-ECT, and VISp. Single-cell suspensions were created and cells were collected using fluorescence activated cell sorting (FACS). FACS gates were selective for cells with fluorescent protein expression from transgenic and/or viral reporters.

Cells were then frozen at -80ºC, and were later processed for scRNA-seq using the SMART-Seq v4 method (Tasic et al., 2018). After sequencing, raw data was quantified using STAR v2.5.3

and were aligned to both a Ref-Seq transcriptome index for the mm10 genome, and a custom index consisting of transgene sequences. PCR duplicates were masked and removed using STAR option ‘bamRemoveDuplicates’. Only uniquely aligned reads were used for gene quantification. Gene read counts were quantified using the summarizeOverlaps function from R GenomicAlignments package using both intronic and exonic reads, and QC was performed as described in (Tasic et al., 2018).

Clustering was performed using house developed R package scrattch.hicat (available via github https://github.com/AllenInstitute/scrattch.hicat). In additional to classical single-cell clustering processing steps provided by other tools such as Seurat, this package features automatically iterative clustering by making finer and finer splits while ensuring all pairs of clusters, even at the finest level, are separable by fairly stringent differential gene expression criteria. The package also performs consensus clustering by repeating iterative clustering step on 80% subsampled set of cells 100 times, and derive the final clustering result based on cell-cell co-clustering probability matrix. This feature enables us to both fine tune clustering boundaries and to assess clustering uncertainty. One critical criterion that determines the clustering resolution is the minimal differential gene expression (DGE) requirement between all pairs of clusters. Using stringent DGE requirement results in fewer clusters with more prominent differences between clusters, while using more relaxed DGE expression result in more clusters captured by more subtle differences. For the whole cortical and hippocampal dataset with ~75,000 cells, we used the standard DGE requirement as in (Tasic et al., 2018). More specificially, q1.th = 0.5 (minimal fraction of cells in a given cluster that express the positive markers), q.diff.th=0.7 (normalized differences in fraction of cells expressing the positive markers between the foreground and background cluster, maximal value is 1), and de.score.th=150 (overall assessment of the statistical significance of all DGE genes). To capture more continuous differences between cell types with more sutble differences, for the isolated cells from the Car3 subclass, we used the relaxed DGE criterial: q1.th = 0.5, q.diff.th=0.5, de.score.th = 100.

Supplemental Information

Table S1. Transgenic mice used for the generation of fMOST imaging datasets, including main metadata information and tamoxifen dosing (see **Methods**) for sparse labeling.

Table S2. List of reconstructed CLA cells and L6PCs, with each neuron's soma location 3D coordinates after registration to CCF.

Table S3. Dendritic and axonal morphological features – quantitative values for each neuron.

Table S4. Single neuron projection matrix used for clustering analysis. In a separate tab, a complete list of all "Summary Structures" in CCFv3 ontology is shown, with the full name and acronym for each structure/region.

Table S5. The 1,997 Car3 IT subclass cells for scRNA-seq analysis, with relevant metadata including retrograde labeling informatoin.

Figure S1. Sparse neuronal labeling in additional Cre lines. MC, Martinotti cells. CA1 OLM, OLM neurons in hippocampal CA1.

Figure S2. CCF registration workflow. Pipeline of 3D registration from fMOST image (subject) to average mouse brain template of CCFv3 (target). Numbers below each panel indicate the pixel sizes in the order of X*Y*Z. See **Methods** for explanation of each step.

Figure S3. *Gnb4* expression in claustrum, endopiriform nucleus and cortex sampled through the entire anterior-posterior span (from ORB to TEa). *Gnb4* gene in situ hybridization (ISH) coronal image planes are shown side-by-side with the corresponding CCF reference plates.

Figure S4. Top-down, coronal and sagittal views of each full CLA neuronal morphology shown within the CCF 3D reference space (with CLA region highlighted). Color scheme: Soma in green, dendrite in magenta, and axon in black.

Figure S5. Top-down, coronal and sagittal views of each full L6PC neuronal morphology shown within the CCF 3D reference space (with CLA region highlighted). Color scheme: Soma in green, dendrite in magenta, and axon in black.

Figure S6. Differential features between neighboring cell clusters, grouped by feature sets: projection pattern, soma location, axon morphological features and dendritic morphological features. Dendrogram was based on co-clustering matrix. Statistical tests were performed at each split of the tree structure. Features with the lowest p-values were selected and labeled at the branch with higher levels. Only features with p-values < 0.1 were selected.

Figure S7. Pie charts showing the composition of each computationally derived cluster by manually annotated cell types, bi- or ipsilateral projection, brain id, soma locations and axon projection targets.

Movie S1. Whole-brain fMOST dataset (down-sampled) of a Cux2-CreERT2;Ai166 brain.

Movie S2. Whole-brain fMOST dataset (down-sampled) of a Plxnd1-CreER;Ai166 brain.

Movie S3. Whole-brain fMOST dataset (down-sampled) of a Fezf2-CreER;Ai166 brain.

Movie S4. Whole-brain fMOST dataset (down-sampled) of a Pvalb-T2A-CreERT2;Ai166 brain.

Movie S5. Whole-brain fMOST dataset (down-sampled) of a Gnb4-IRES2-CreERT2;Ai140;Ai82 brain.

Movie S6. Whole-brain fMOST dataset (down-sampled) of a Slc6a4-CreERT2;Ai166 brain.

Movie S7. Whole-brain fMOST dataset (down-sampled) of a Tnnt1-IRES2-CreERT2;Ai140;Ai82 brain.

Movie S8. Whole-brain fMOST dataset (down-sampled) of a Vipr2-IRES2-Cre-neo;Ai166 brain.

Movie S9. CCF-Registered image dataset of Gnb4-IRES2-CreERT2;Ai140;Ai82 brain #17109.

Movie S10. CCF-Registered image dataset of Gnb4-IRES2-CreERT2;Ai140;Ai82 brain #17781.

Movie S11. CCF-Registered image dataset of Gnb4-IRES2-CreERT2;Ai140;Ai82 brain #17782.

Movie S12. CCF-Registered image dataset of Gnb4-IRES2-CreERT2;Ai140;Ai82 brain #236174.

Figure 1

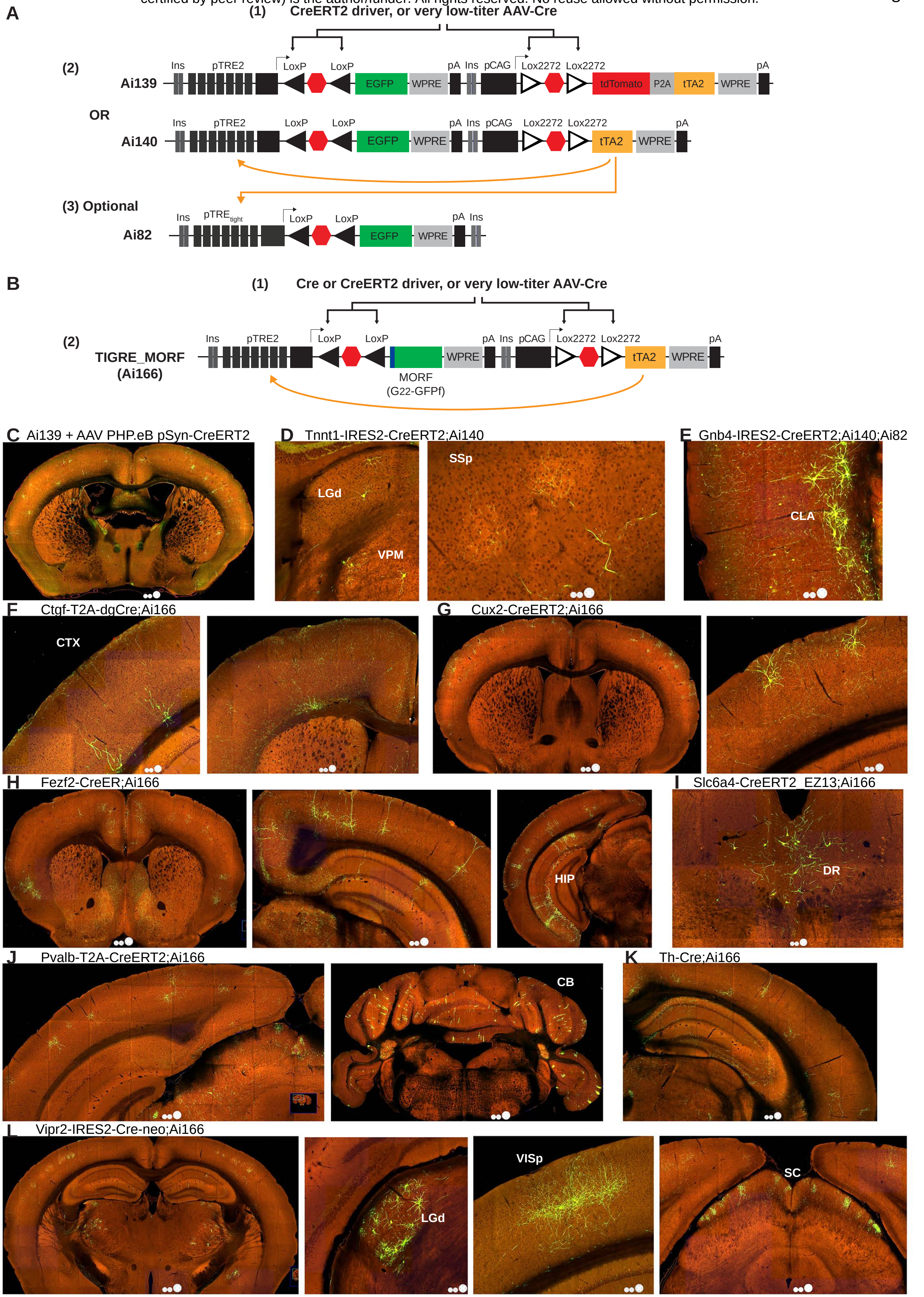

A
(1)
CreERT2 driver, or very low-titer AAV-Cre
(2)
Ai139
Ins
pTRE2
LoxP
LoxP
EGFP
WPRE
pA
Ins
pCAG
Lox2272
Lox2272
tdTomato
P2A
tTA2
WPRE
pA
OR
Ai140
Ins
pTRE2
LoxP
LoxP
EGFP
WPRE
pA
Ins
pCAG
Lox2272
Lox2272
tTA2
WPRE
pA
(3) Optional
Ai82
Ins
pTRE$_{tight}$
LoxP
LoxP
EGFP
WPRE
pA
Ins
B
(1)
Cre or CreERT2 driver, or very low-titer AAV-Cre
(2)
TIGRE_MORF (Ai166)
Ins
pTRE2
LoxP
LoxP
MORF (G22-GFPf)
WPRE
pA
Ins
pCAG
Lox2272
Lox2272
tTA2
WPRE
pA
C
Ai139 + AAV PHP.eB pSyn-CreERT2
D
Tnnt1-IRES2-CreERT2;Ai140
LGd
VPM
SSp
E
Gnb4-IRES2-CreERT2;Ai140;Ai82
CLA
F
Ctgf-T2A-dgCre;Ai166
CTX
G
Cux2-CreERT2;Ai166
H
Fezf2-CreER;Ai166
HIP
I
Slc6a4-CreERT2_EZ13;Ai166
DR
J
Pvalb-T2A-CreERT2;Ai166
CB
K
Th-Cre;Ai166
L
Vipr2-IRES2-Cre-neo;Ai166
LGd
VISp
SC

Figure 2

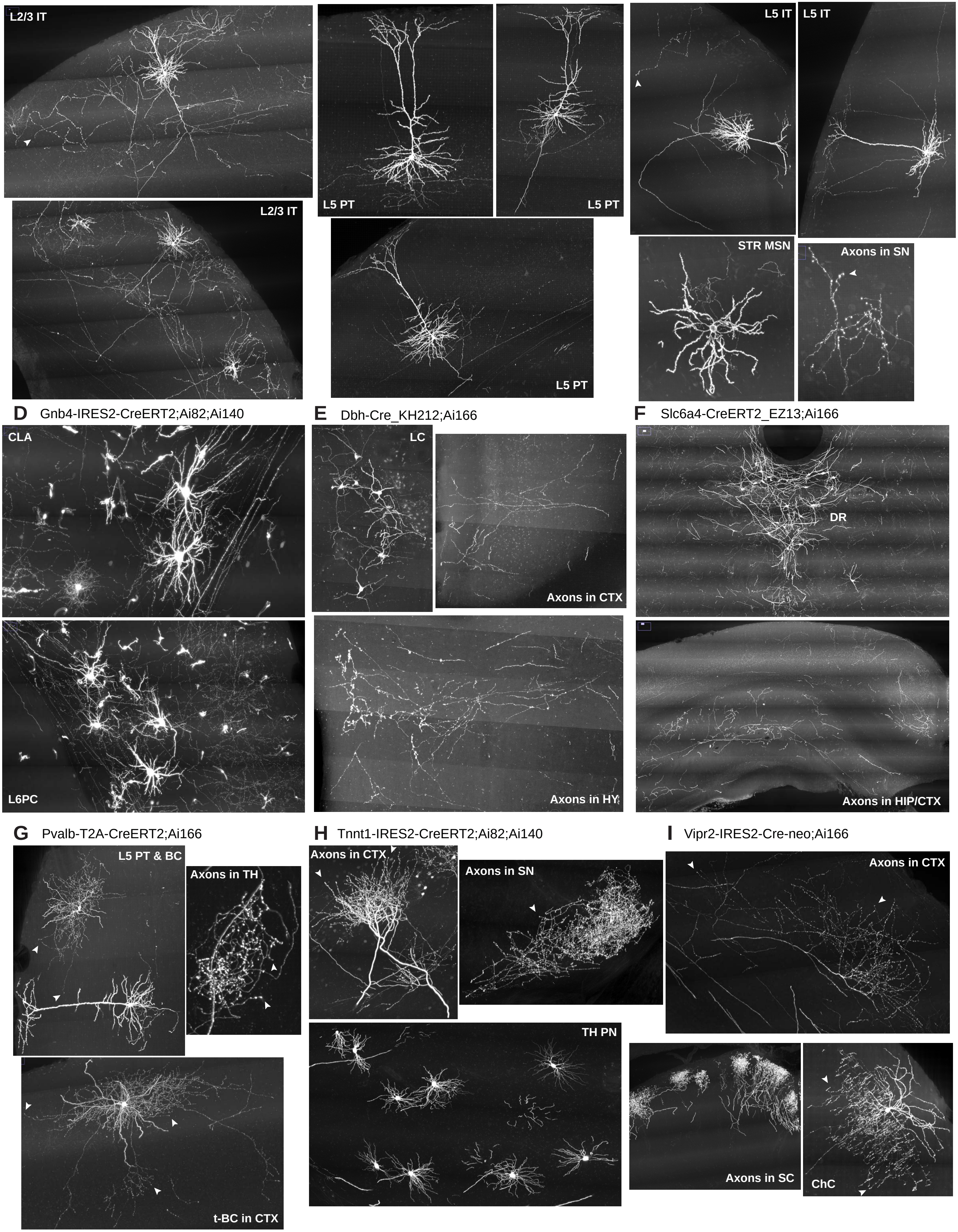
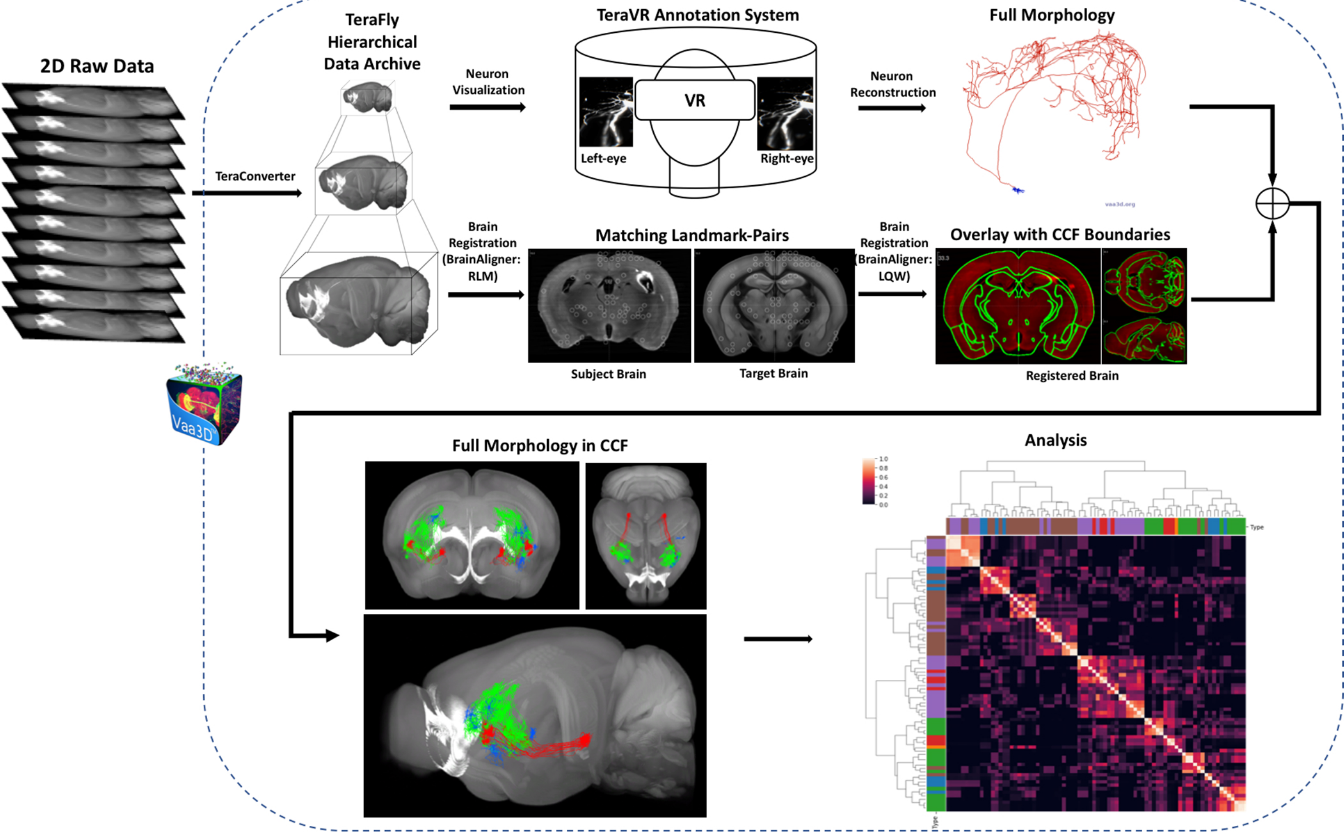

A Cux2-CreERT2;Ai166
L2/3 IT
L2/3 IT
B Fezf2-CreER;Ai166
L5 PT
L5 PT
L5 PT
C Plxnd1-CreER;Ai166
L5 IT
L5 IT
STR MSN
Axons in SN
D Gnb4-IRES2-CreERT2;Ai82;Ai140
CLA
L6PC
E Dbh-Cre_KH212;Ai166
LC
Axons in CTX
Axons in HY
F Slc6a4-CreERT2_EZ13;Ai166
DR
Axons in HIP/CTX
G Pvalb-T2A-CreERT2;Ai166
L5 PT & BC
Axons in TH
t-BC in CTX
H Tnnt1-IRES2-CreERT2;Ai82;Ai140
Axons in CTX
Axons in SN
TH PN
I Vipr2-IRES2-Cre-neo;Ai166
Axons in CTX
Axons in SC
ChC

Figure 3



2D Raw Data
TeraConverter
TeraFly Hierarchical Data Archive
Neuron Visualization
TeraVR Annotation System
VR
Left-eye
Right-eye
Neuron Reconstruction
Full Morphology
vaa3d.org
Brain Registration (BrainAligner: RLM)
Matching Landmark-Pairs
Subject Brain
Target Brain
Brain Registration (BrainAligner: LQW)
Overlay with CCF Boundaries
33.3
Registered Brain
Vaa3D
Full Morphology in CCF
Analysis
1.0
0.8
0.6
0.4
0.2
0.0
Type
Type

Figure 4

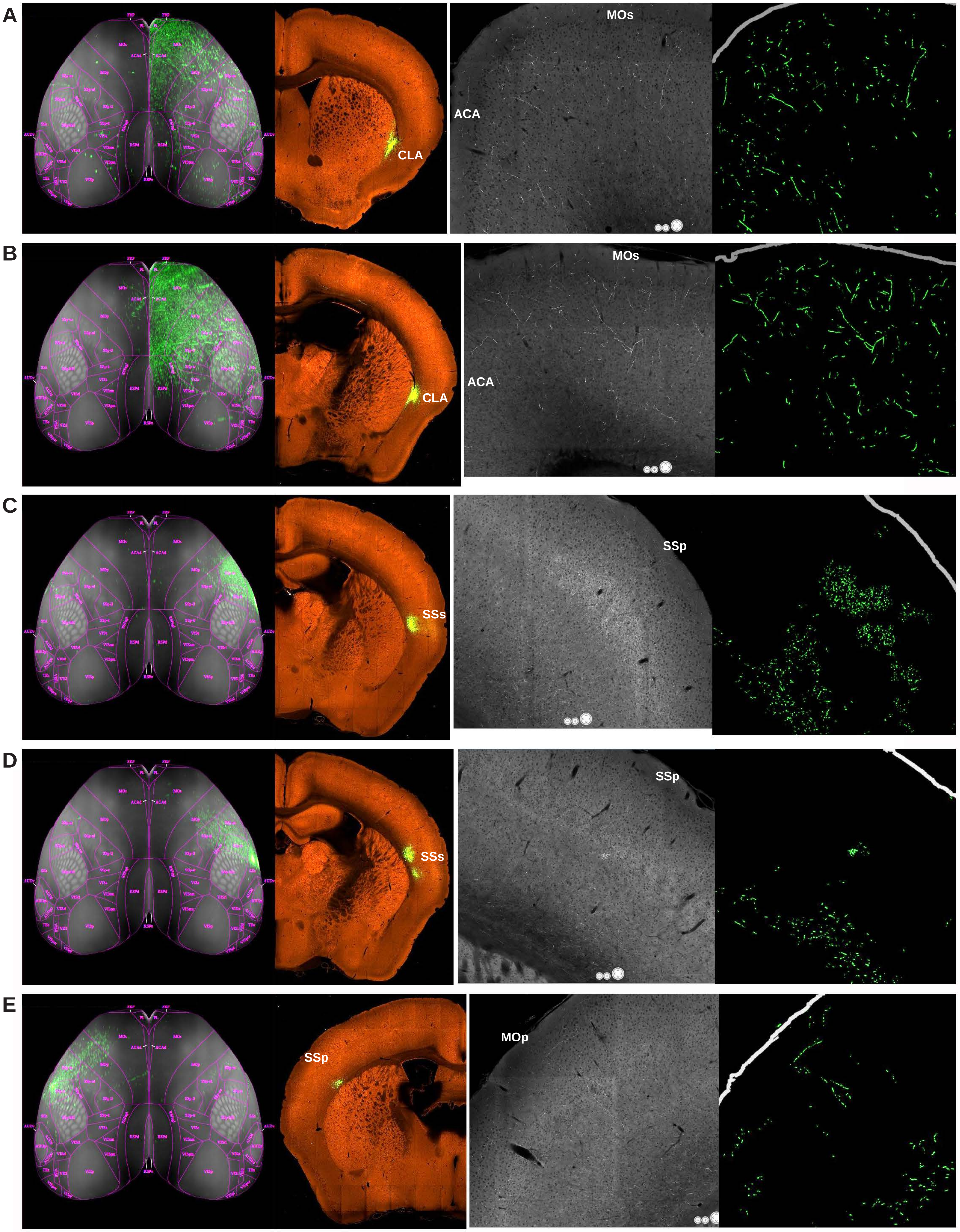


Figure 5

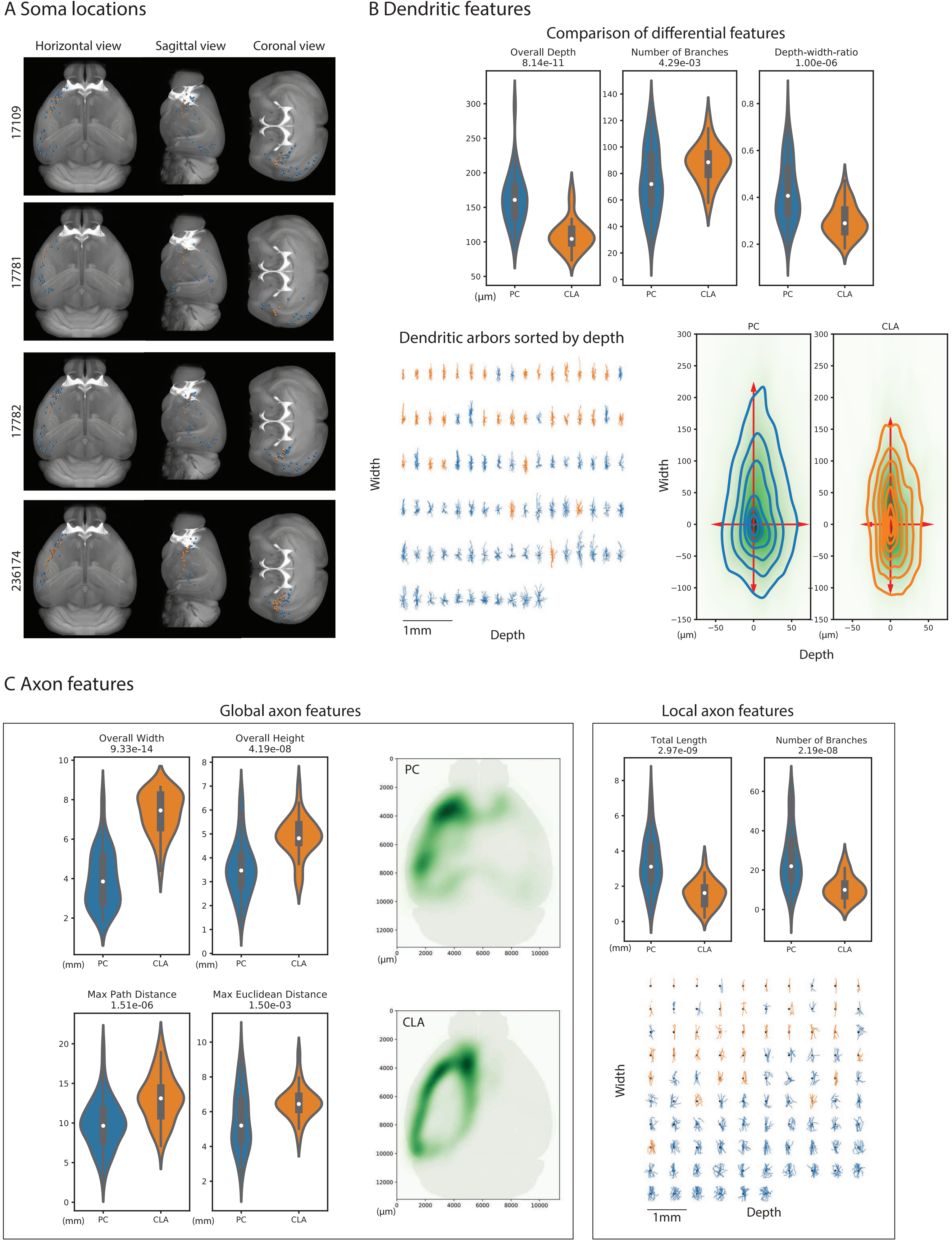

A Soma locations
Horizontal view
Sagittal view
Coronal view
17109
17781
17782
236174
B Dendritic features
Comparison of differential features
Overall Depth
8.14e-11
Number of Branches
4.29e-03
Depth-width-ratio
1.00e-06
300
250
200
150
100
50
(μm)
PC
CLA
140
120
100
80
60
40
20
0
0.8
0.6
0.4
0.2
Dendritic arbors sorted by depth
Width
1mm
Depth
PC
CLA
300
250
200
150
100
50
0
−50
−100
−150
(μm)
−50
0
50
Width
Depth
C Axon features
Global axon features
Overall Width
9.33e-14
Overall Height
4.19e-08
(mm)
PC
CLA
Max Path Distance
1.51e-06
Max Euclidean Distance
1.50e-03
PC
CLA
0
2000
4000
6000
8000
10000
12000
(μm)
Local axon features
Total Length
2.97e-09
Number of Branches
2.19e-08
(mm)
Width
1mm
Depth

Figure 6

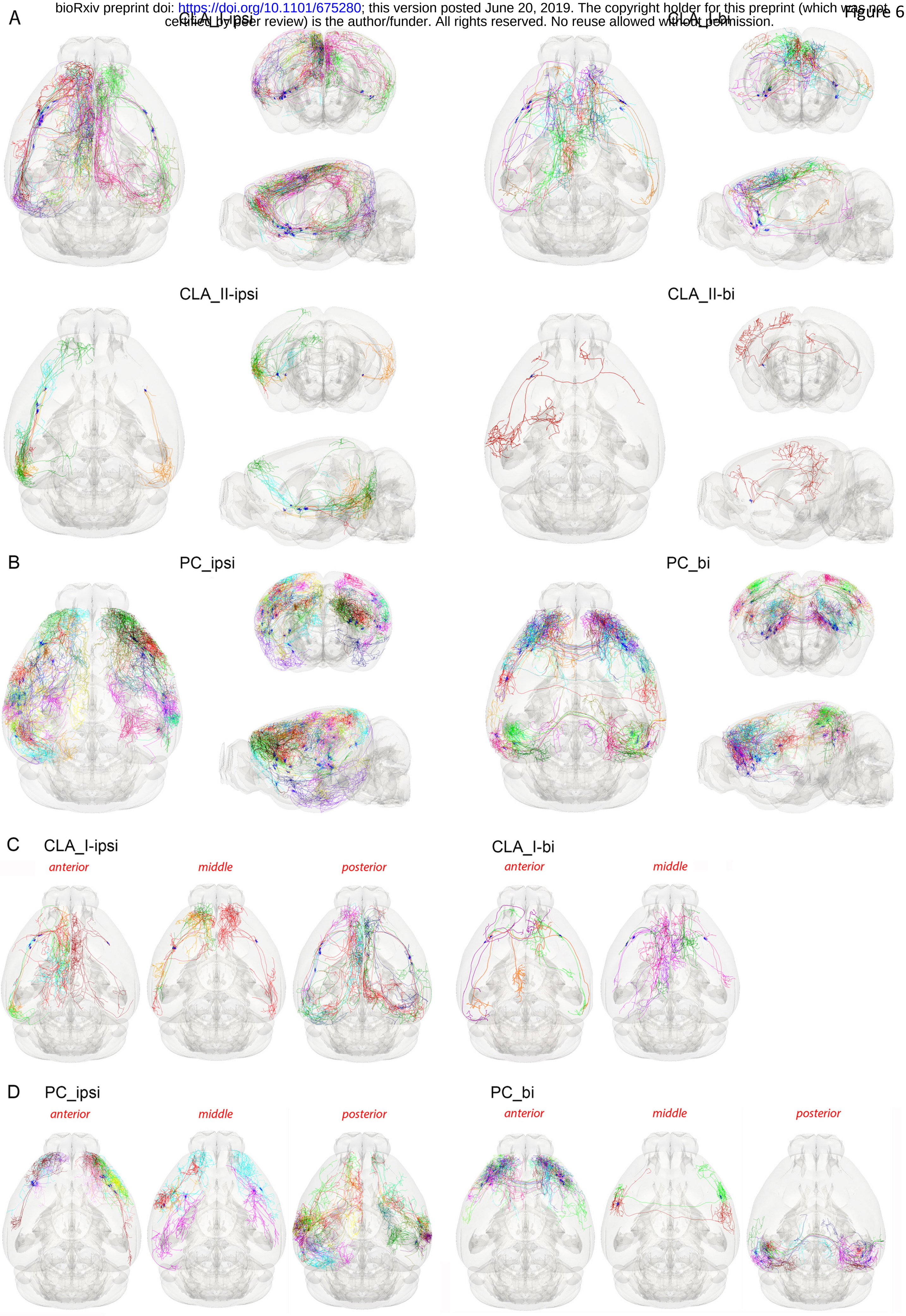

A
CLA_I-ipsi
CLA_I-bi
CLA_II-ipsi
CLA_II-bi
B
PC_ipsi
PC_bi
C
CLA_I-ipsi
anterior
middle
posterior
CLA_I-bi
anterior
middle
D
PC_ipsi
anterior
middle
posterior
PC_bi
anterior
middle
posterior

Figure 7

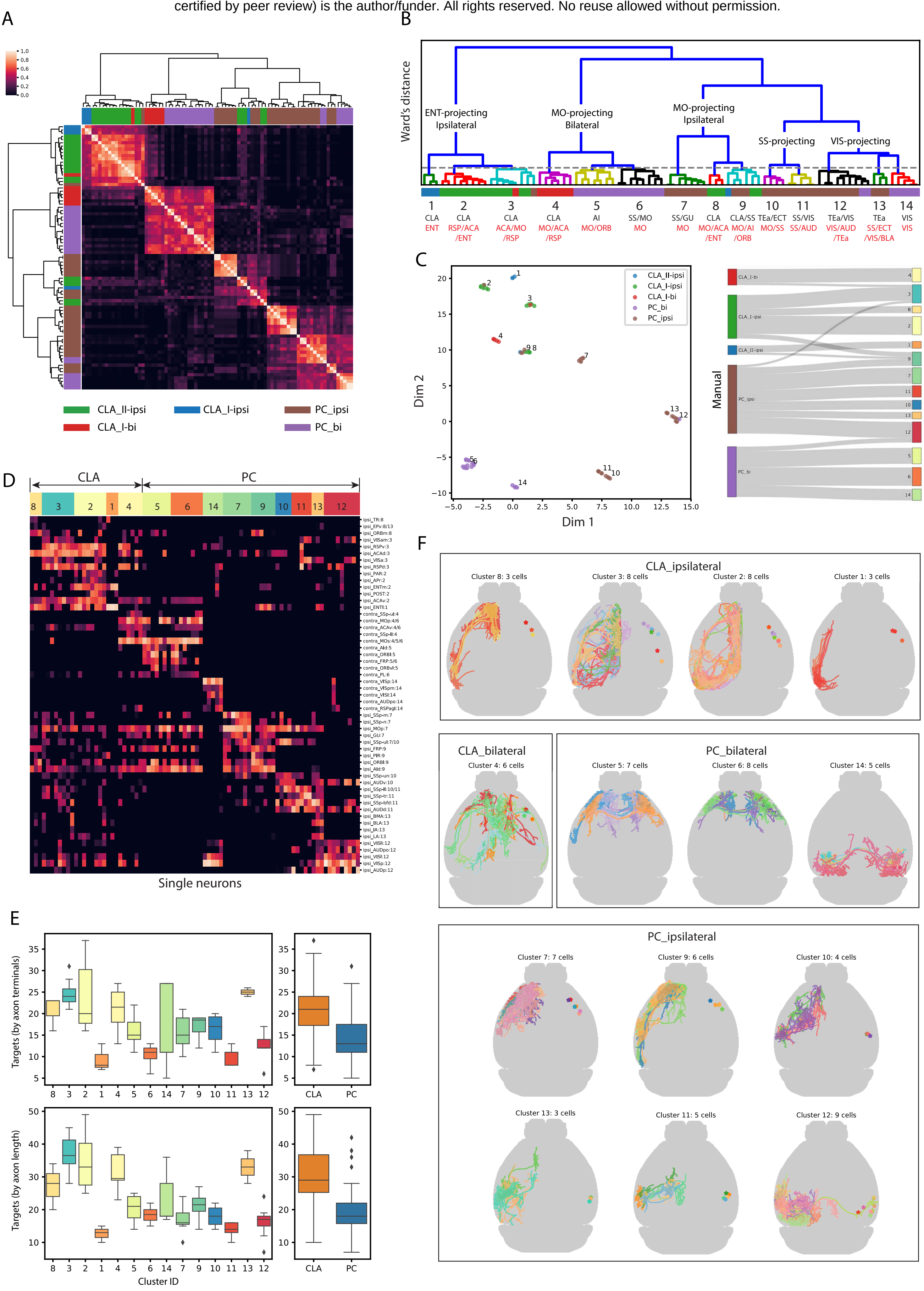

A
1.0
0.8
0.6
0.4
0.2
0.0
CLA_II-ipsi
CLA_I-ipsi
PC_ipsi
CLA_I-bi
PC_bi
B
Ward's distance
ENT-projecting Ipsilateral
MO-projecting Bilateral
MO-projecting Ipsilateral
SS-projecting
VIS-projecting
1 CLA ENT
2 CLA RSP/ACA /ENT
3 CLA ACA/MO /RSP
4 CLA MO/ACA /RSP
5 AI MO/ORB
6 SS/MO MO
7 SS/GU MO
8 CLA MO/ACA /ENT
9 CLA/SS MO/AI /ORB
10 TEa/ECT MO/SS
11 SS/VIS SS/AUD
12 TEa/VIS VIS/AUD /TEa
13 TEa SS/ECT /VIS/BLA
14 VIS VIS
C
Dim 2
Dim 1
CLA_II-ipsi
CLA_I-ipsi
CLA_I-bi
PC_bi
PC_ipsi
Manual
Computational
CLA_I-bi
CLA_I-ipsi
CLA_II-ipsi
PC_ipsi
PC_bi
D
CLA
PC
8 3 2 1 4 5 6 14 7 9 10 11 13 12
Single neurons
E
Targets (by axon terminals)
Targets (by axon length)
Cluster ID
CLA
PC
F
CLA_ipsilateral
Cluster 8: 3 cells
Cluster 3: 8 cells
Cluster 2: 8 cells
Cluster 1: 3 cells
CLA_bilateral
Cluster 4: 6 cells
PC_bilateral
Cluster 5: 7 cells
Cluster 6: 8 cells
Cluster 14: 5 cells
PC_ipsilateral
Cluster 7: 7 cells
Cluster 9: 6 cells
Cluster 10: 4 cells
Cluster 13: 3 cells
Cluster 11: 5 cells
Cluster 12: 9 cells

Figure 8

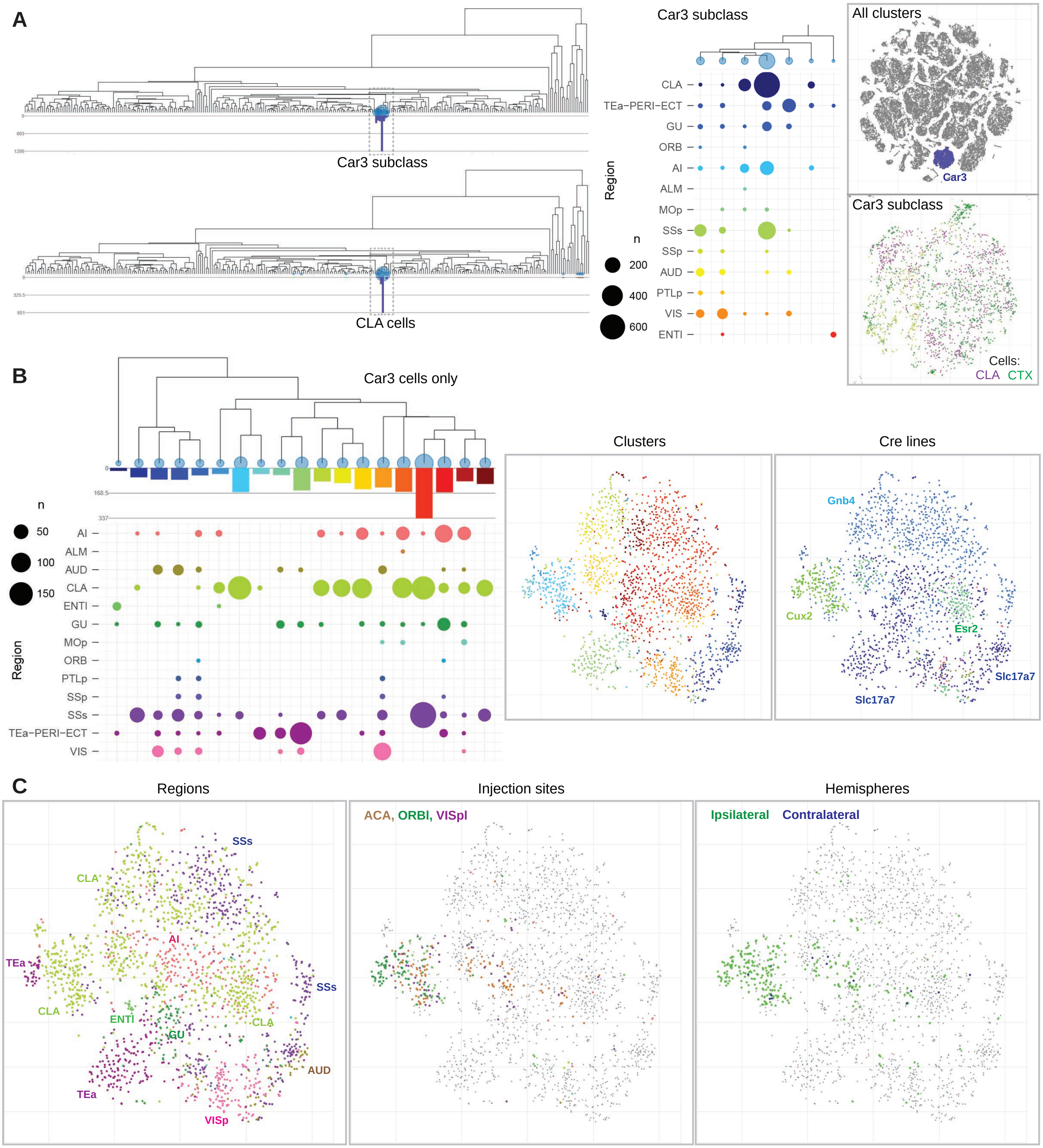

A
Car3 subclass
CLA cells
Car3 subclass
Region
n
200
400
600
CLA
TEa−PERI−ECT
GU
ORB
AI
ALM
MOp
SSs
SSp
AUD
PTLp
VIS
ENTI
All clusters
Car3
Car3 subclass
Cells:
CLA CTX
B
Car3 cells only
n
50
100
150
Region
AI
ALM
AUD
CLA
ENTI
GU
MOp
ORB
PTLp
SSp
SSs
TEa−PERI−ECT
VIS
168.5
337
Clusters
Cre lines
Gnb4
Cux2
Esr2
Slc17a7
Slc17a7
C
Regions
SSs
CLA
AI
TEa
SSs
CLA
ENTI
CLA
GU
AUD
TEa
VISp
Injection sites
ACA, ORBl, VISpl
Hemispheres
Ipsilateral
Contralateral

Figure S1

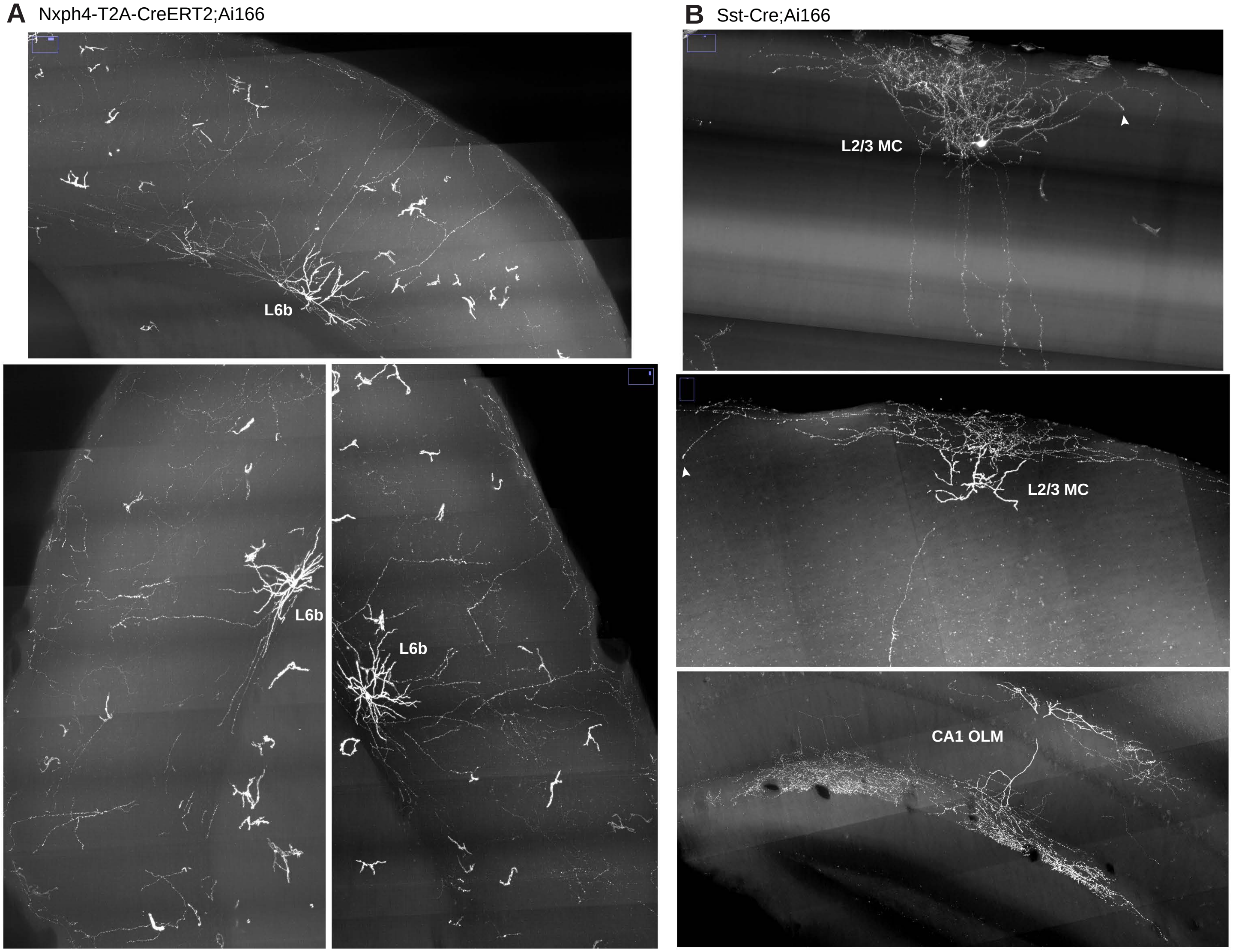

A
Nxph4-T2A-CreERT2;Ai166
L6b
L6b
L6b
B
Sst-Cre;Ai166
L2/3 MC
L2/3 MC
CA1 OLM

Figure S2

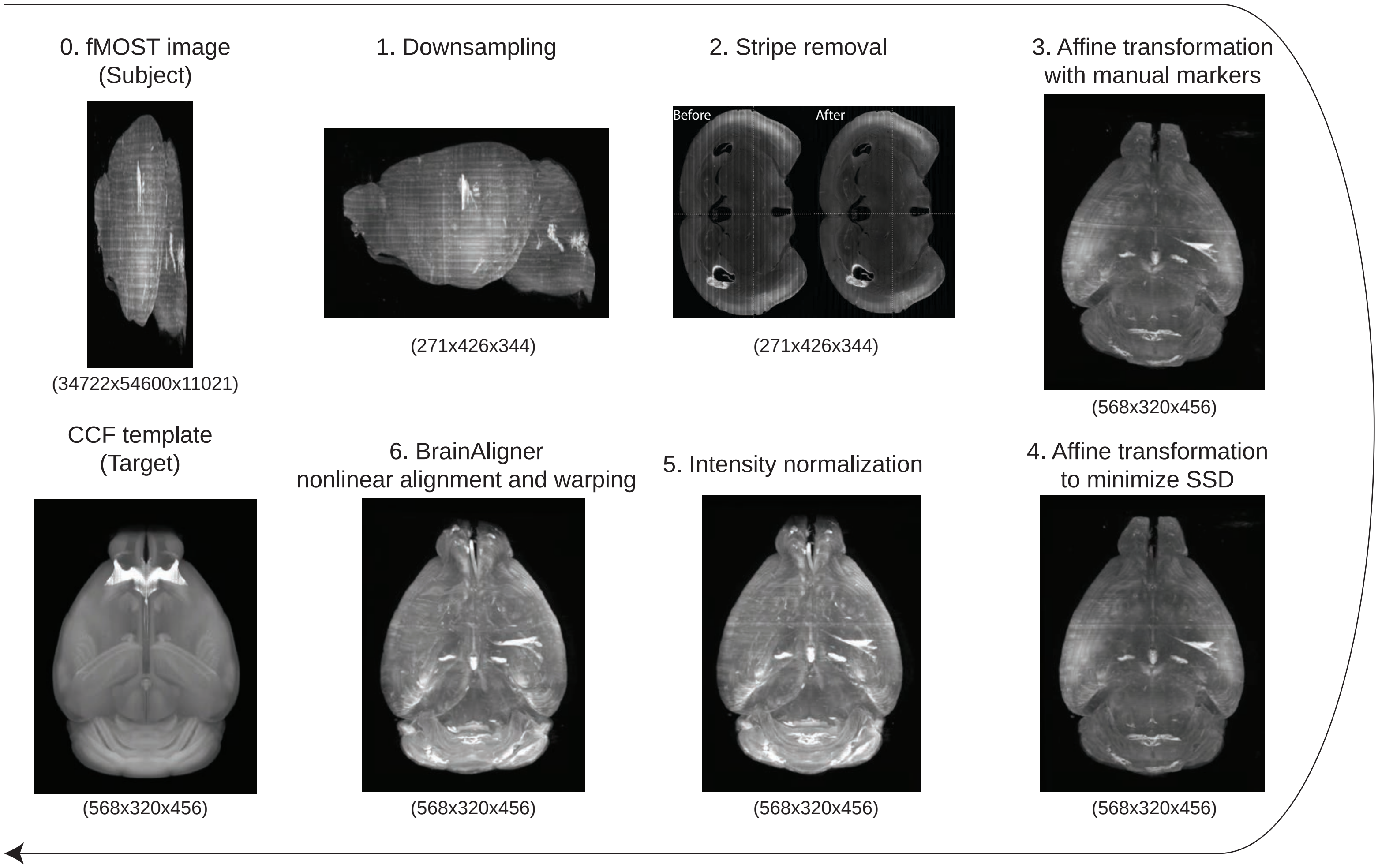
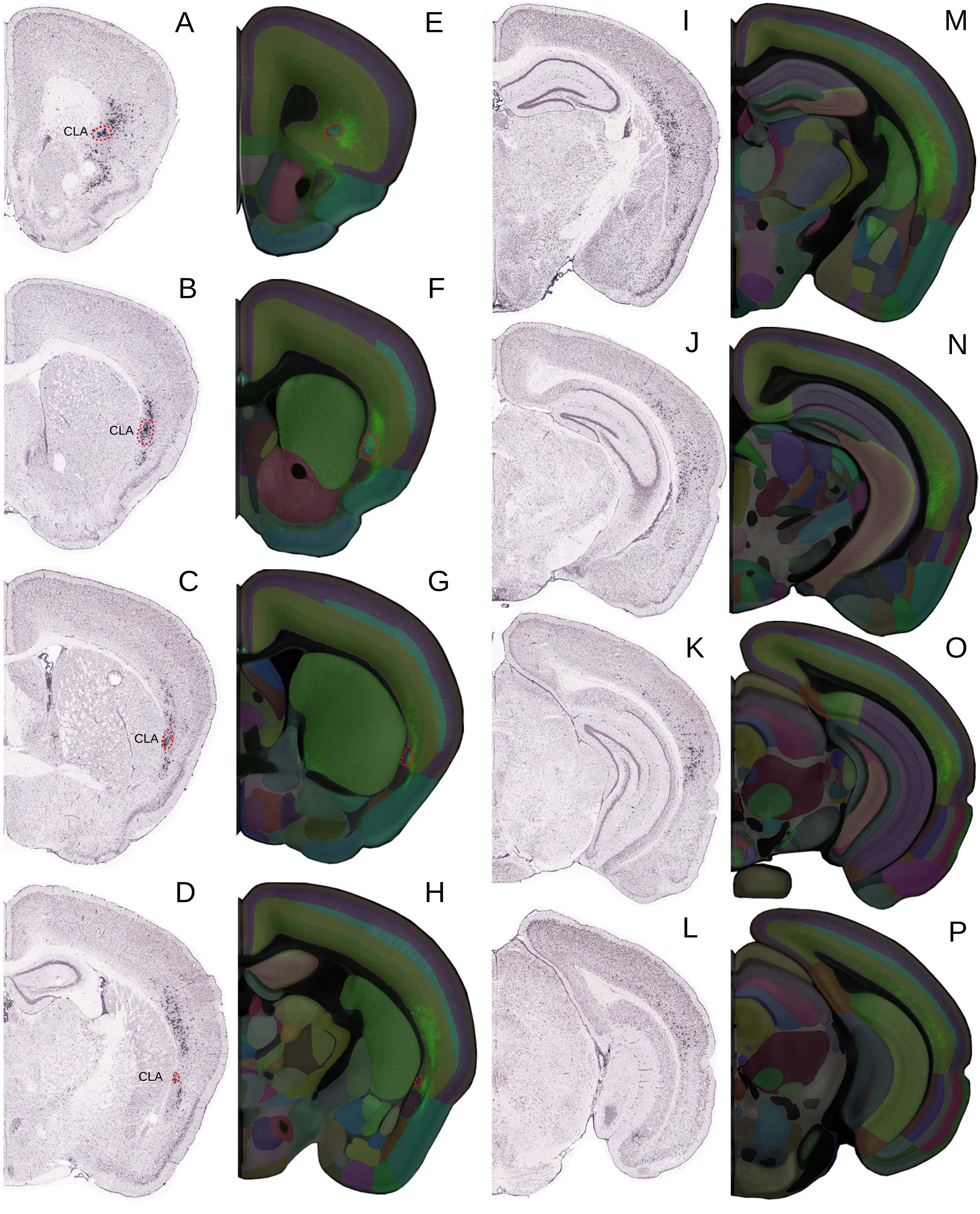
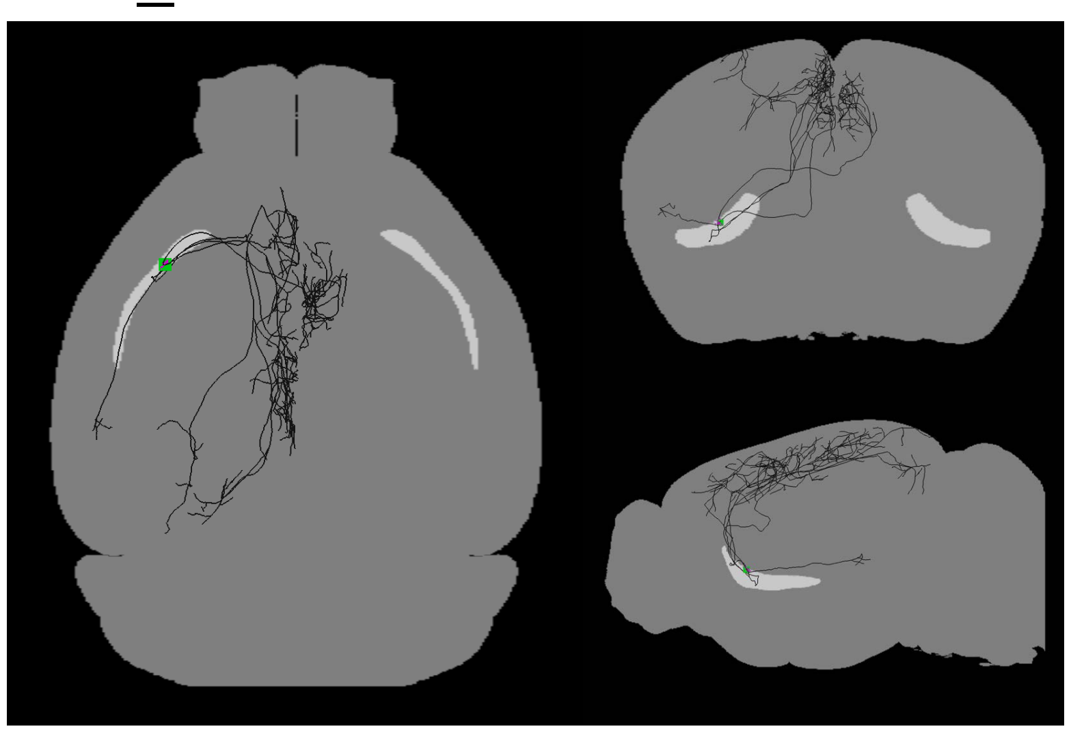
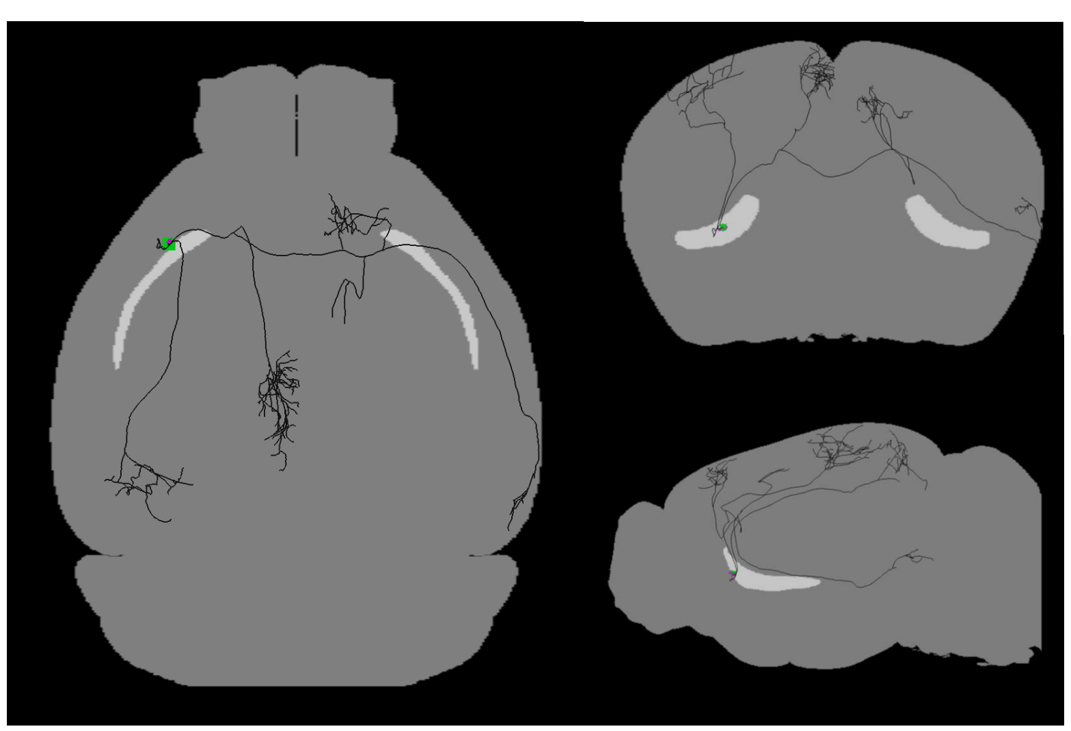

0. fMOST image
(Subject)
(34722x54600x11021)
1. Downsampling
(271x426x344)
2. Stripe removal
Before
After
(271x426x344)
3. Affine transformation
with manual markers
(568x320x456)
CCF template
(Target)
(568x320x456)
6. BrainAligner
nonlinear alignment and warping
(568x320x456)
5. Intensity normalization
(568x320x456)
4. Affine transformation
to minimize SSD
(568x320x456)

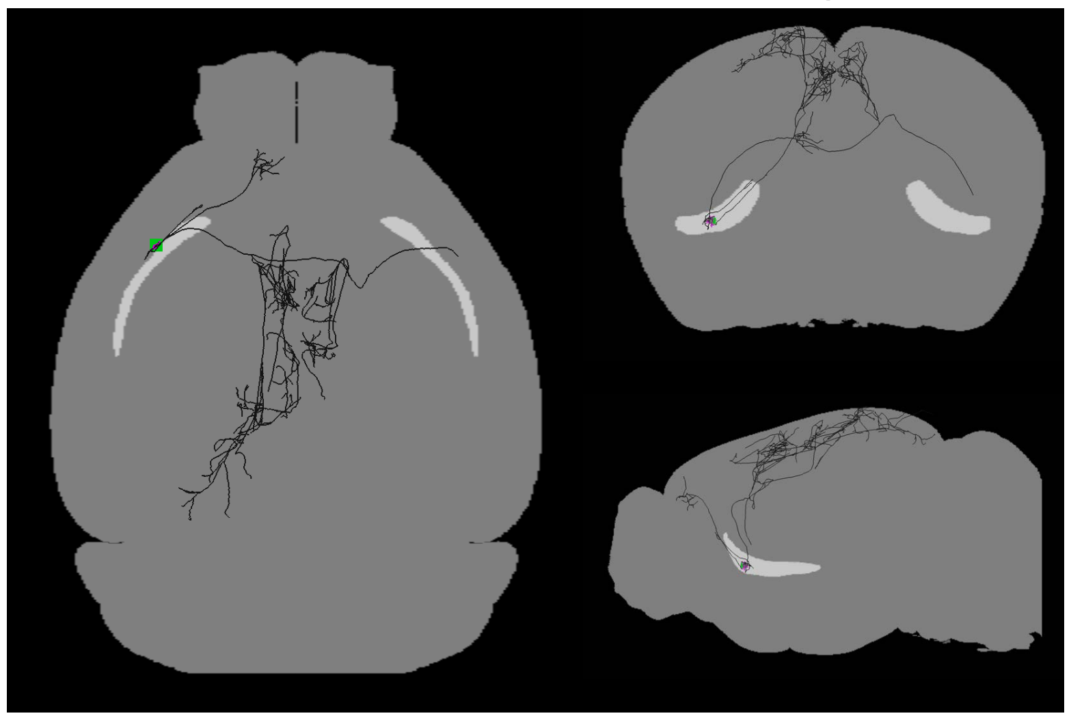
Figure S3

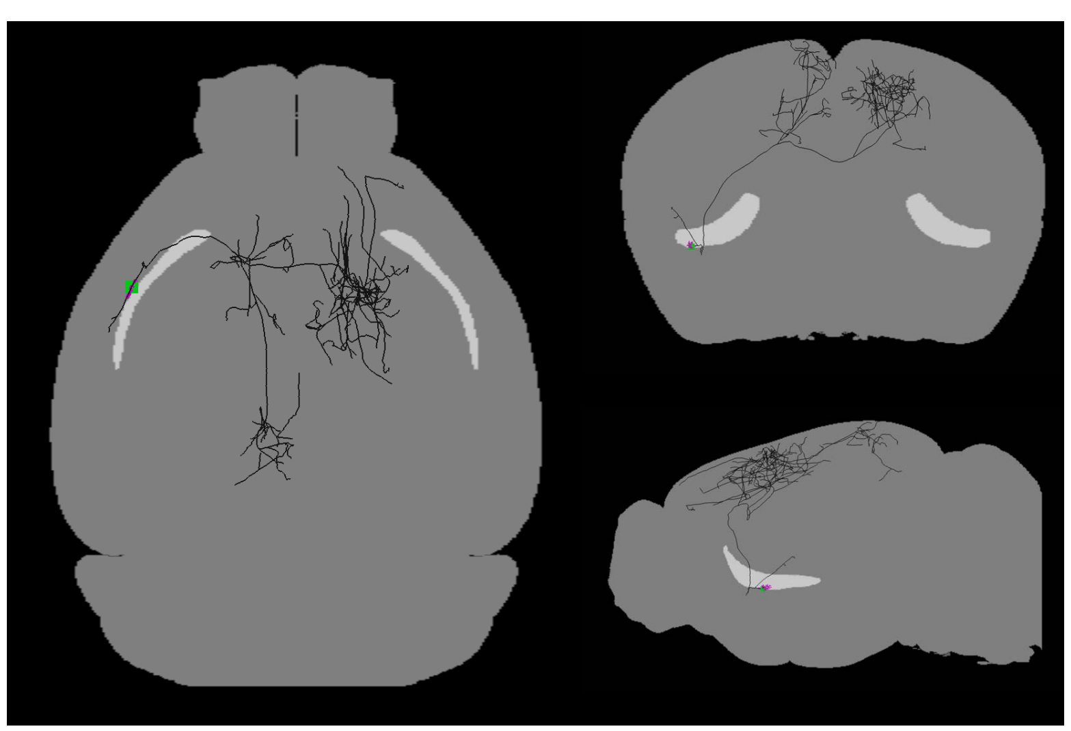


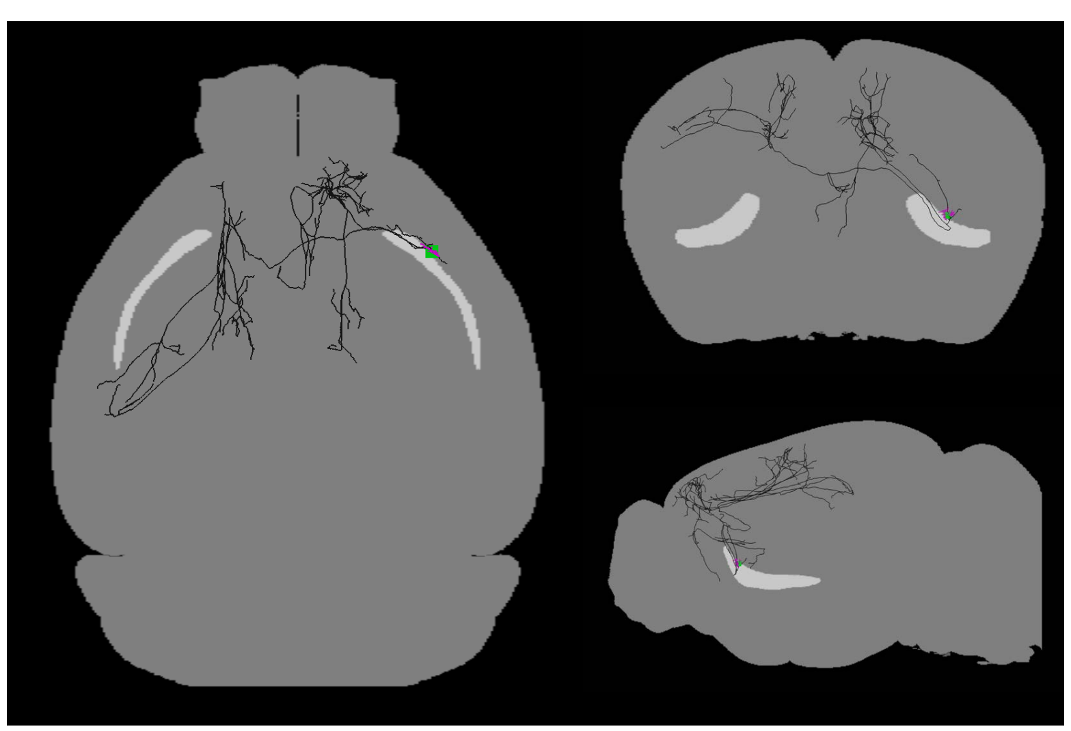
CLA_bilateral

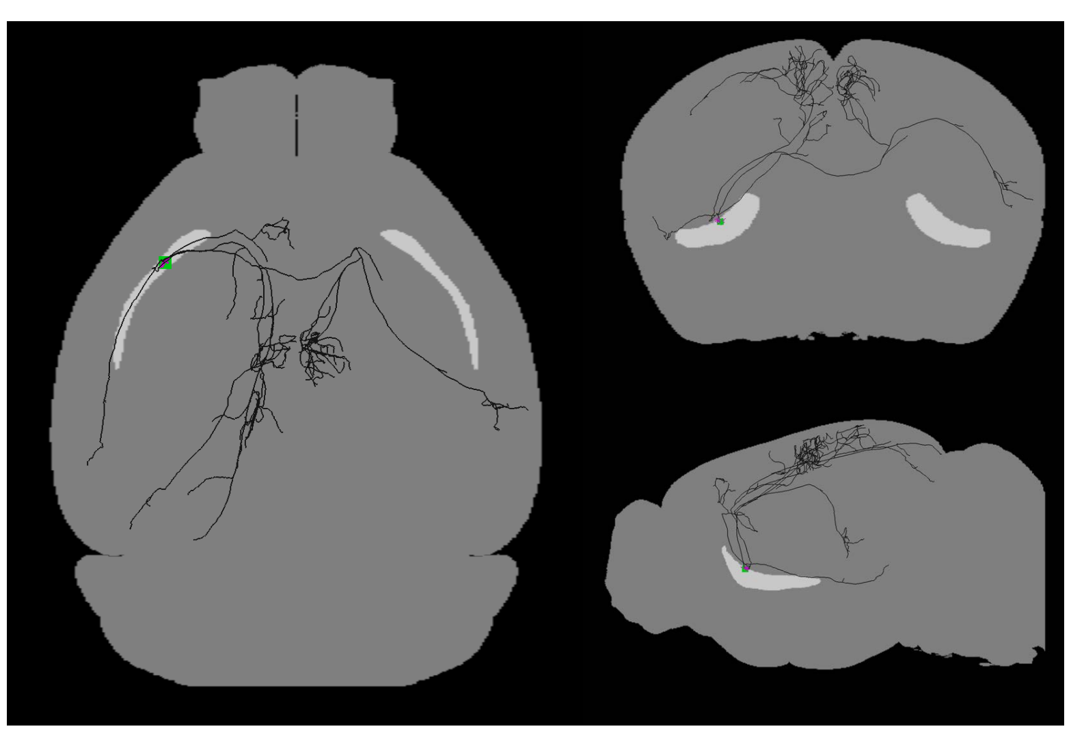

236174_03429_03528_X12632_Y10625_CLA_Bilateral_C3

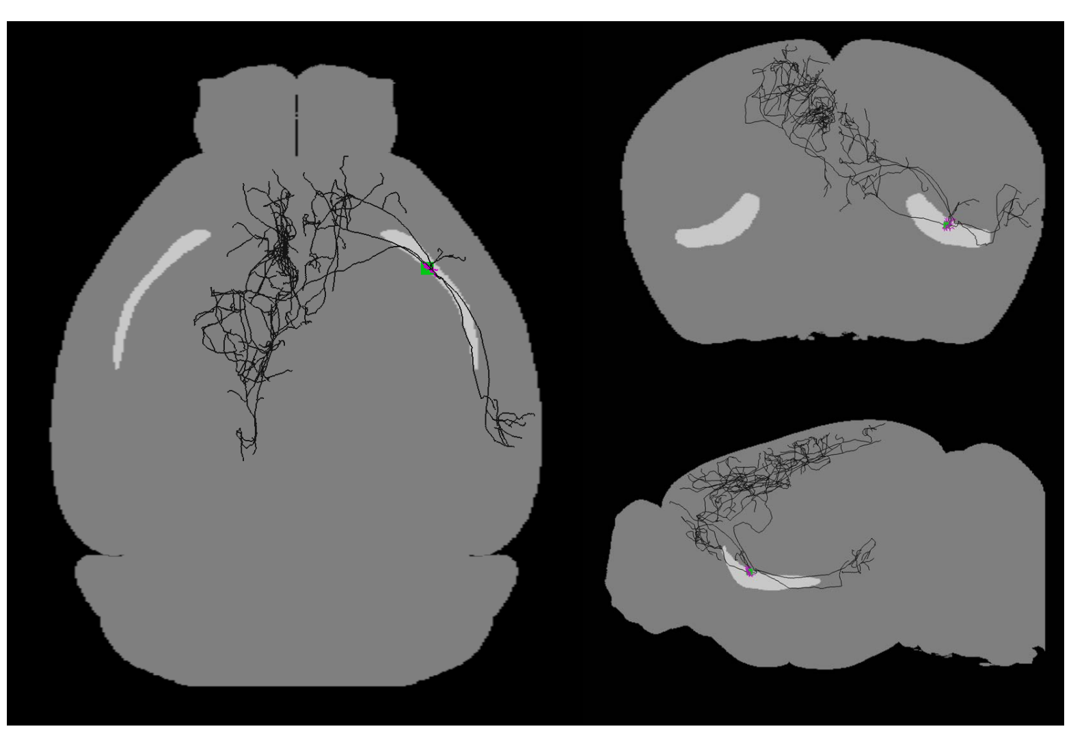

17109_02401_02500_X09695_Y09693_CLA_Bilateral_C4

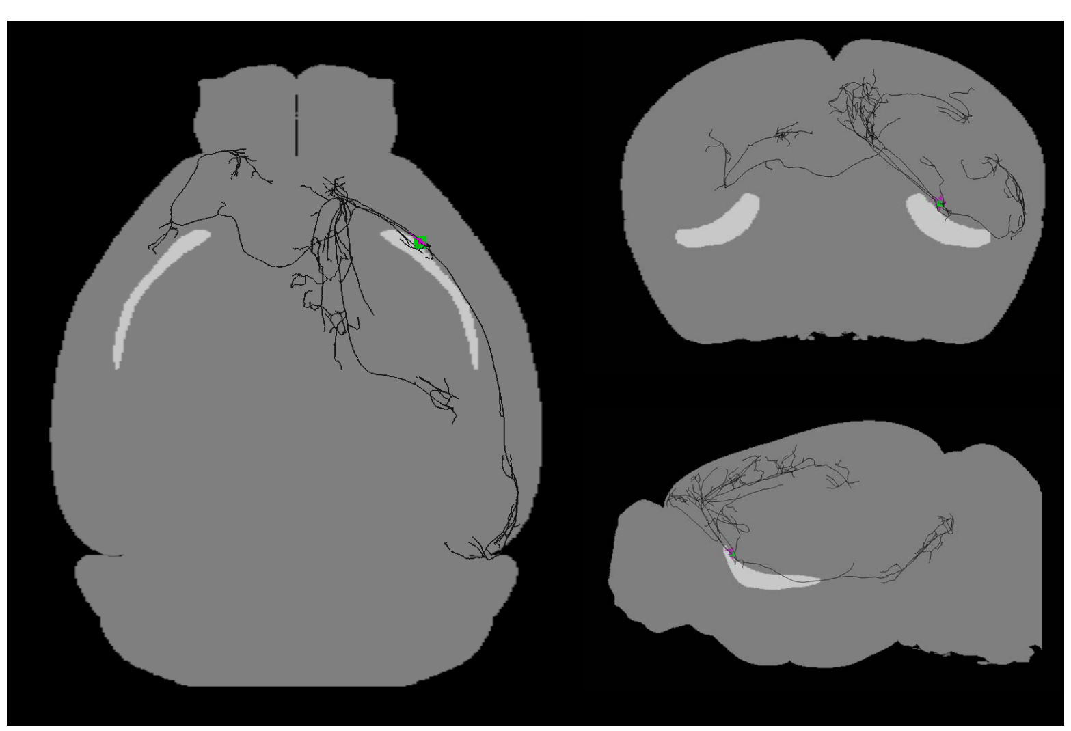

17109_02601_02700_X10213_Y08783_CLA_Bilateral_C4

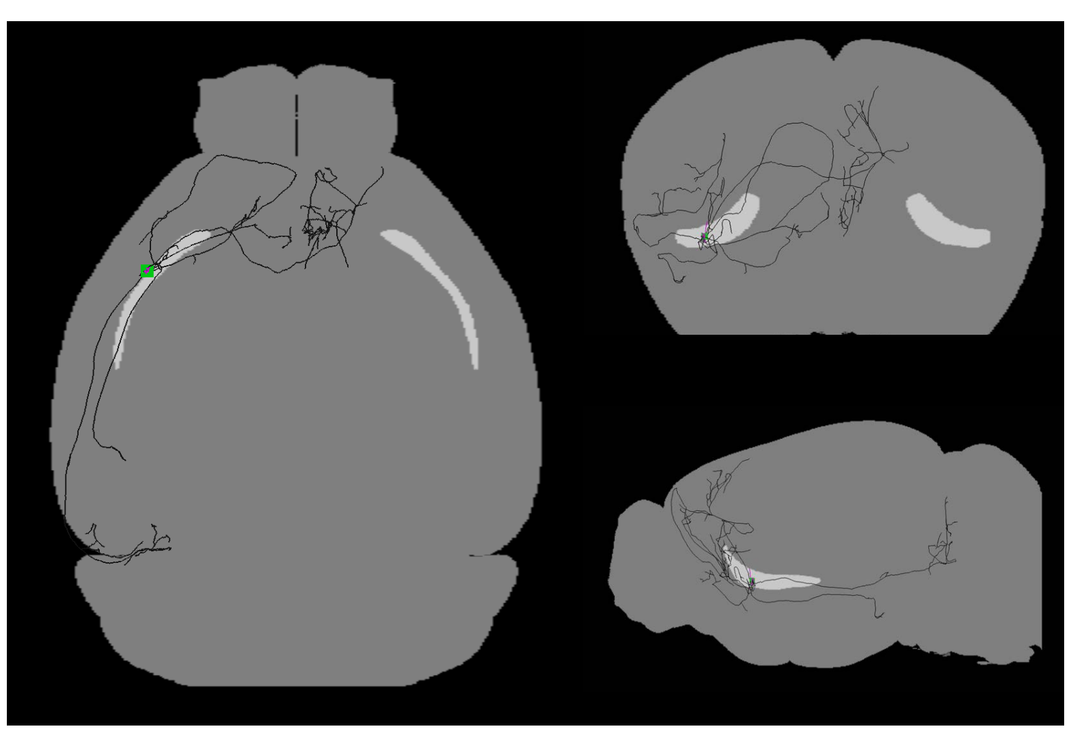

17109_03101_03200_X10824_Y07188_CLA_Bilateral_C4

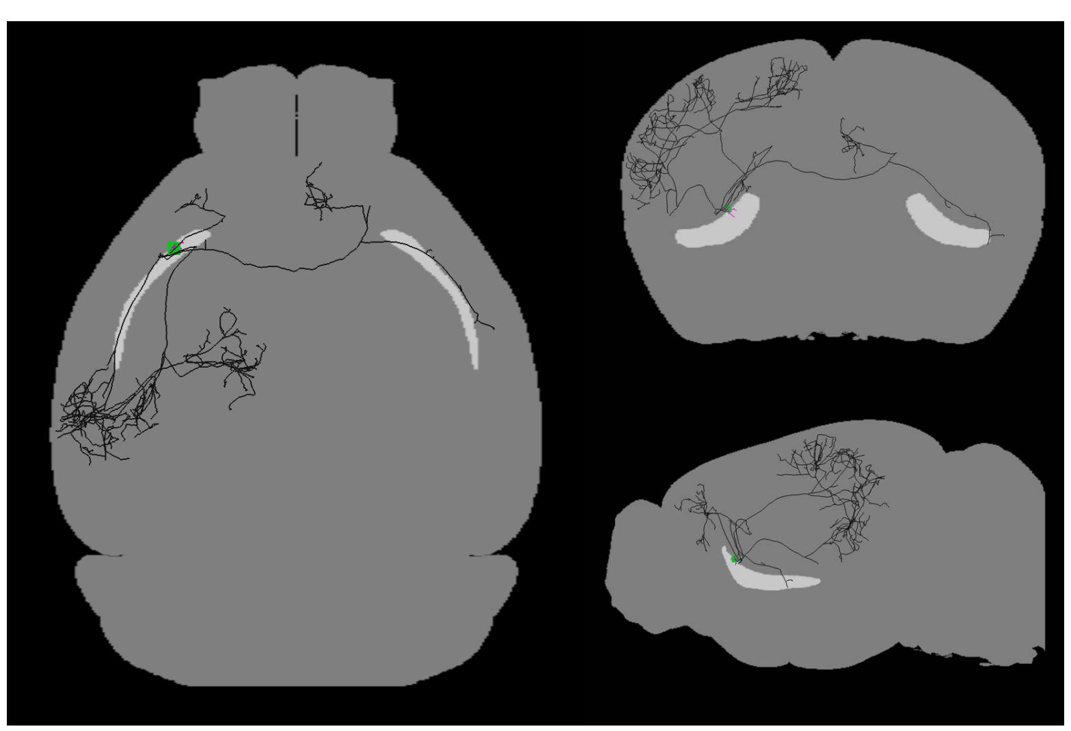

17782_03487_03487_X11014_Y17041_CLA_Bilateral_C4


236174_03447_03459_X12562_Y10626_CLA_Bilateral_C4

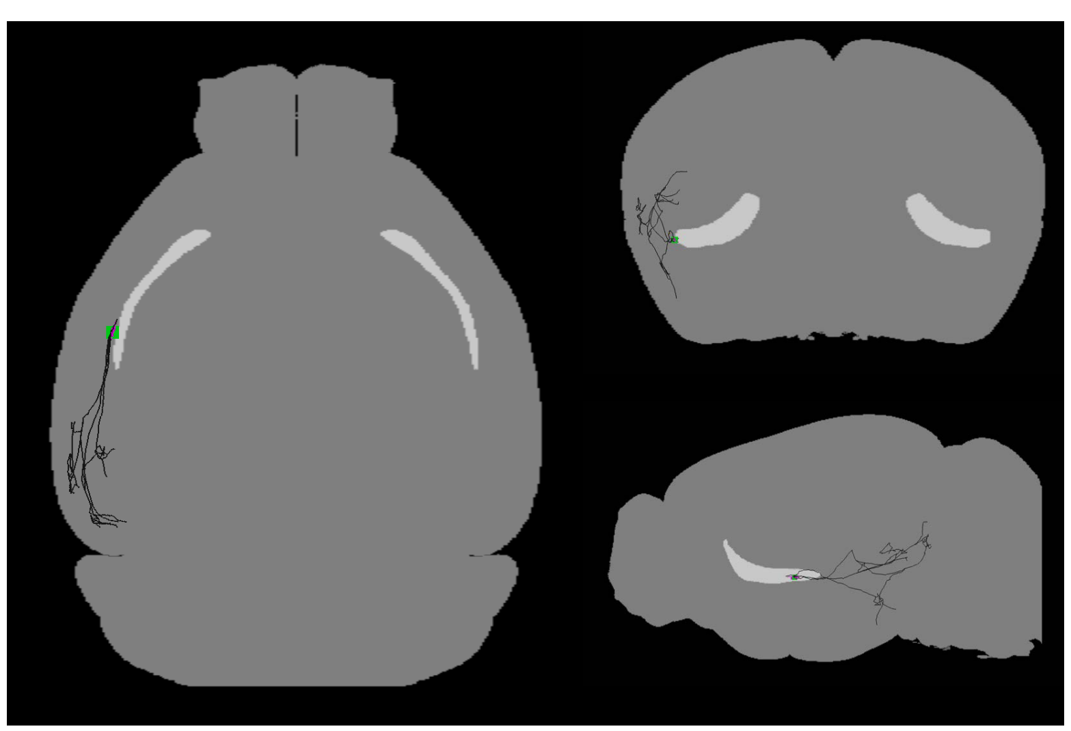

236174_03536_03545_X15159_Y25525_CLA_Bilateral_C4

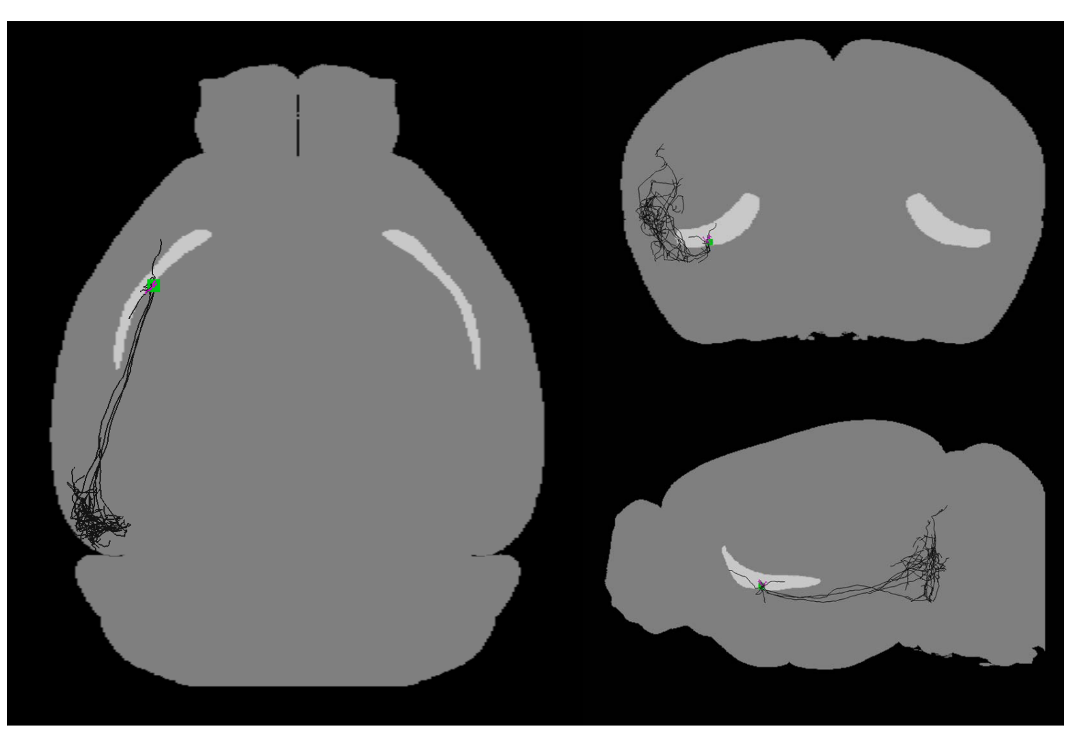

17782_03284_03284_X11909_Y16428_CLA_Bilateral_Cnan

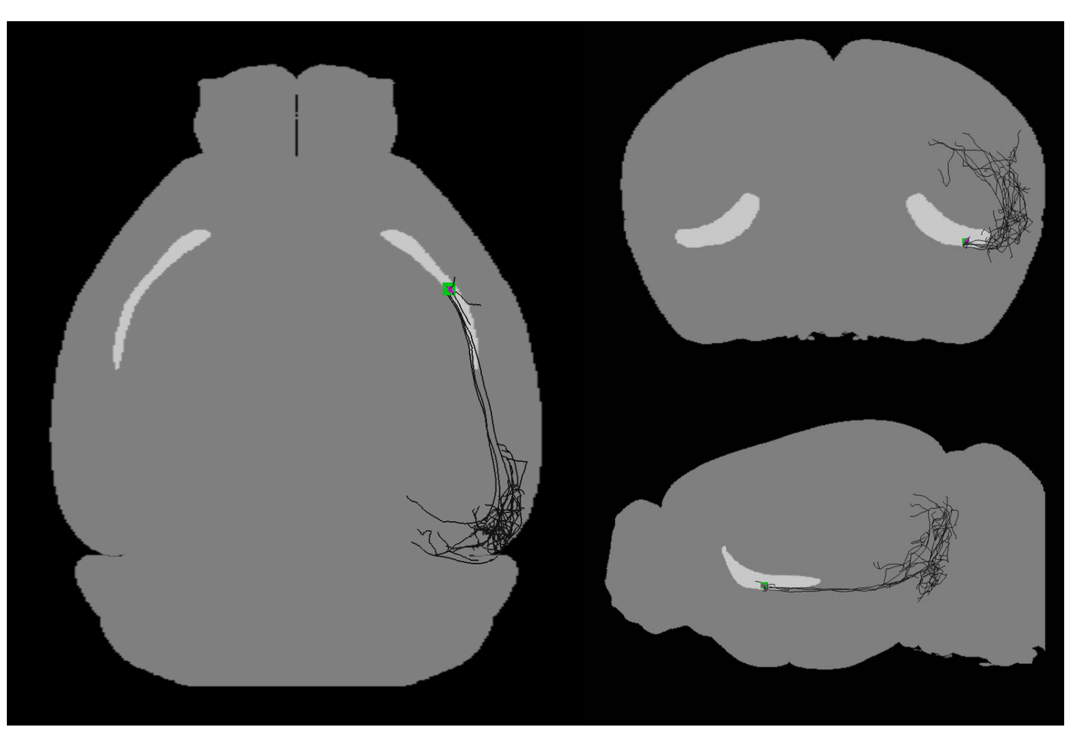

17782_03672_03672_X34784_Y19432_CLA_Bilateral_Cnan

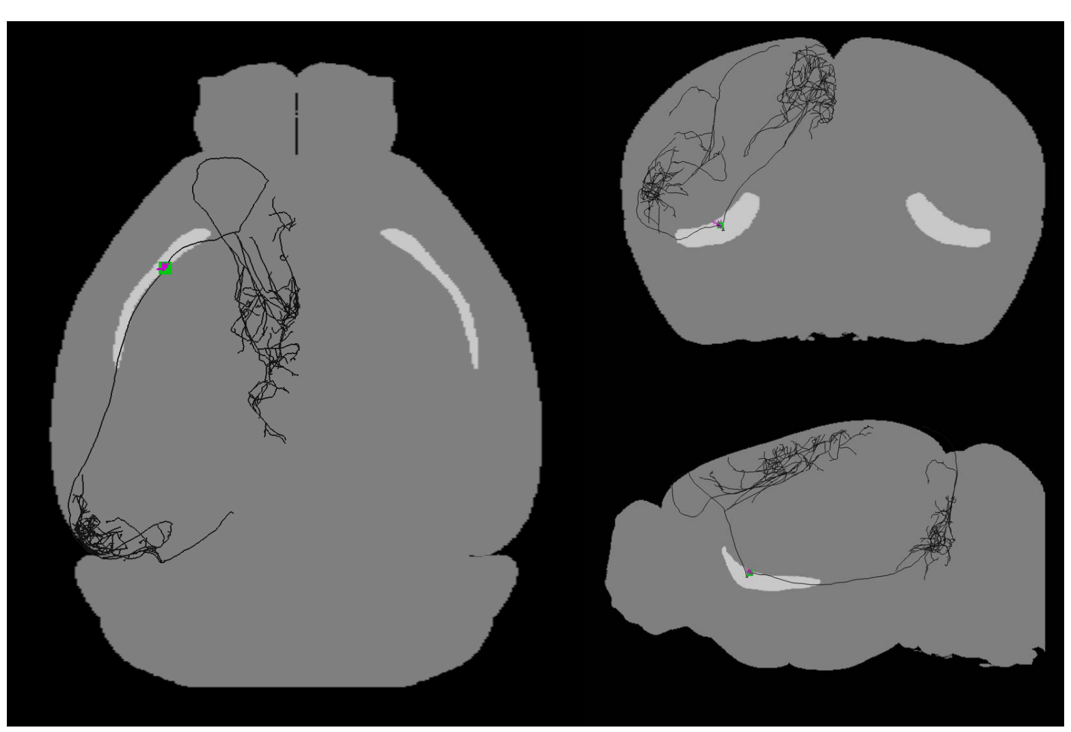

236174_03229_03328_X11950_Y11335_CLA_Bilateral_Cnan

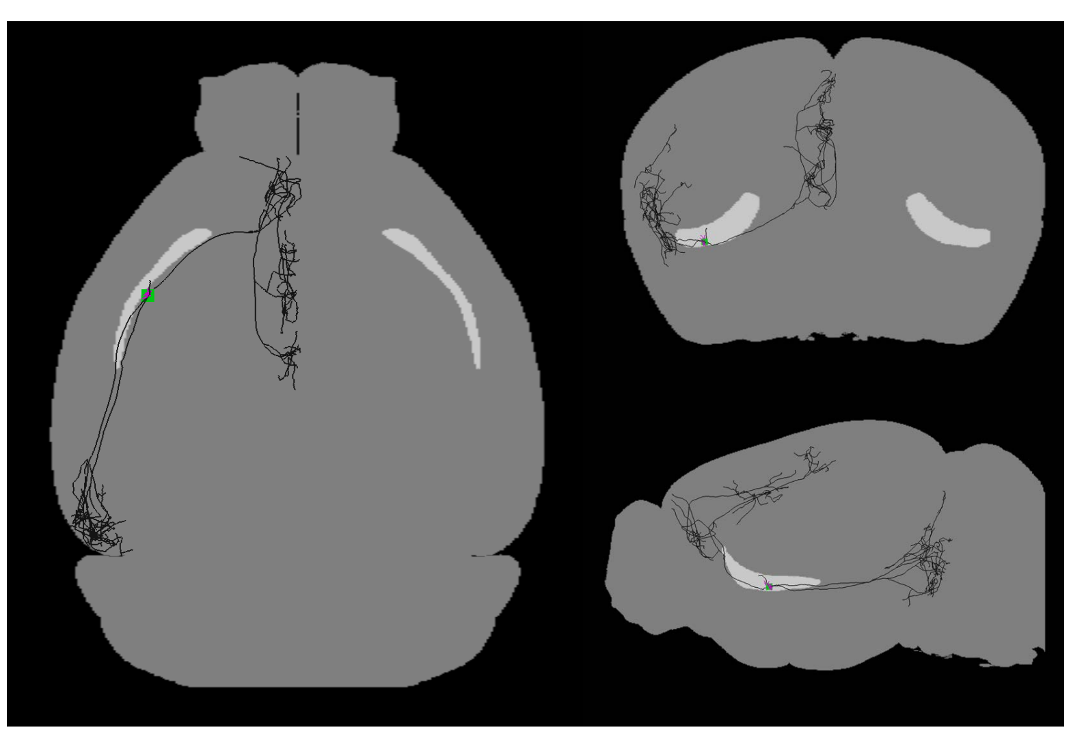
CLA_ipsilateral

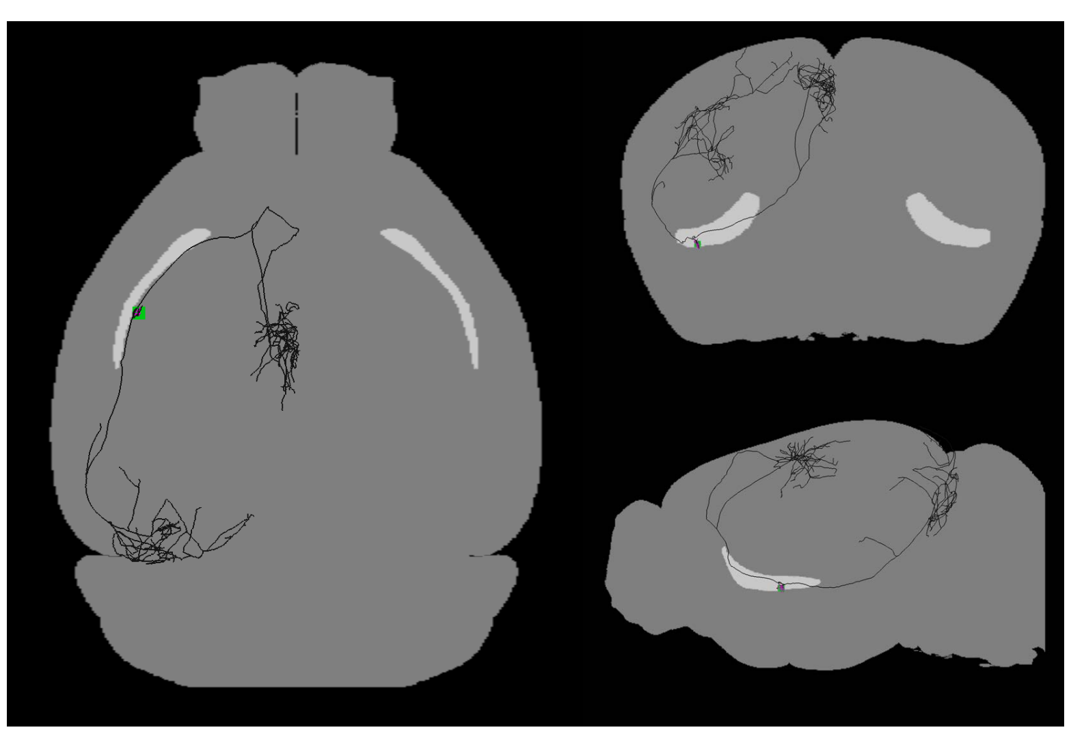

17781_04698_04698_X17857_Y11456_CLA_Ipsilateral_C1

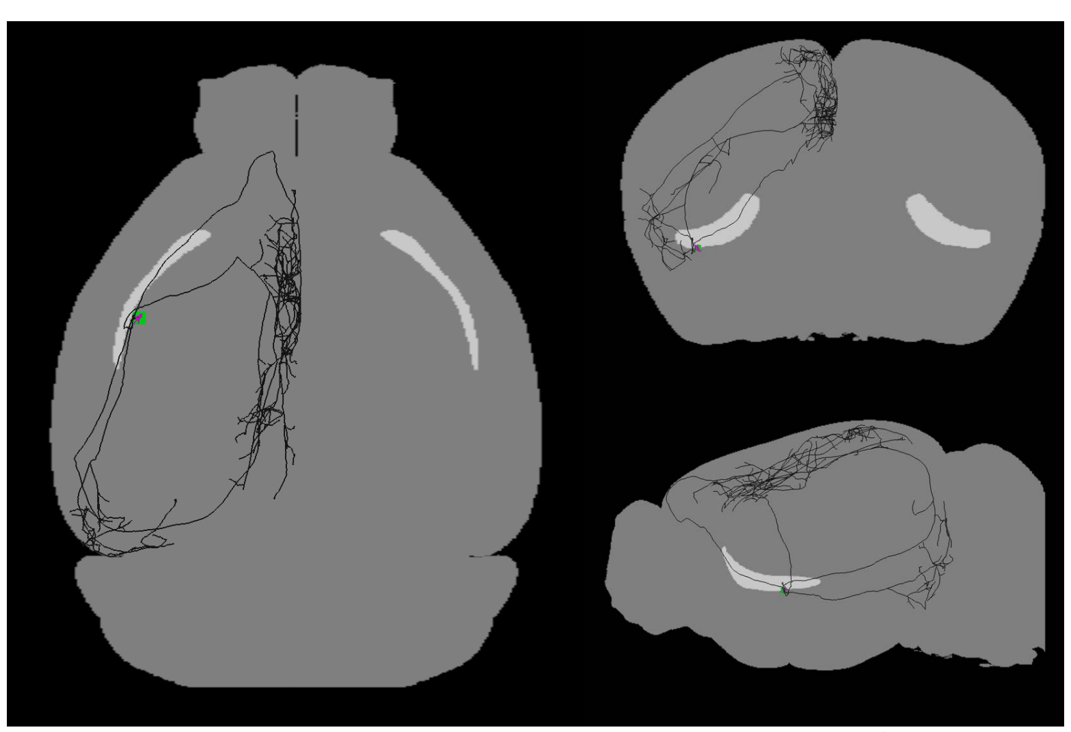

236174_03729_03828_X13645_Y09551_CLA_Ipsilateral_C1

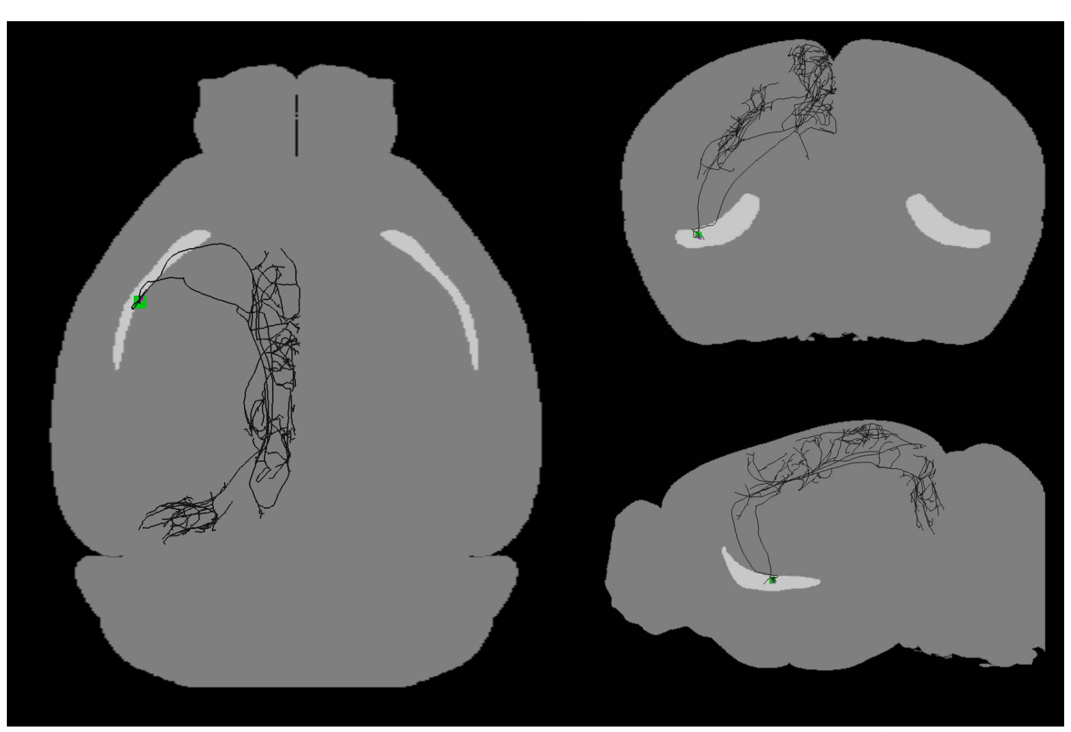

236174_03829_03928_X16301_Y26647_CLA_Ipsilateral_C1

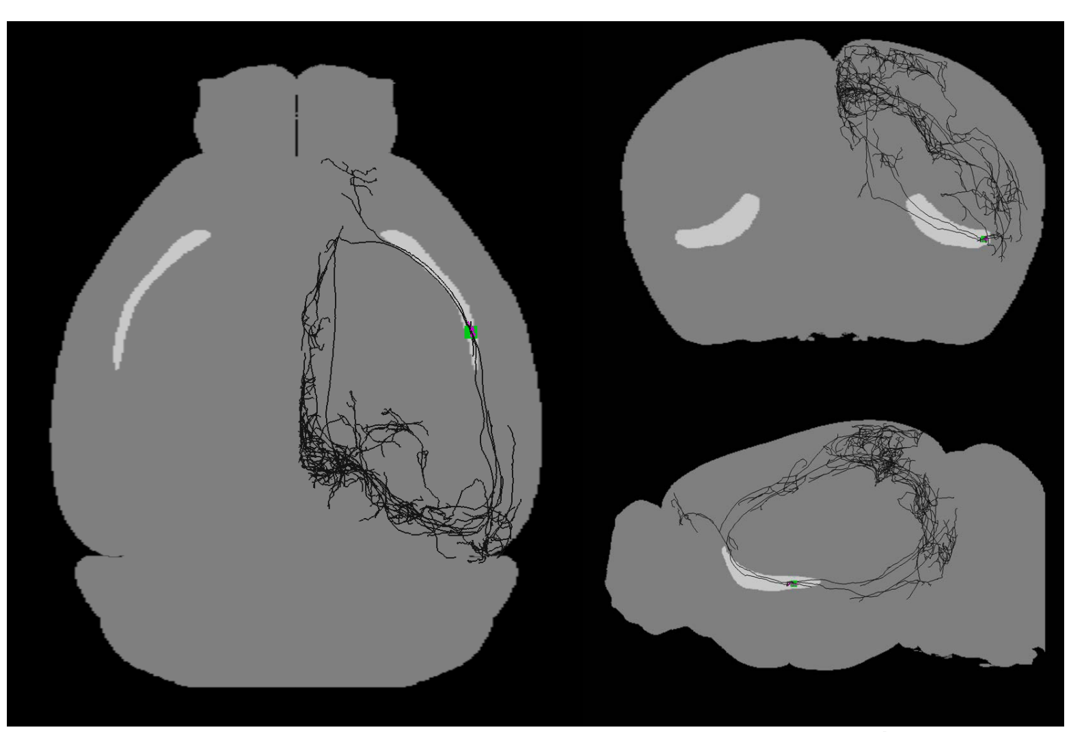

236174_03529_03628_X12805_Y10541_CLA_Ipsilateral_C2

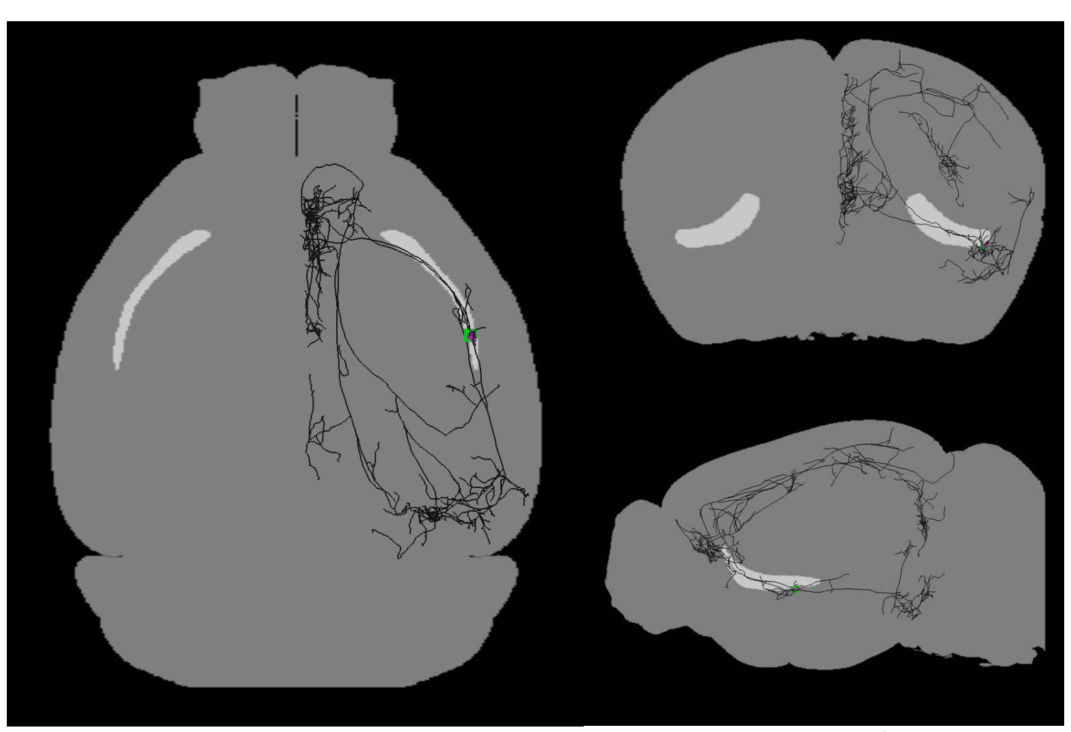

236174_03929_04028_X13599_Y09165_CLA_Ipsilateral_C2

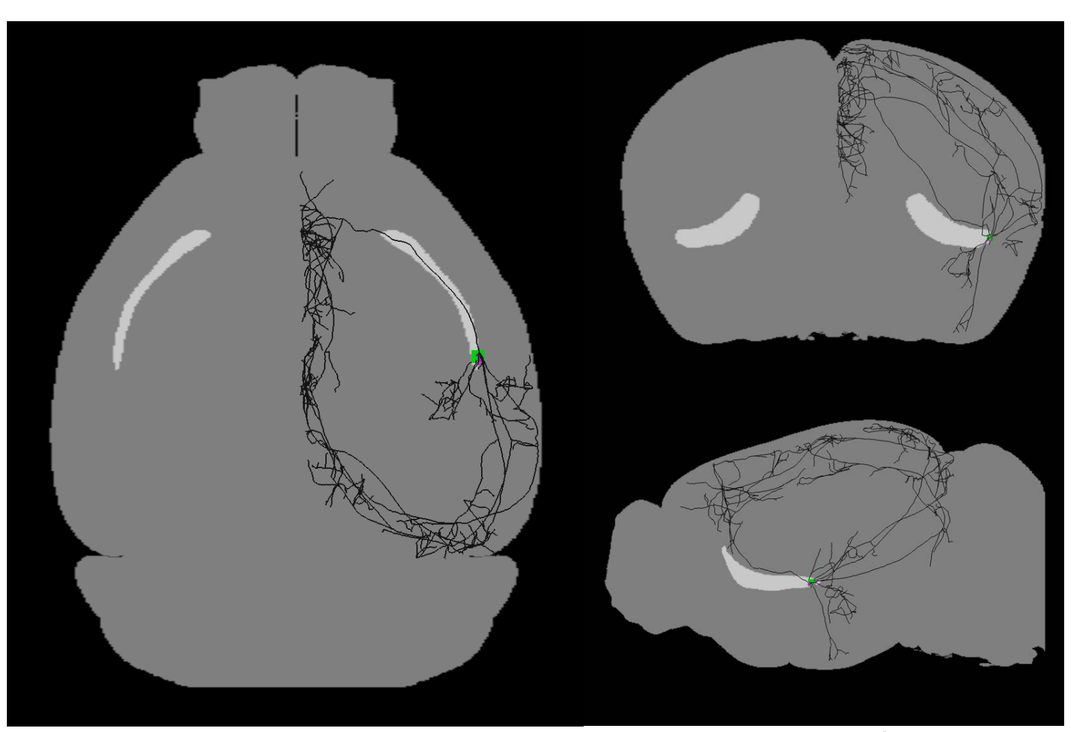

236174_04229_04328_X13663_Y08589_CLA_Ipsilateral_C2

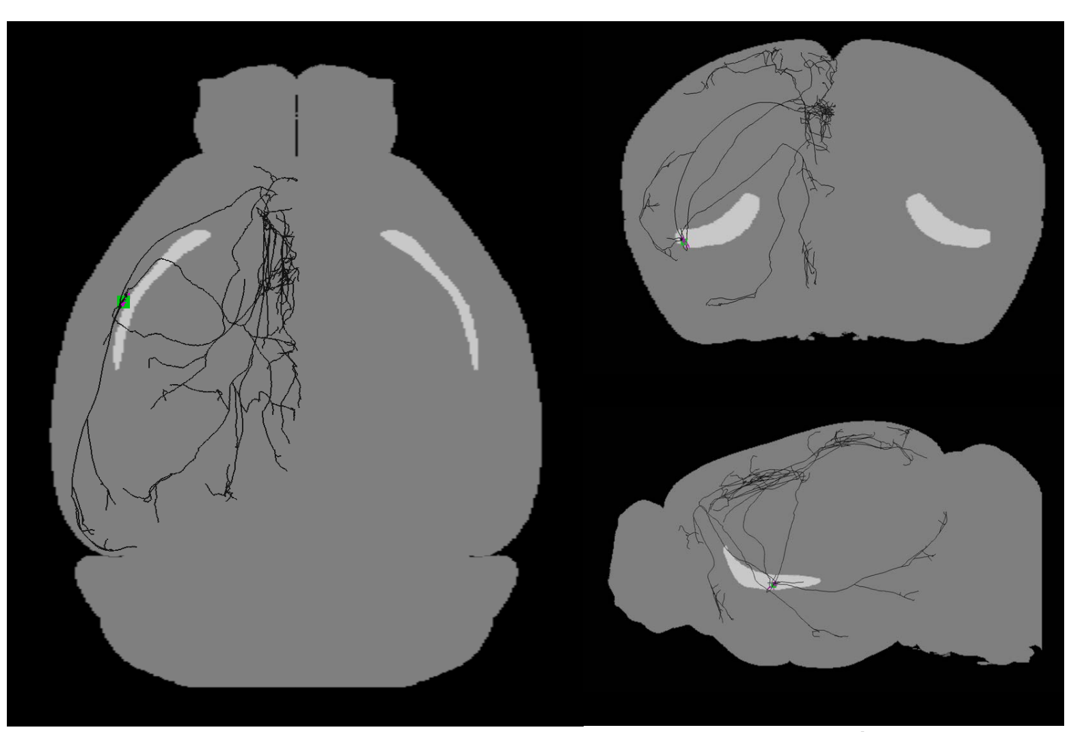

236174_04266_04266_X13848_Y08550_CLA_Ipsilateral_C2

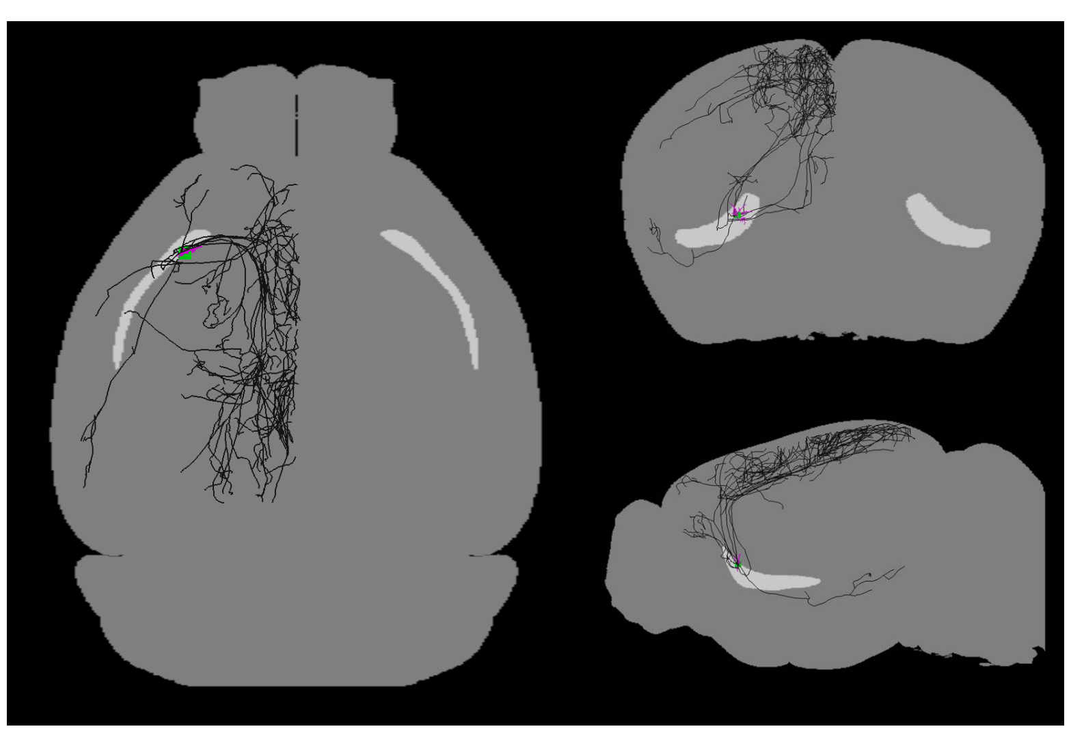

236174_04029_04128_X13079_Y08858_CLA_Ipsilateral_C2

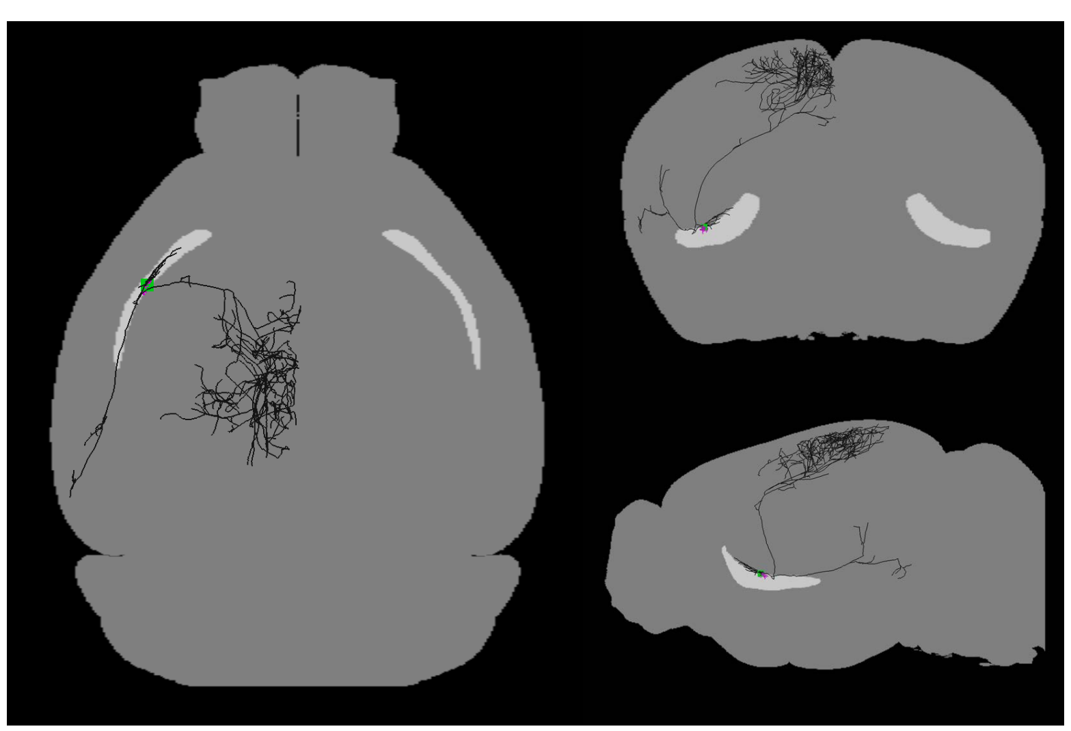

236174_04647_04647_X16405_Y27845_CLA_Ipsilateral_C2

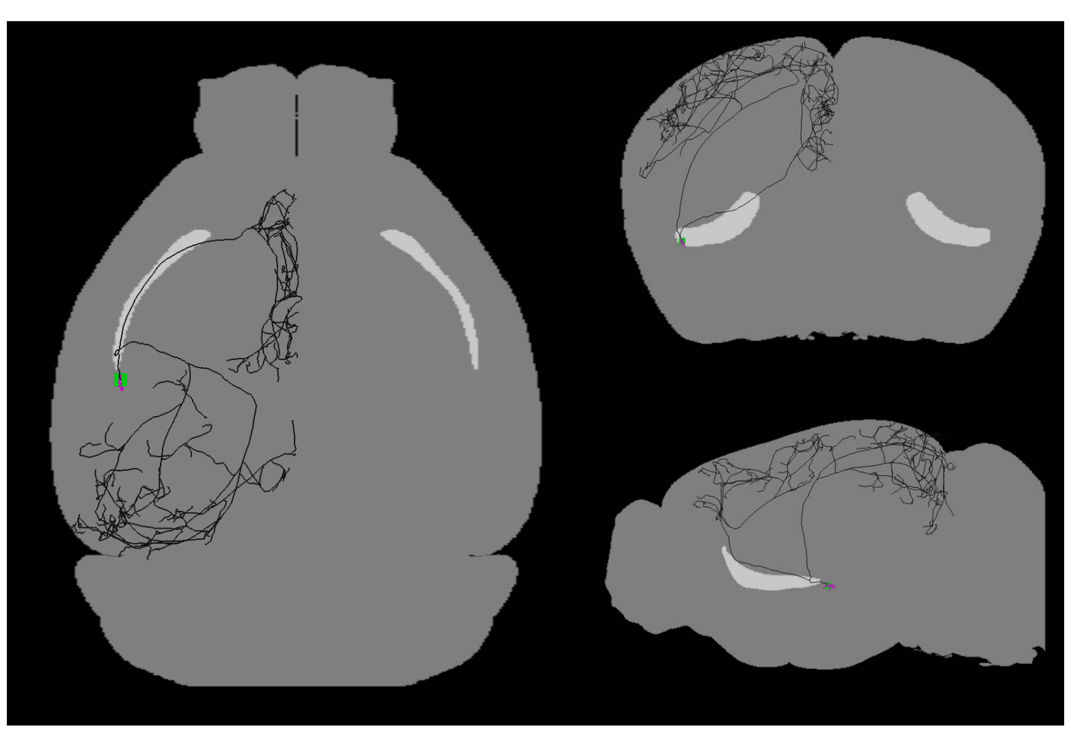

236174_04729_04829_X16869_Y27809_CLA_Ipsilateral_C2

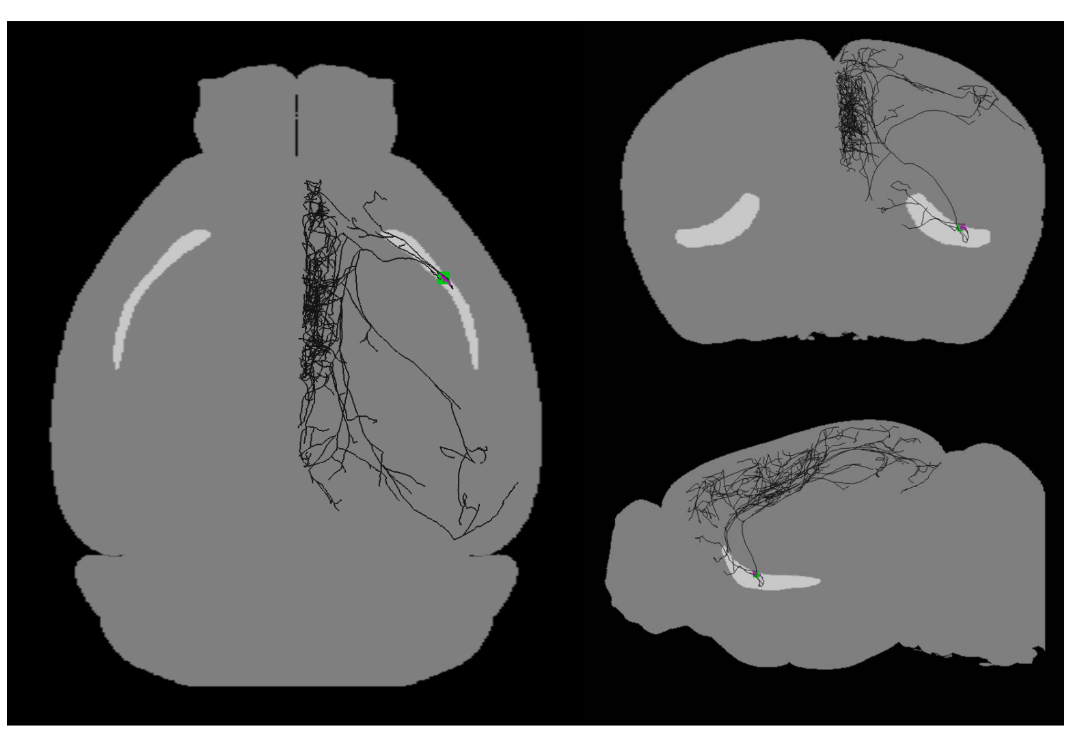

236174_05138_05237_X16501_Y28259_CLA_Ipsilateral_C2

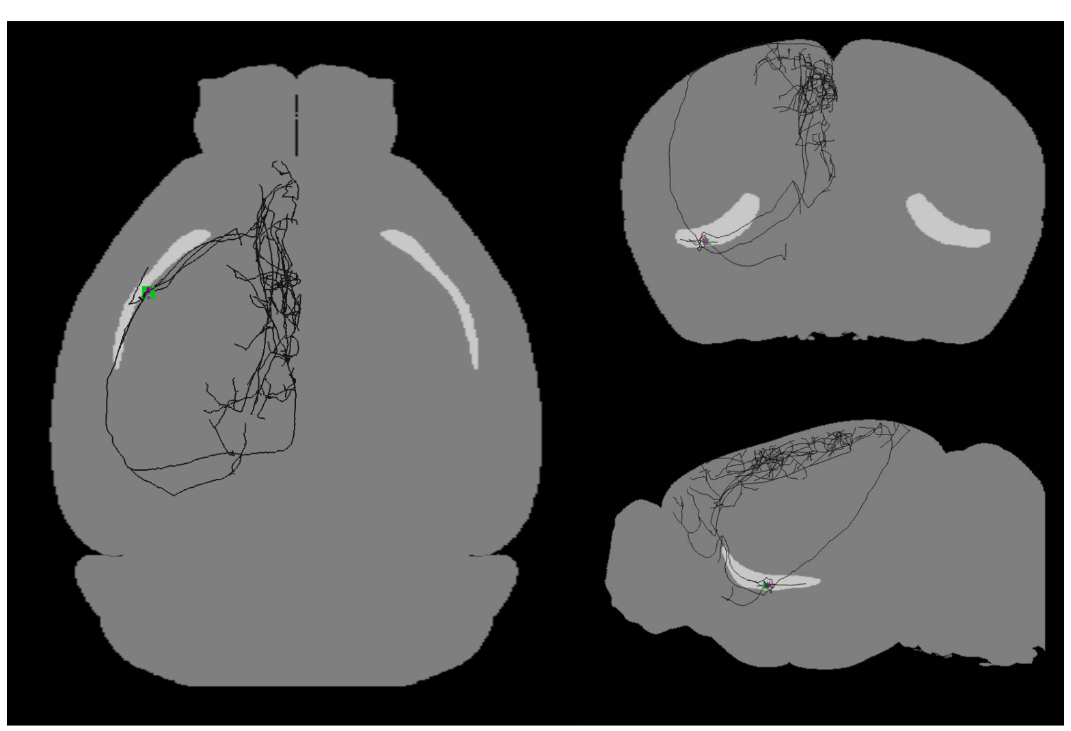

17781_04095_04095_X17570_Y12460_CLA_Ipsilateral_C3

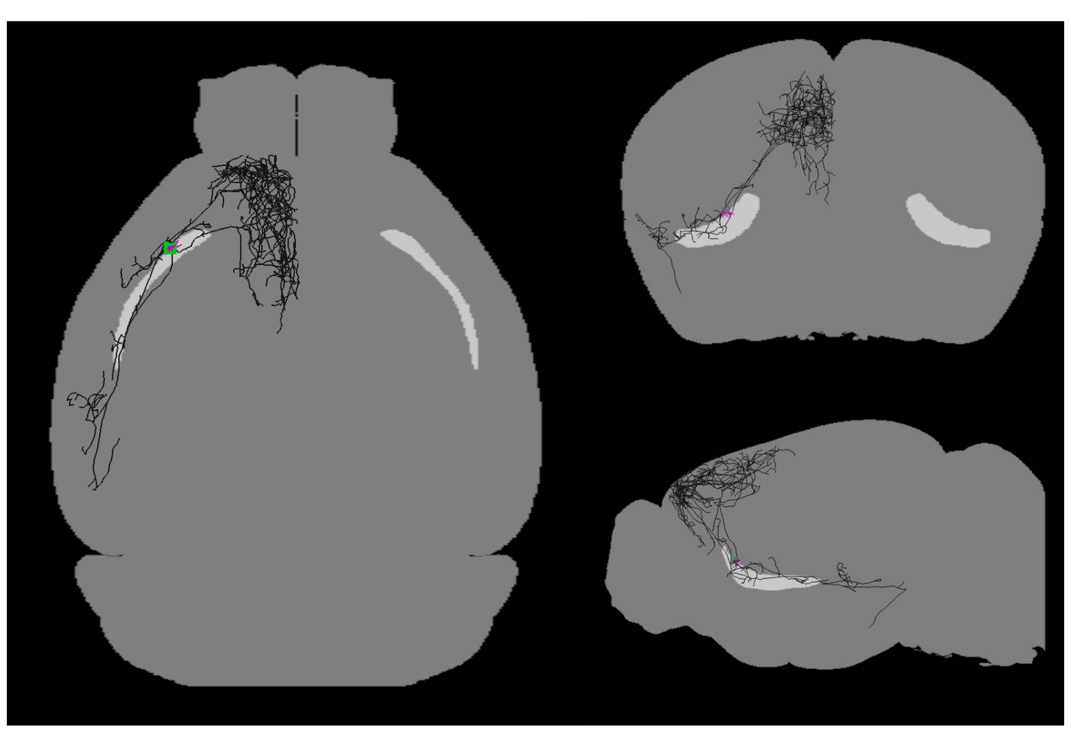

236174_03229_03328_X12413_Y11831_CLA_Ipsilateral_C3

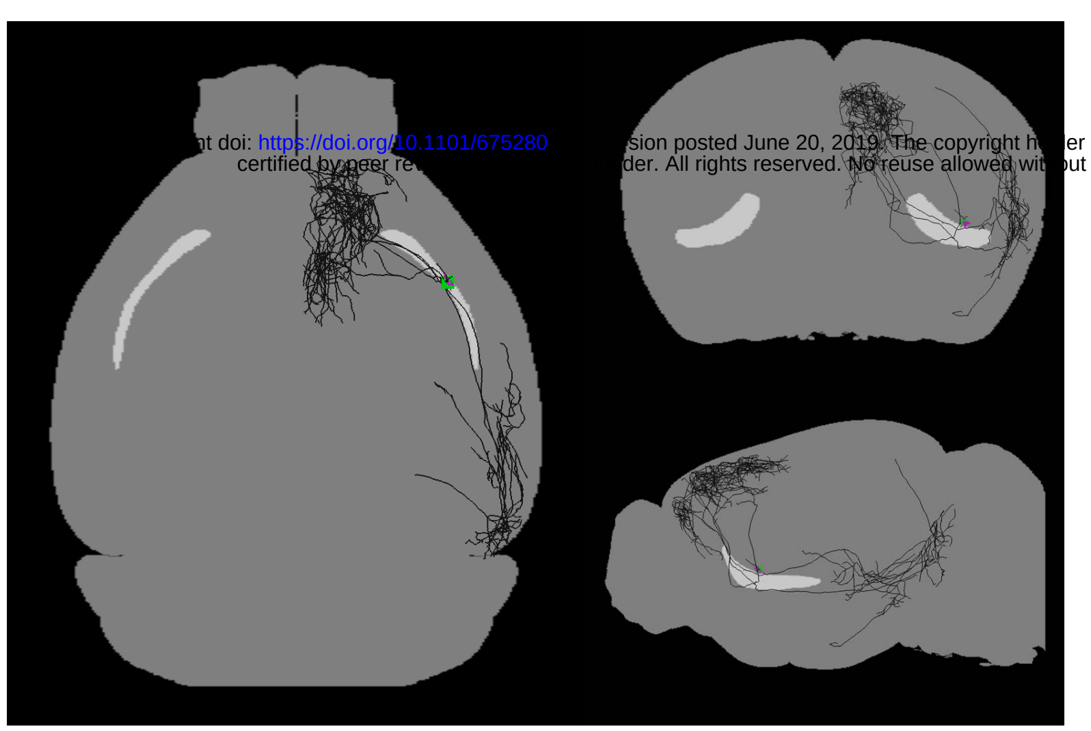

236174_03729_03828_X12692_Y09419_CLA_Ipsilateral_C3

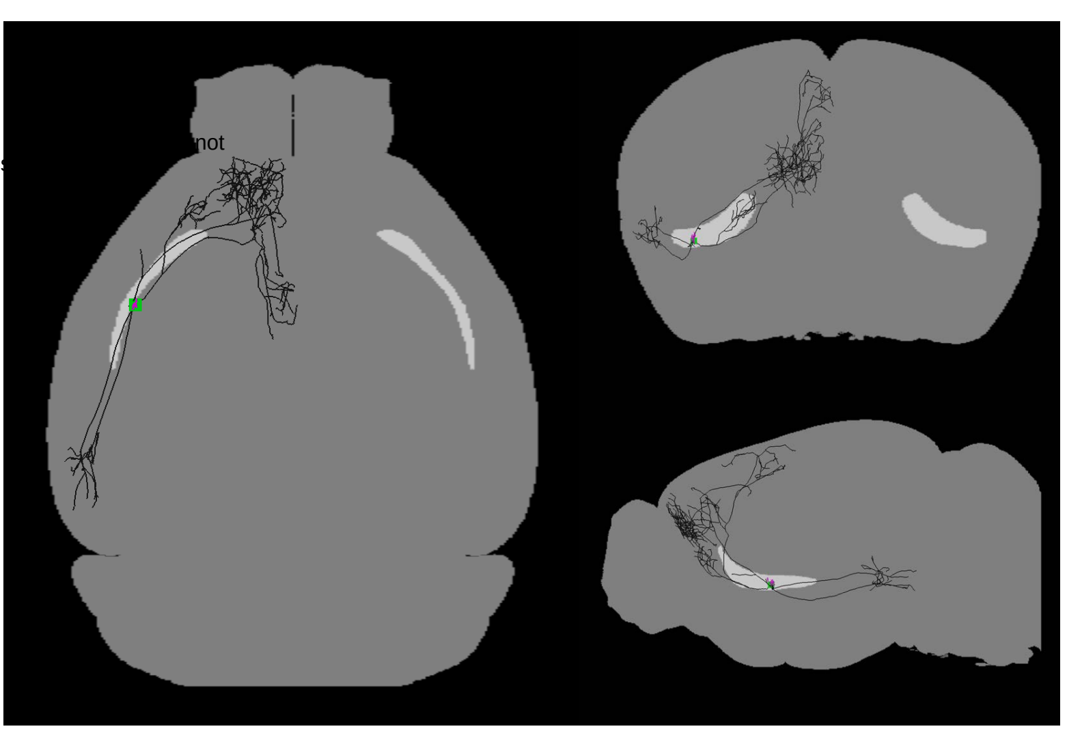

236174_05338_05437_X13590_Y07348_CLA_Ipsilateral_C3

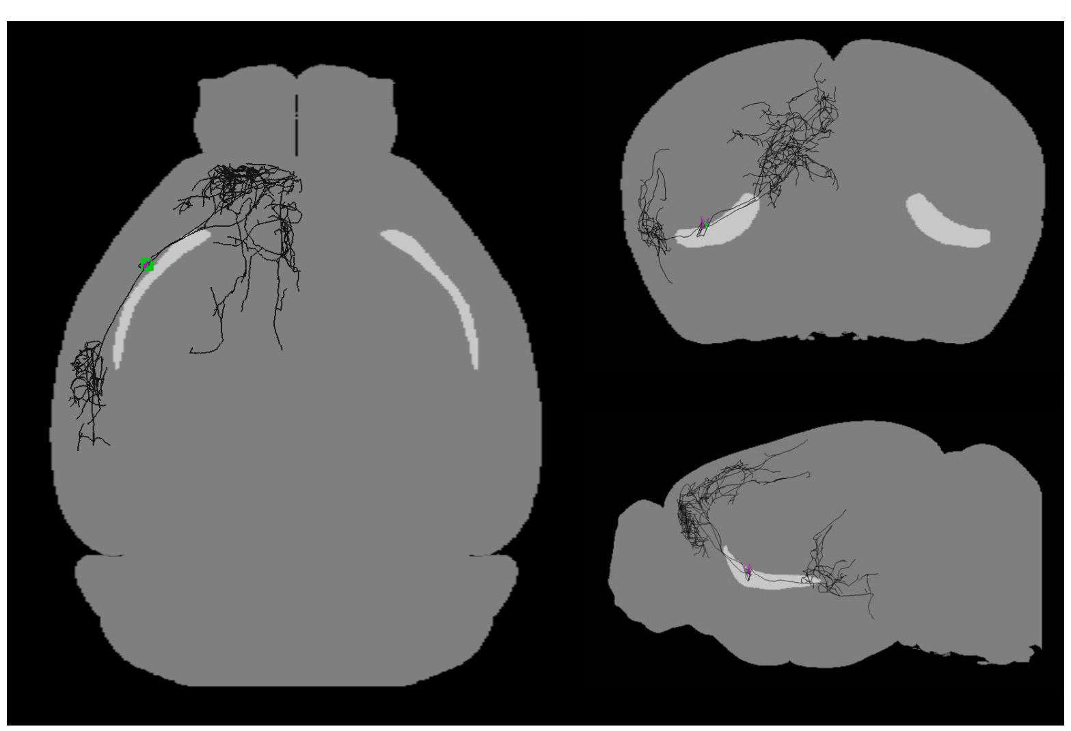

236174_03729_03828_X15443_Y26410_CLA_Ipsilateral_C3

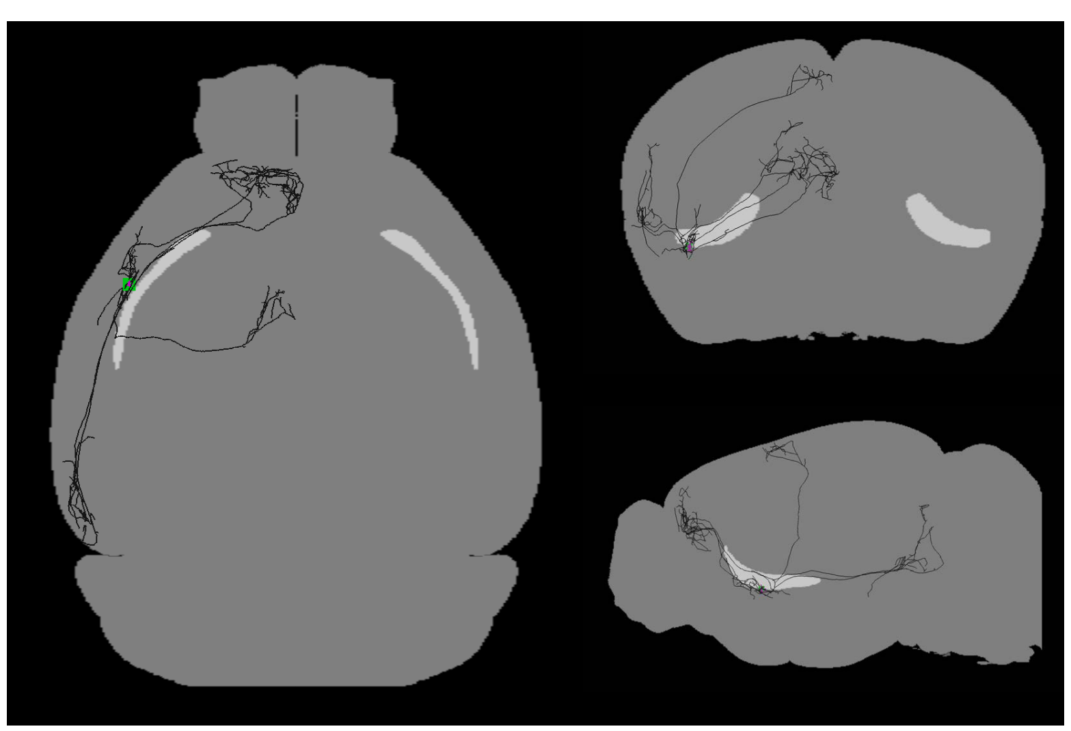

236174_03829_03928_X13590_Y09284_CLA_Ipsilateral_C3

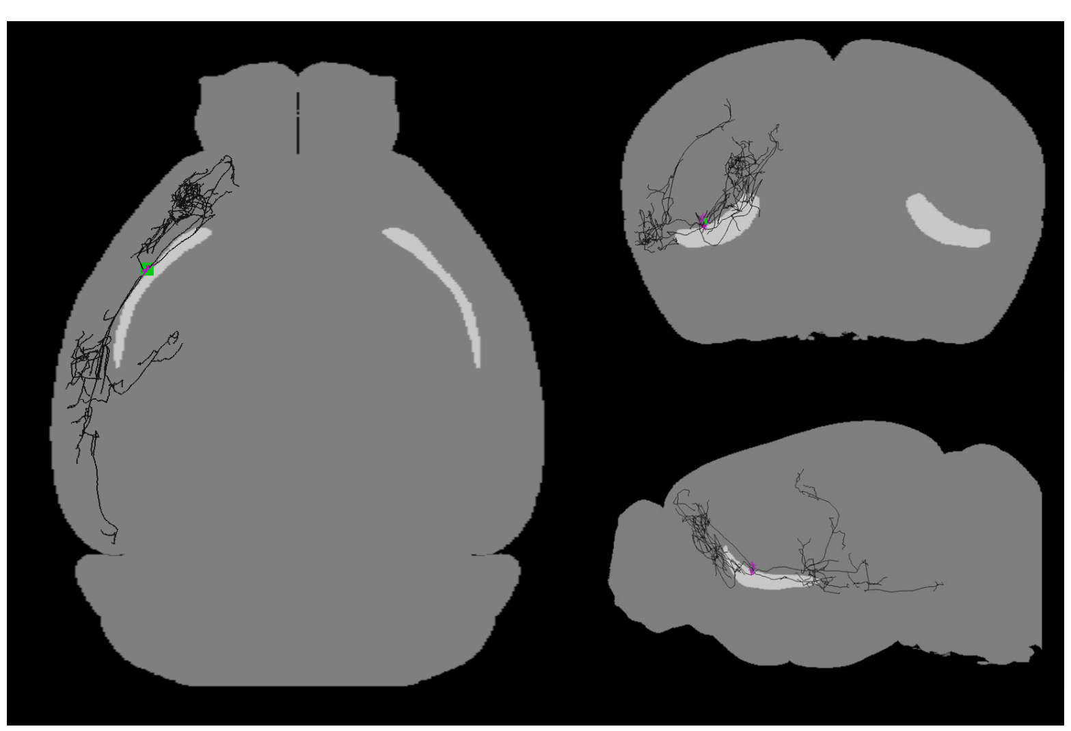

236174_03215_03215_X11999_Y11133_CLA_Ipsilateral_C8

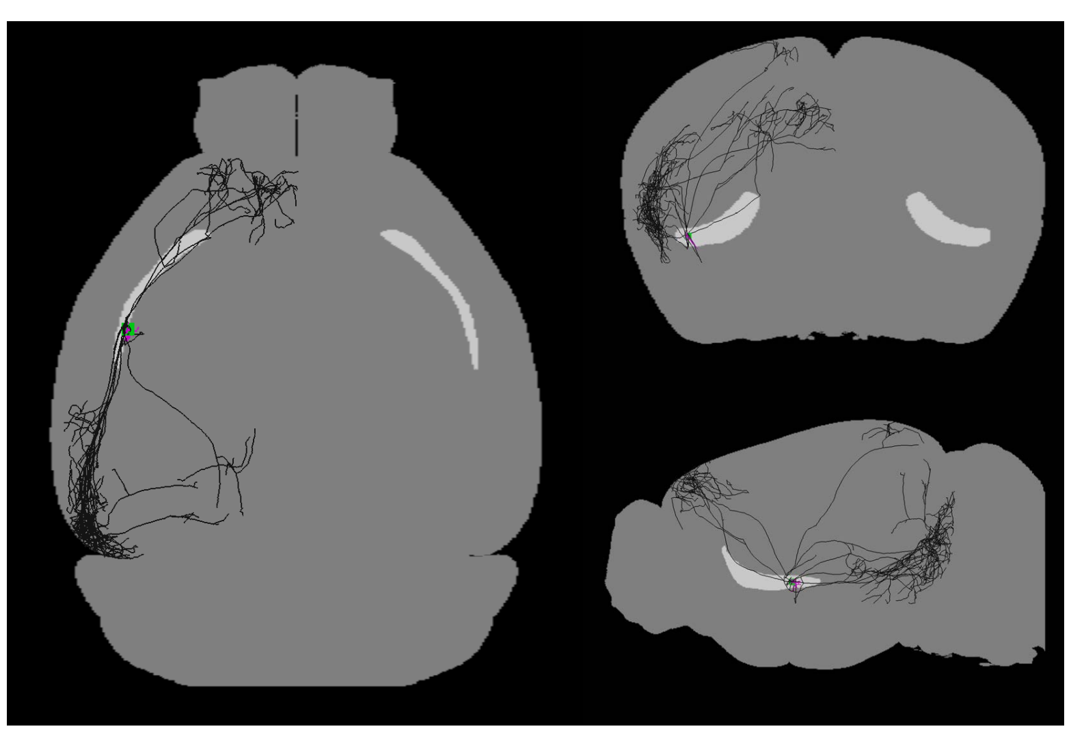

236174_03729_03828_X15151_Y26698_CLA_Ipsilateral_C8


236174_03970_04170_X13439_Y08678_CLA_Ipsilateral_C8


17781_03320_03320_X16423_Y14748_CLA_Ipsilateral_C9

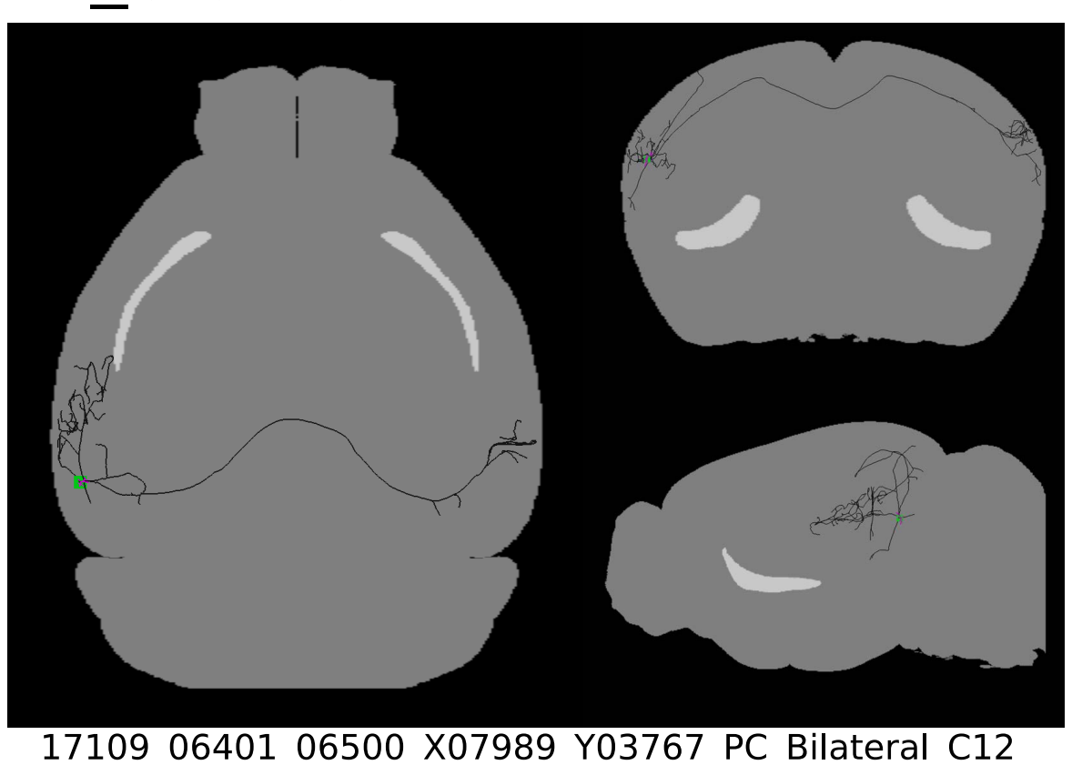

17781_03668_03668_X17825_Y13313_CLA_Ipsilateral_C9

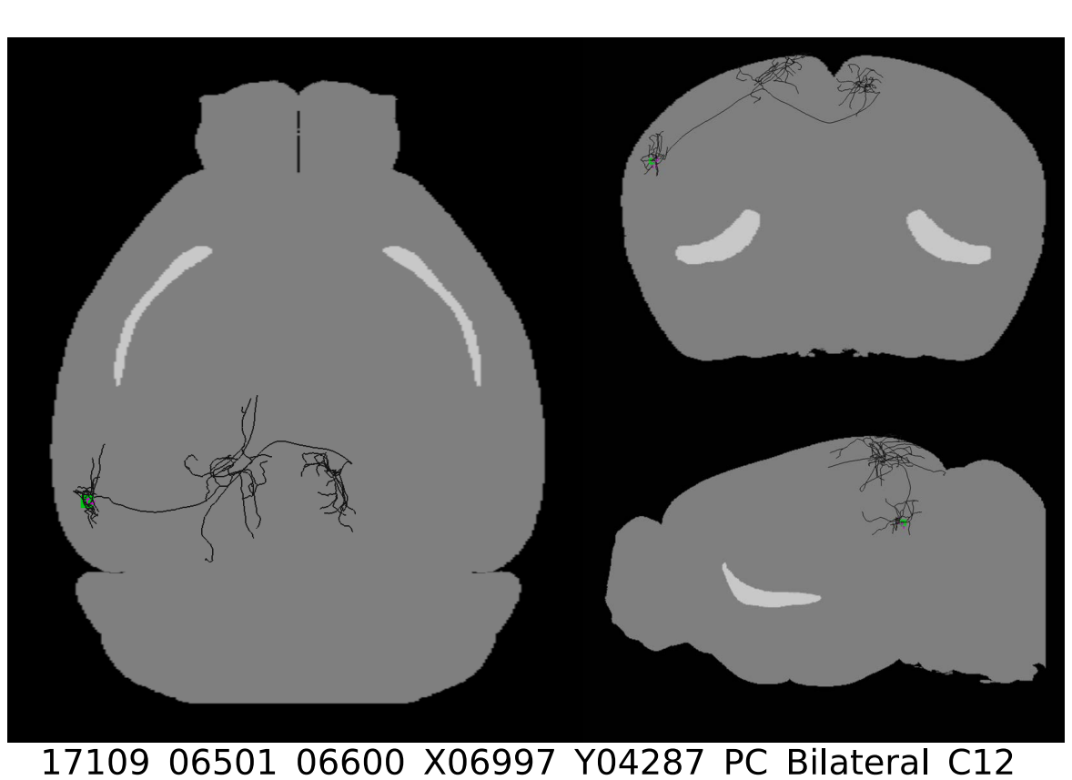

17782_03651_03651_X35286_Y18512_CLA_Ipsilateral_C9

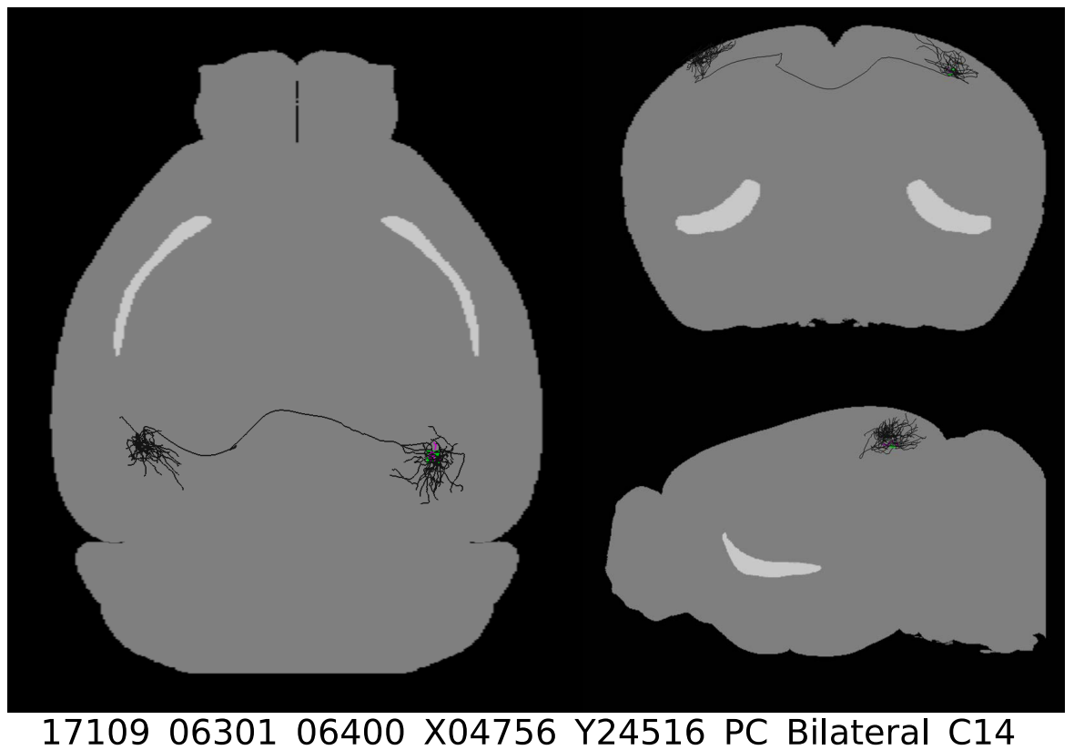

236174_04429_04528_X13147_Y08003_CLA_Ipsilateral_Cnan

PC_bilateral

Figure S5

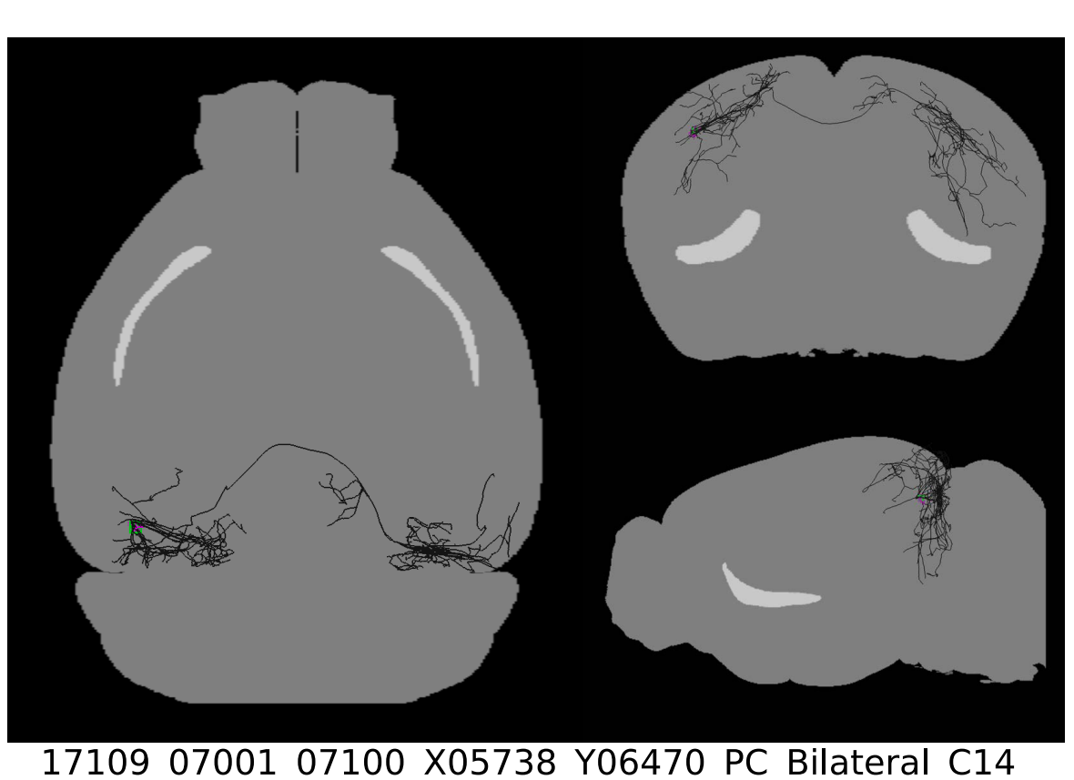

17109_06401_06500_X07989_Y03767_PC_Bilateral_C12

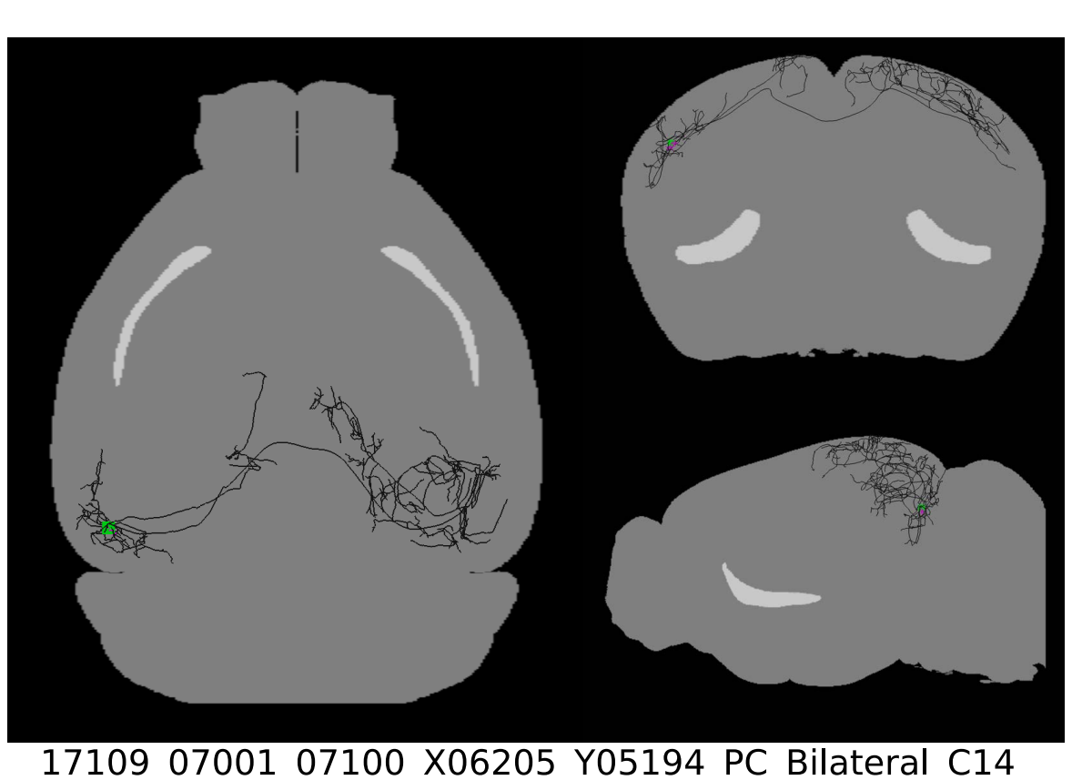

17109_06501_06600_X06997_Y04287_PC_Bilateral_C12

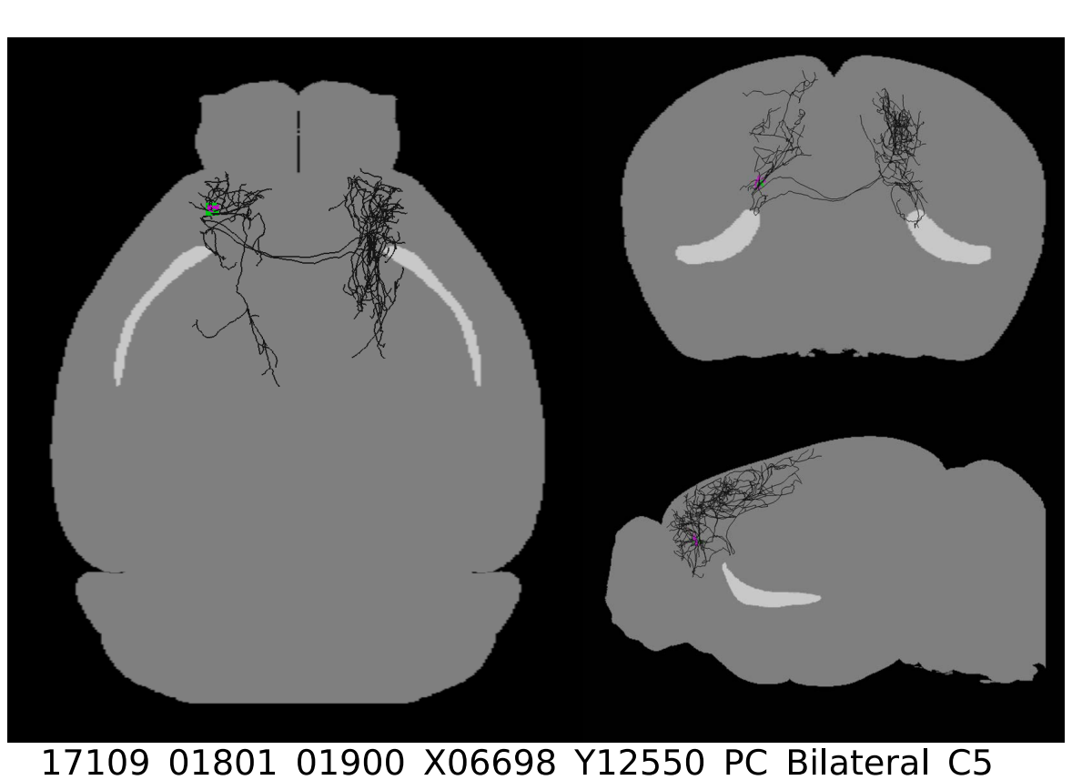

17109_06301_06400_X04756_Y24516_PC_Bilateral_C14


17109_06601_06700_X04384_Y07436_PC_Bilateral_C14

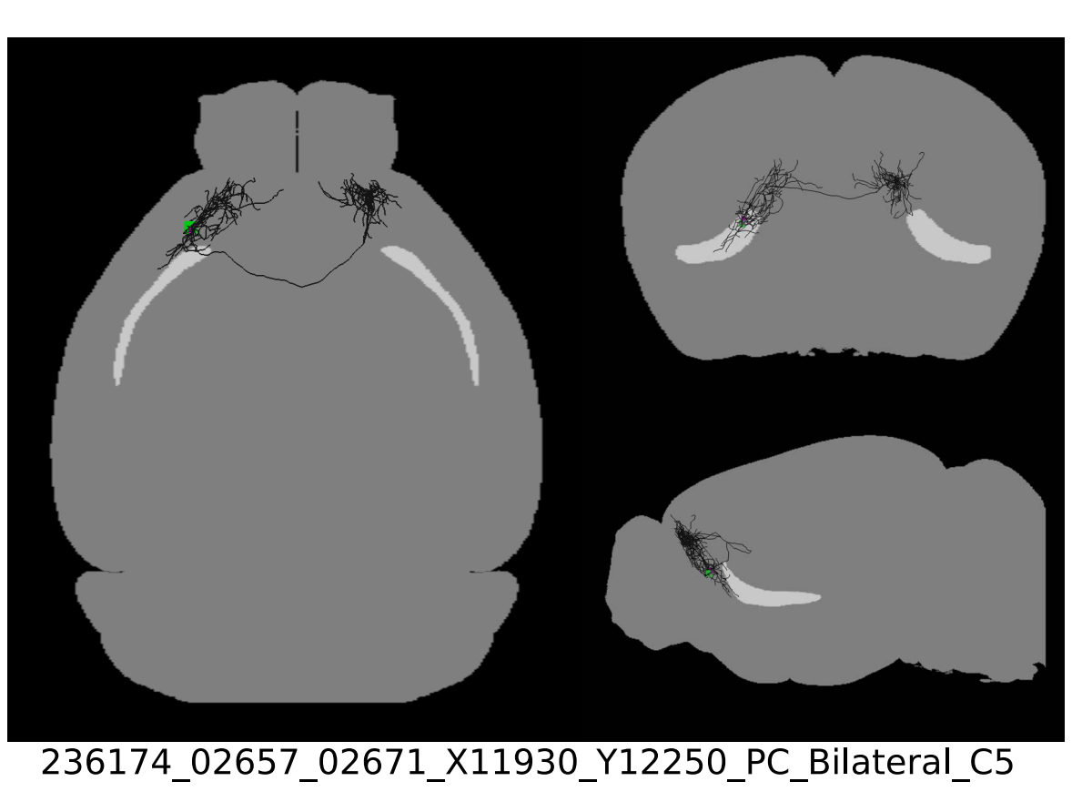

17109_06601_06700_X05417_Y25287_PC_Bilateral_C14

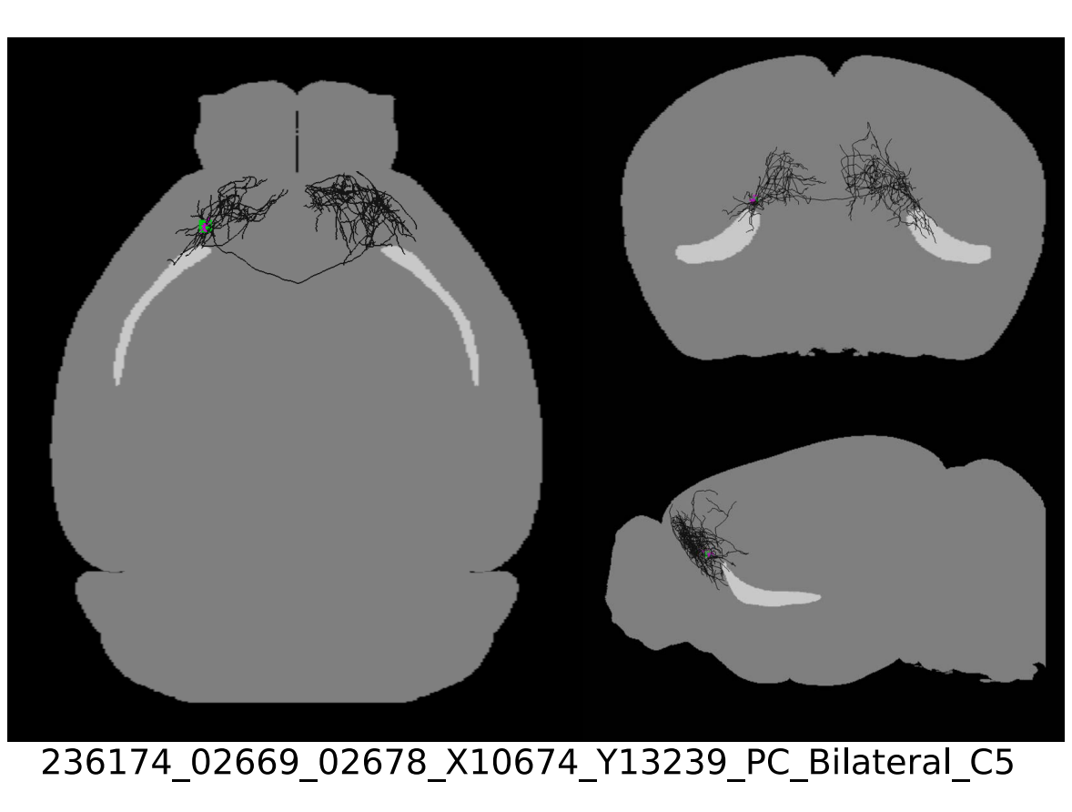

17109_07001_07100_X05738_Y06470_PC_Bilateral_C14

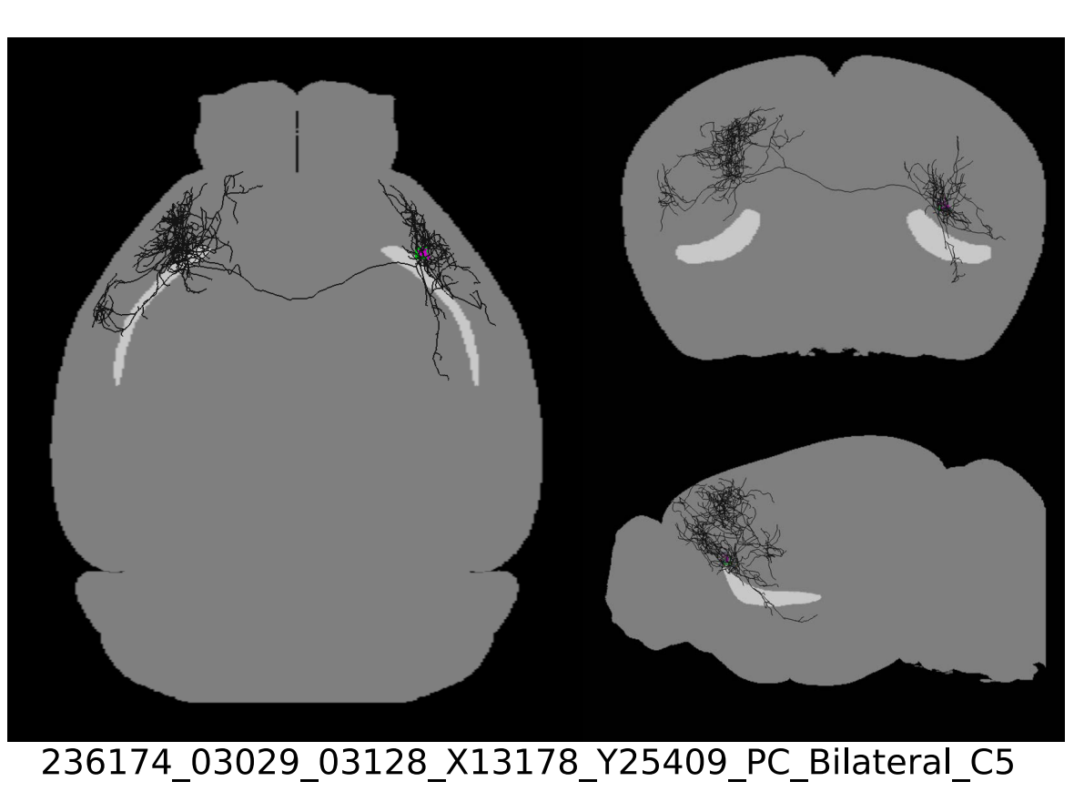

17109_07001_07100_X06205_Y05194_PC_Bilateral_C14

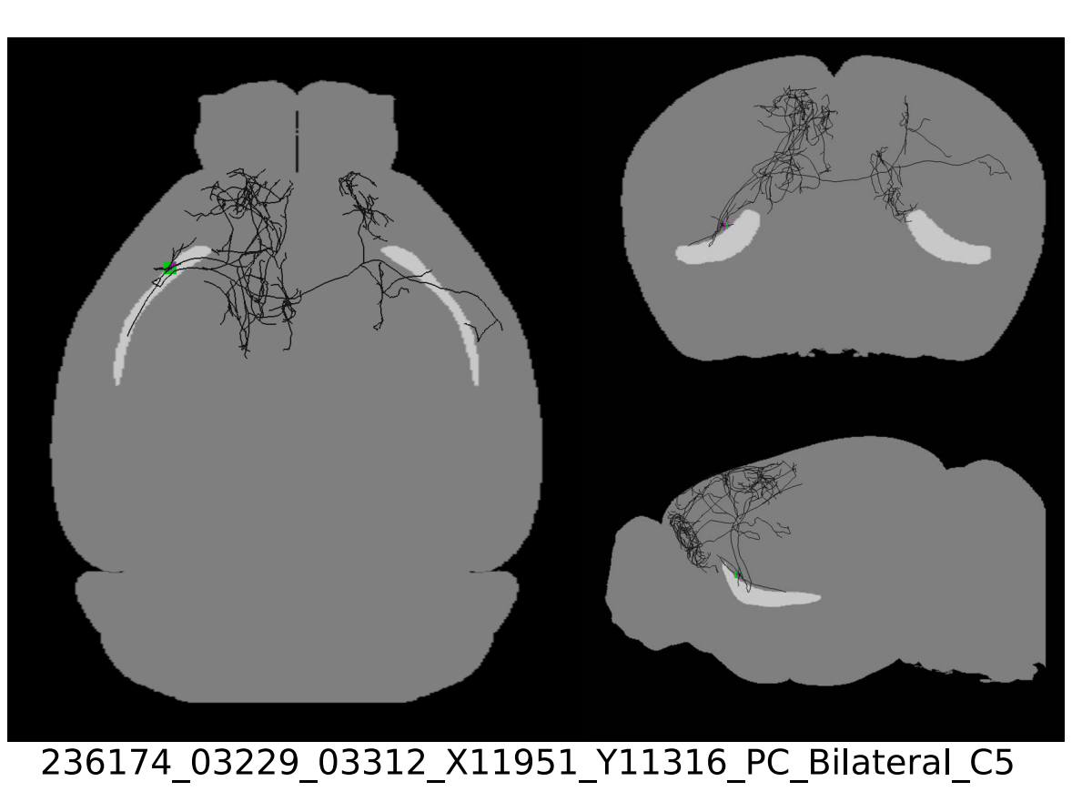

17109_01801_01900_X06698_Y12550_PC_Bilateral_C5


17782_03138_03138_X32805_Y18352_PC_Bilateral_C5

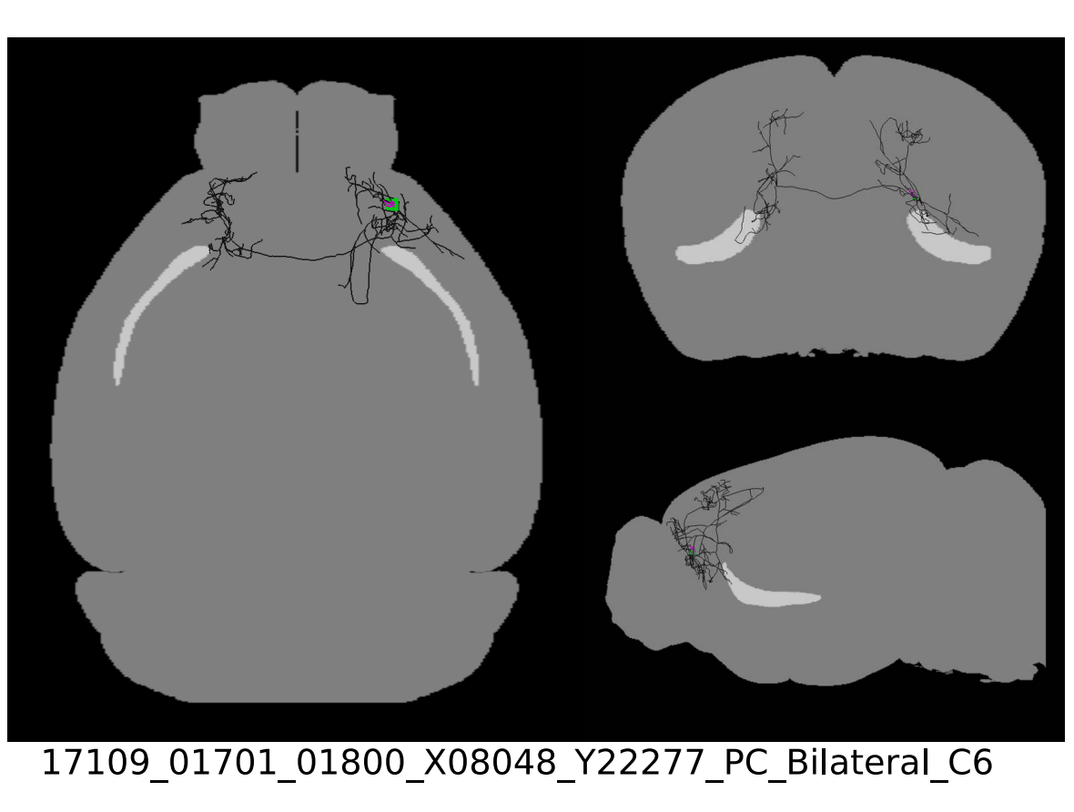

236174_02657_02671_X11930_Y12250_PC_Bilateral_C5

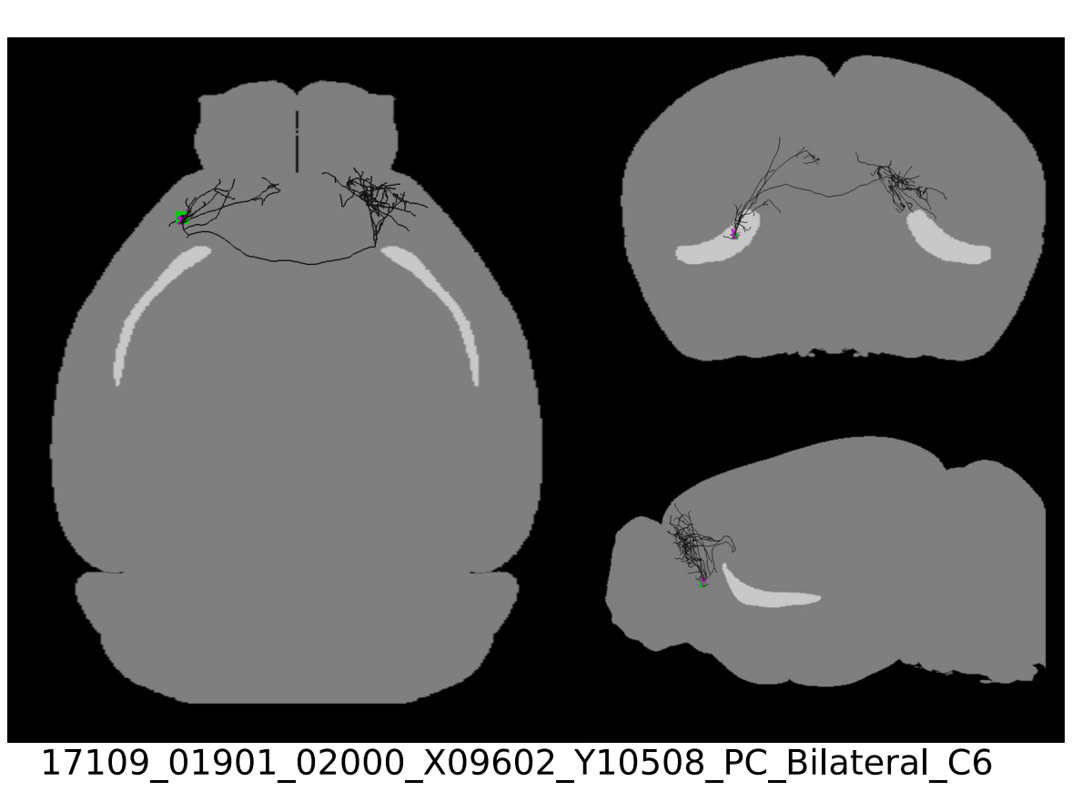

236174_02669_02678_X10674_Y13239_PC_Bilateral_C5

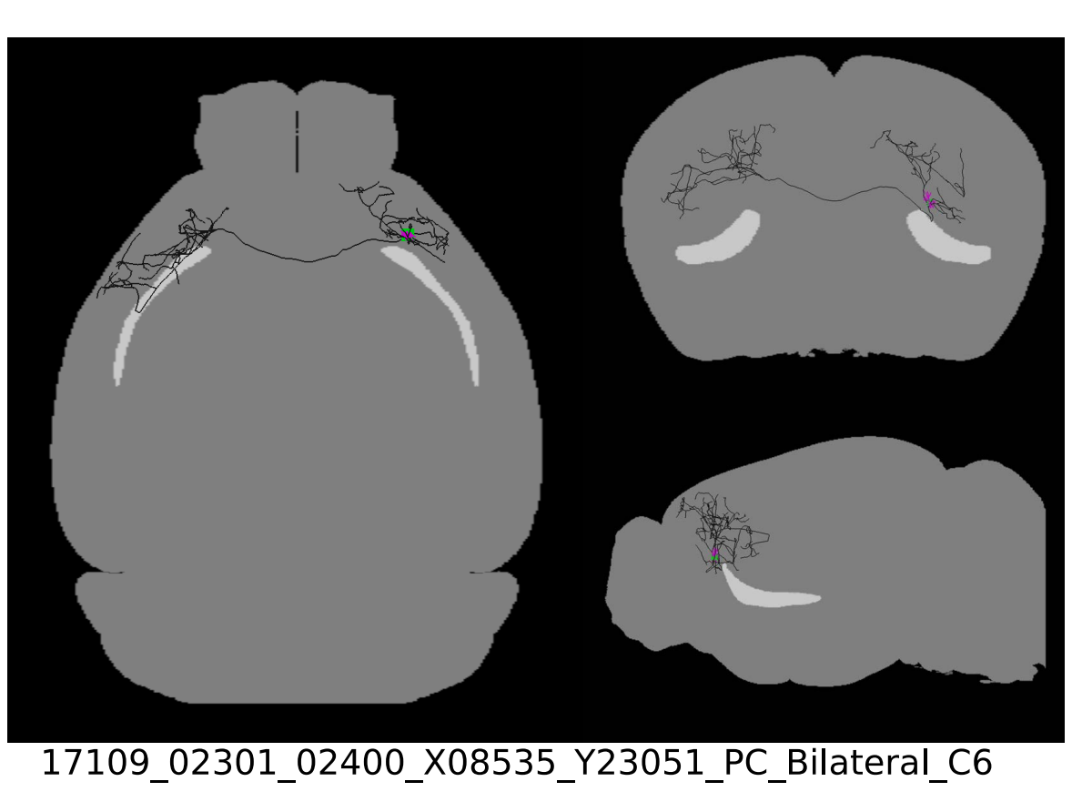

236174_03029_03128_X13178_Y25409_PC_Bilateral_C5

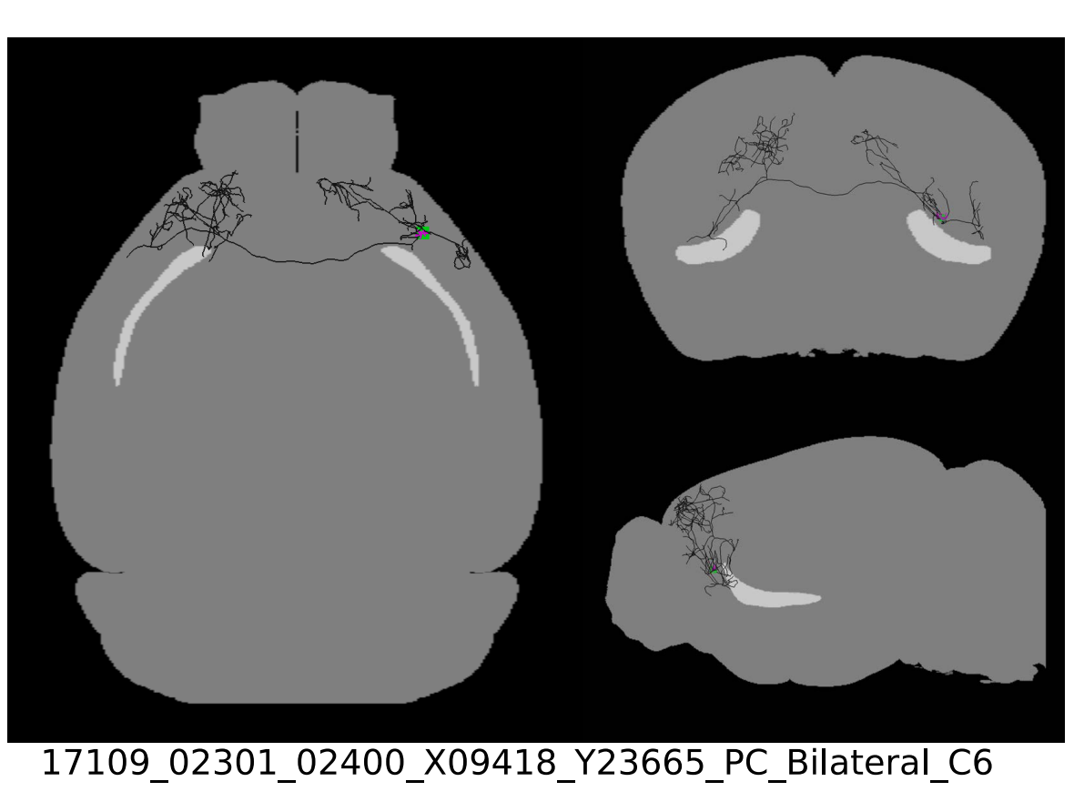

236174_03229_03312_X11951_Y11316_PC_Bilateral_C5

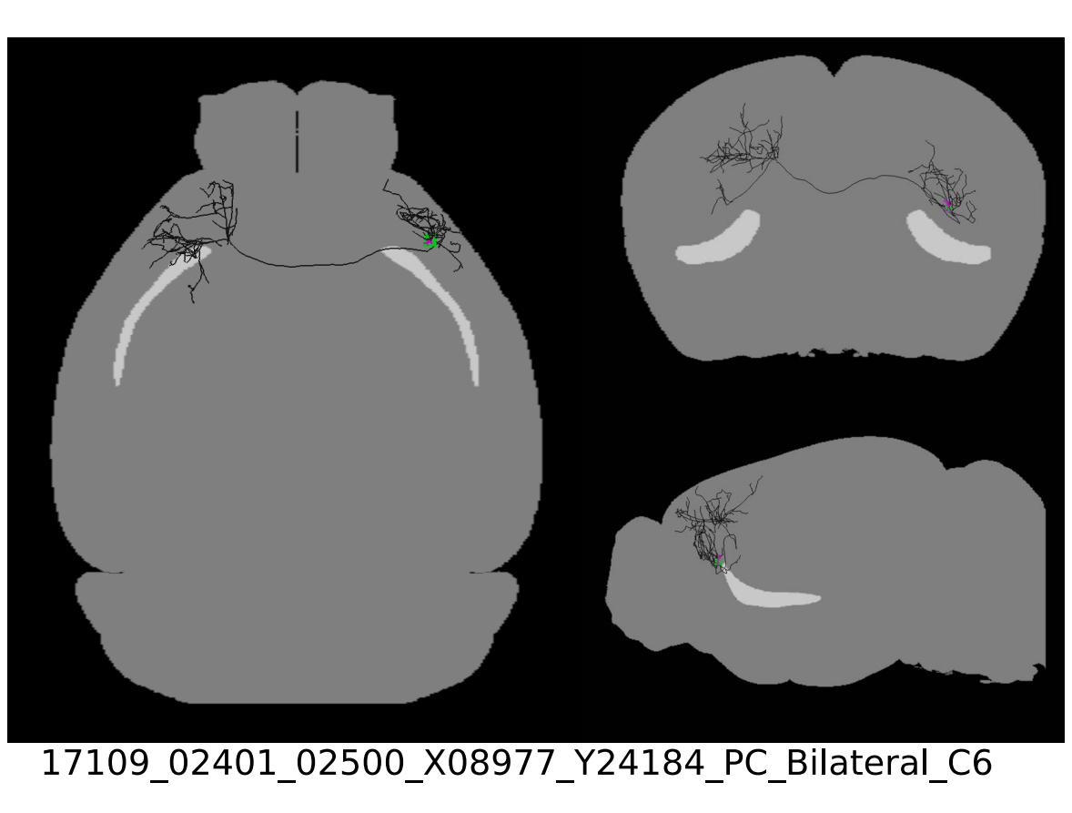

236174_03329_03428_X13938_Y26099_PC_Bilateral_C5


17109_01701_01800_X08048_Y22277_PC_Bilateral_C6

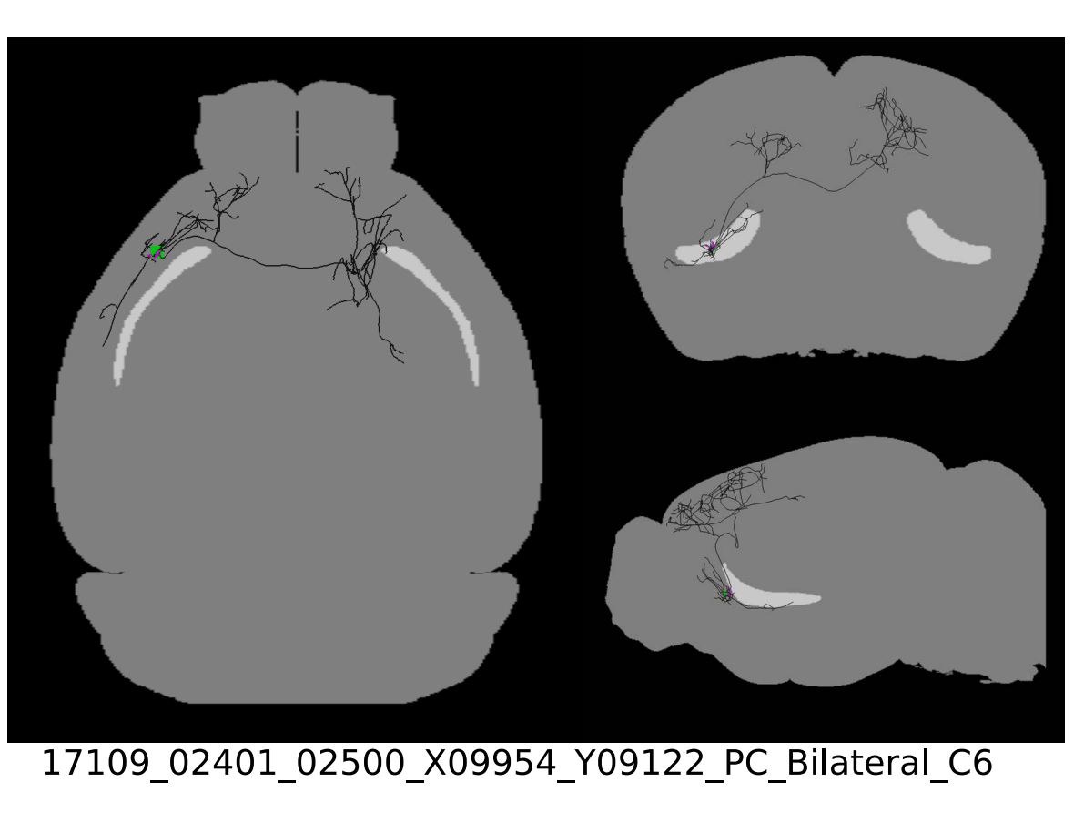

17109_01901_02000_X09602_Y10508_PC_Bilateral_C6

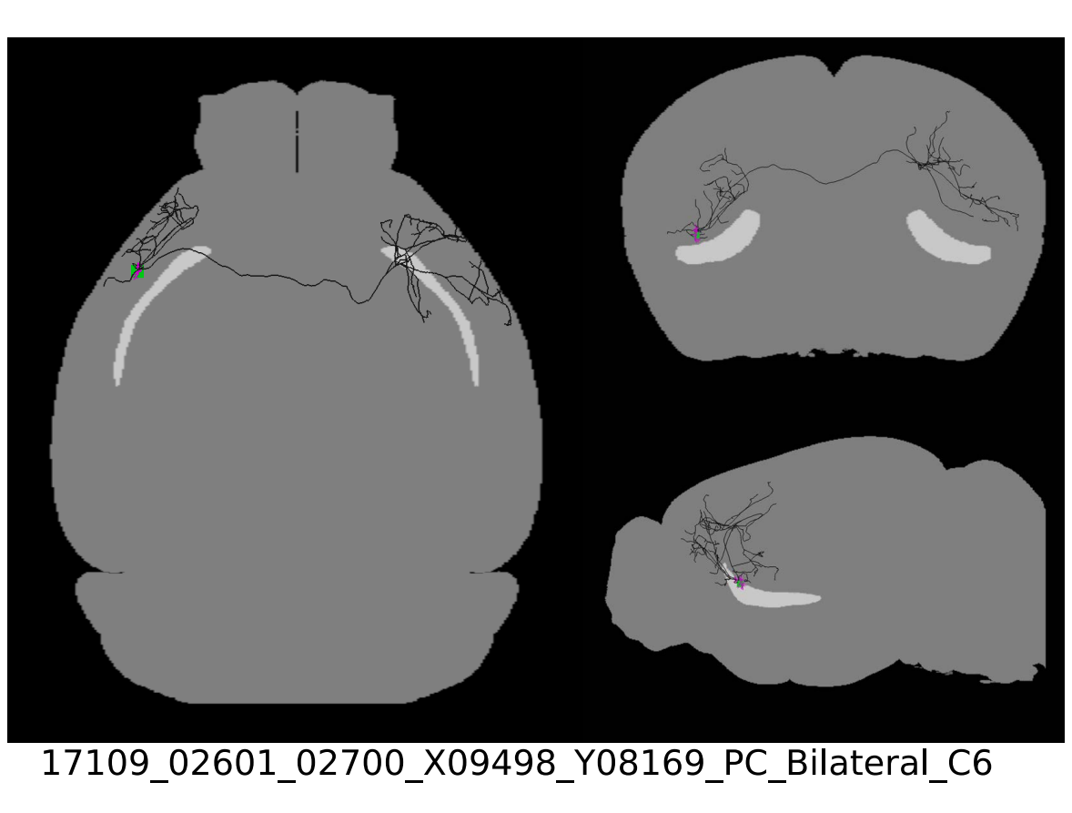

17109_02301_02400_X08535_Y23051_PC_Bilateral_C6

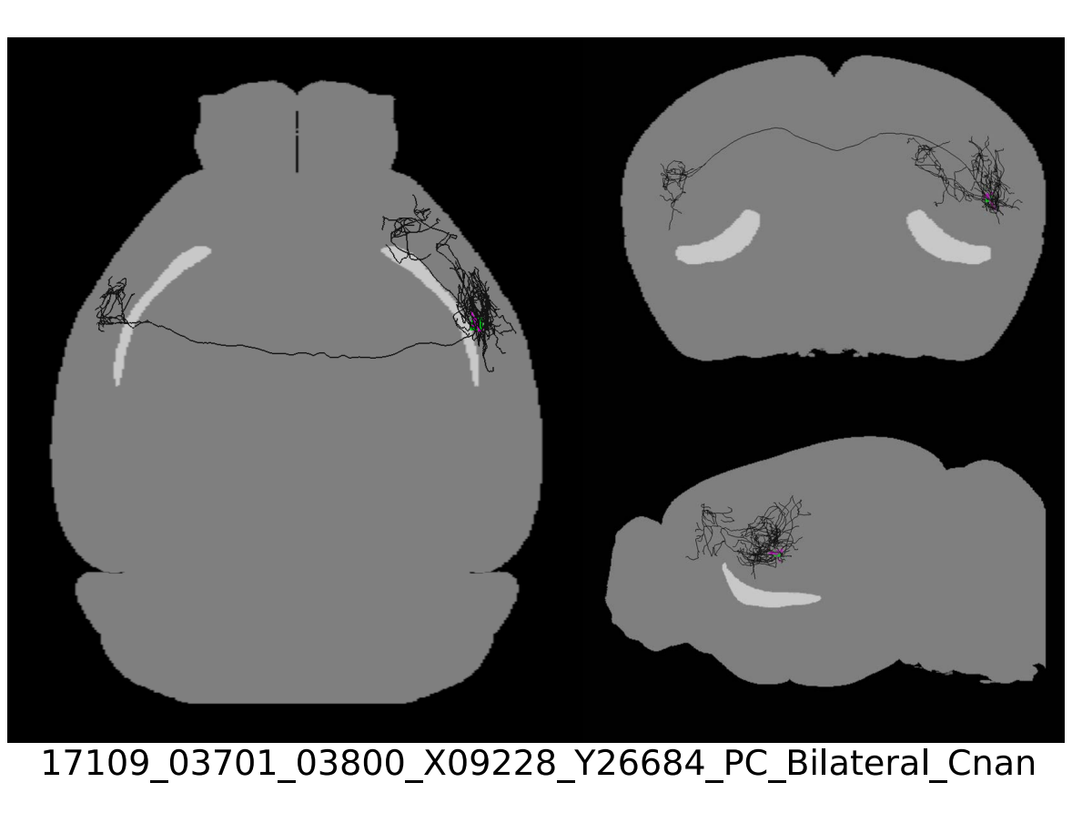

17109_02301_02400_X09418_Y23665_PC_Bilateral_C6


17109_02401_02500_X08977_Y24184_PC_Bilateral_C6

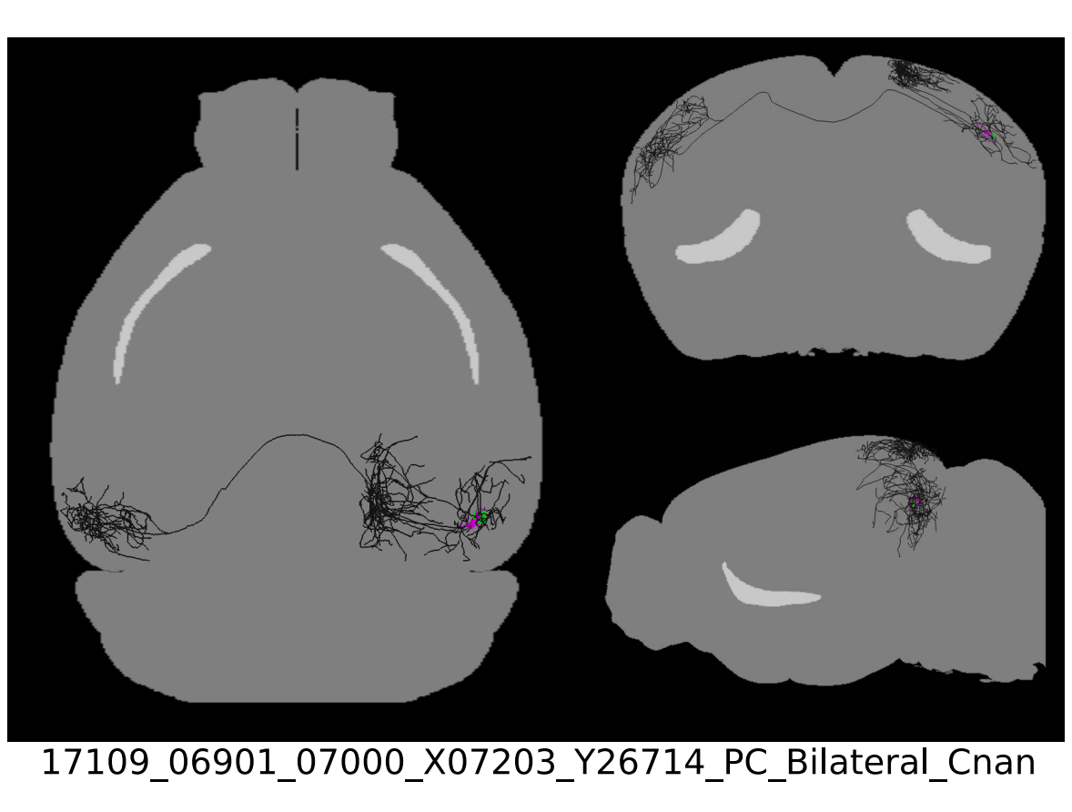

17109_02401_02500_X09338_Y02394_PC_Bilateral_C6

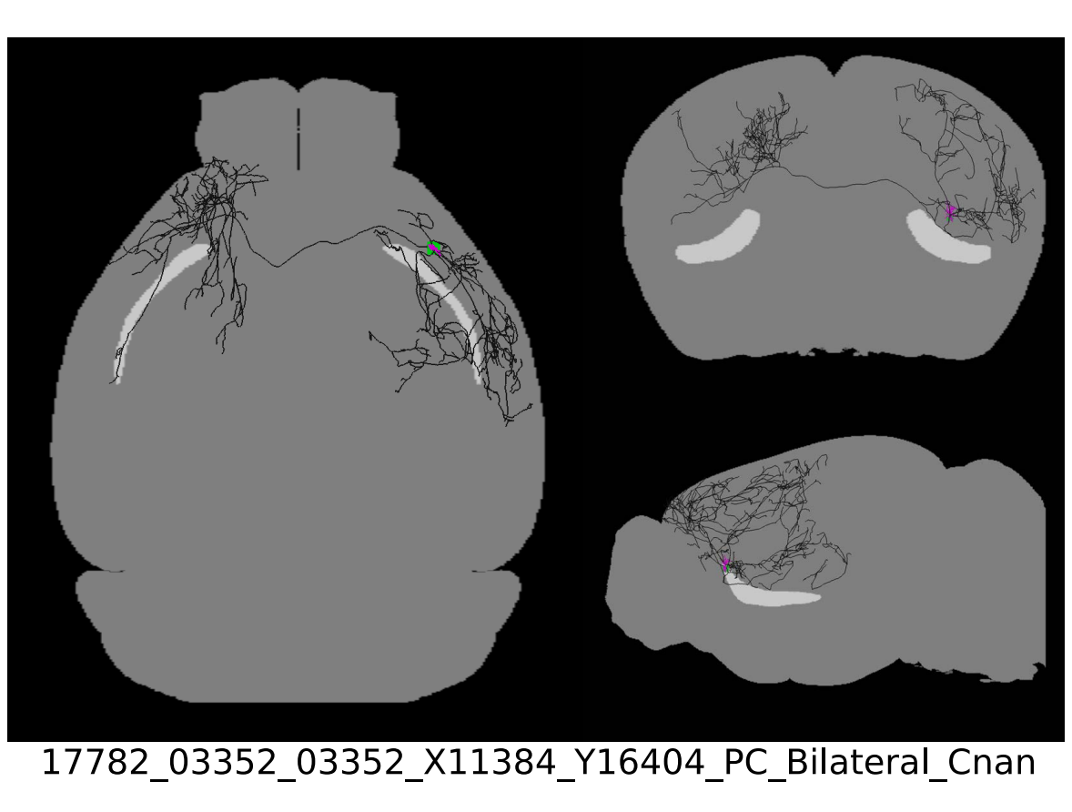

17109_02401_02500_X09954_Y09122_PC_Bilateral_C6


17109_02601_02700_X09498_Y08169_PC_Bilateral_C6


17109_03701_03800_X09228_Y26684_PC_Bilateral_Cnan

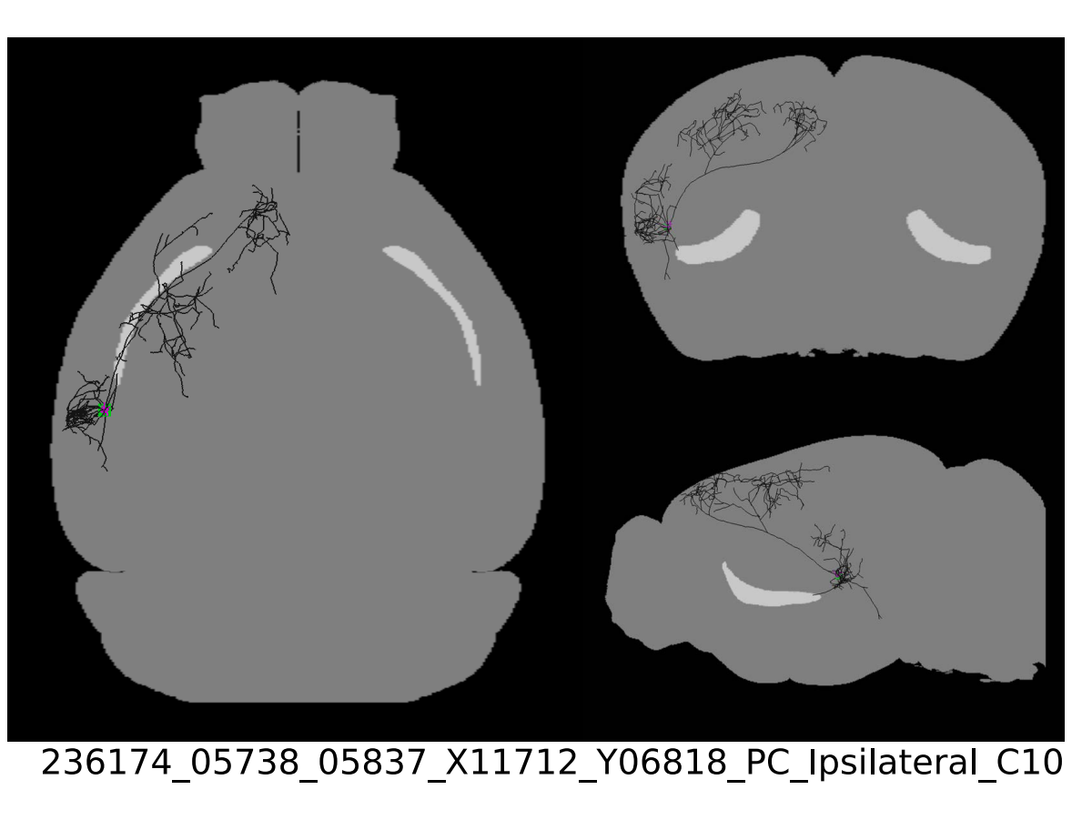

17109_04101_04200_X06753_Y06197_PC_Bilateral_Cnan

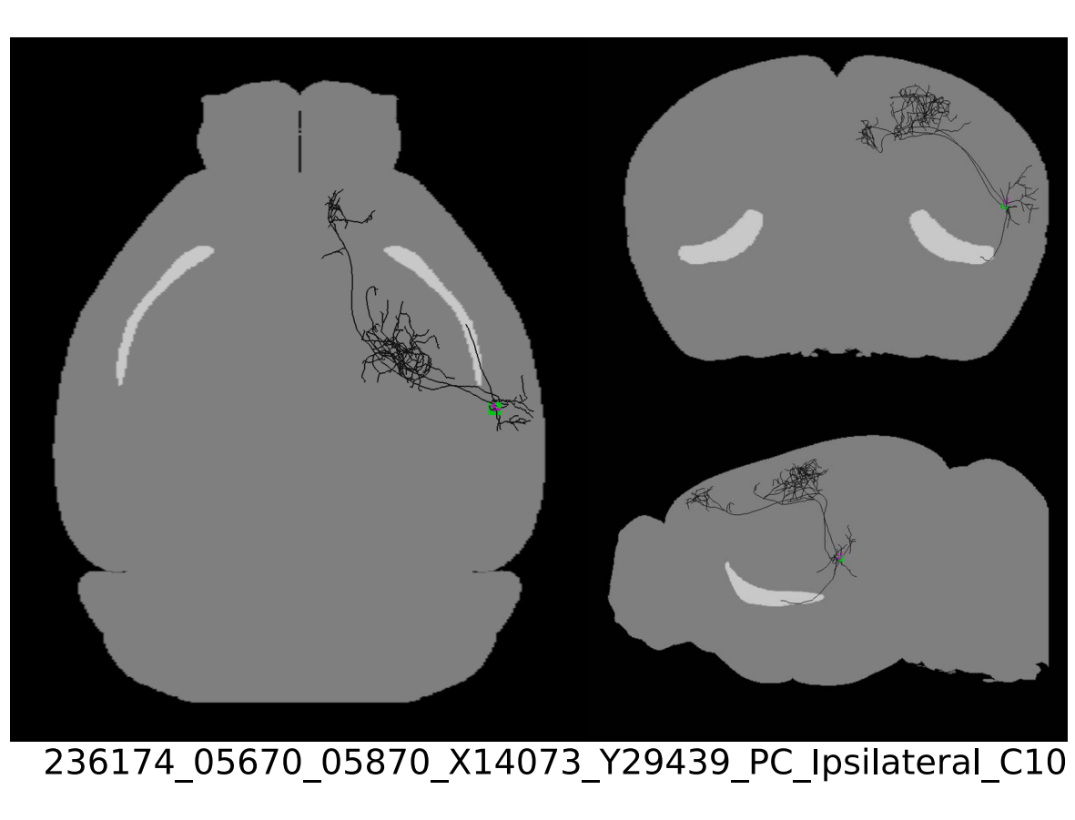

17109_06901_07000_X07203_Y26714_PC_Bilateral_Cnan

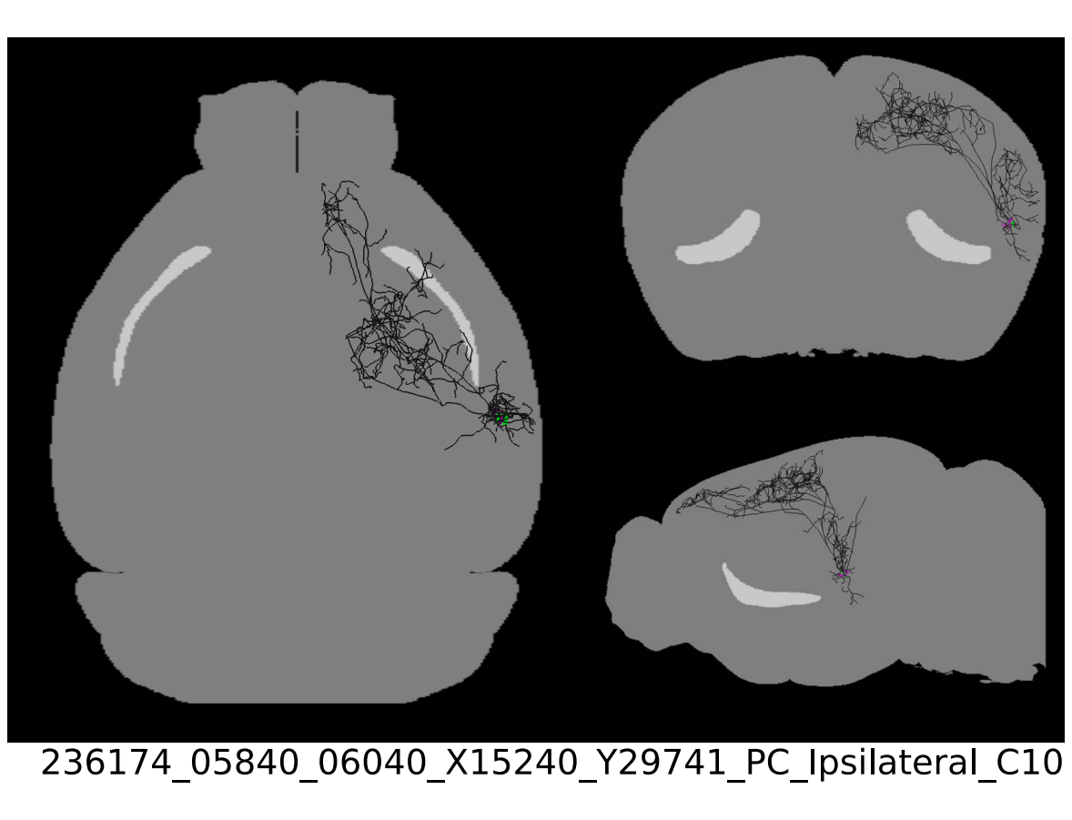

17782_03352_03352_X11384_Y16404_PC_Bilateral_Cnan

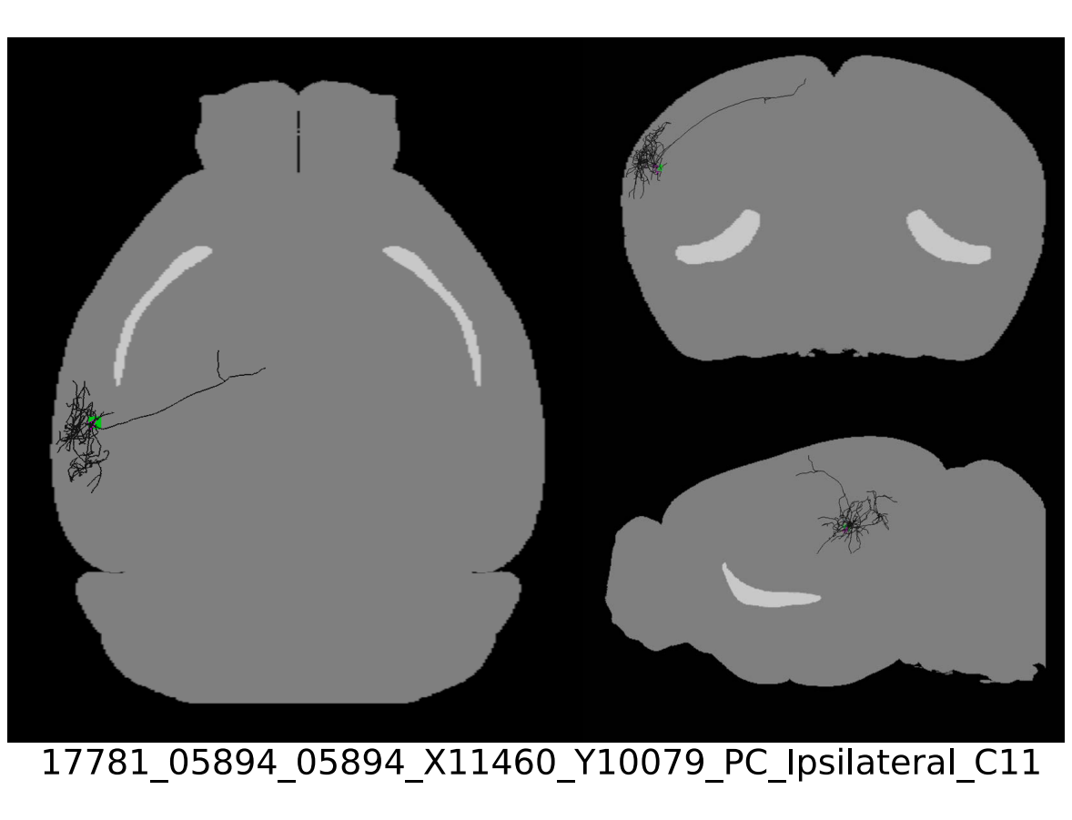
PC_ipsilateral

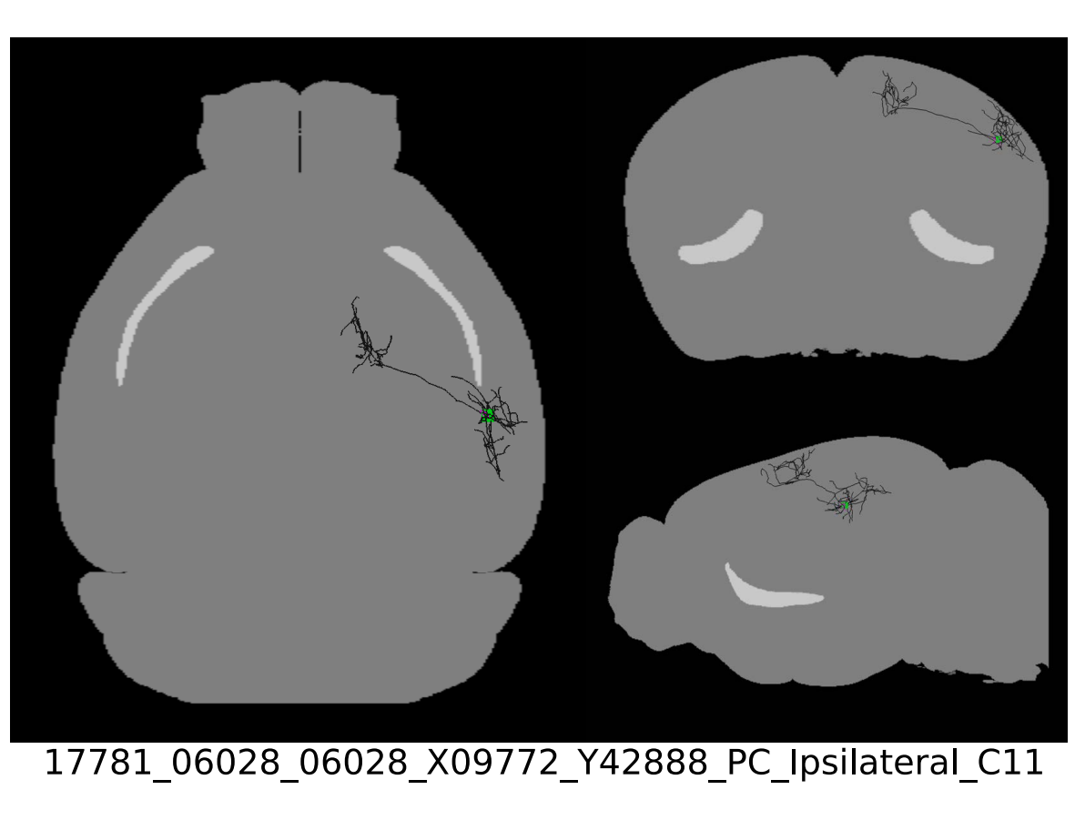

17781_05655_05655_X10641_Y11102_PC_Ipsilateral_C10

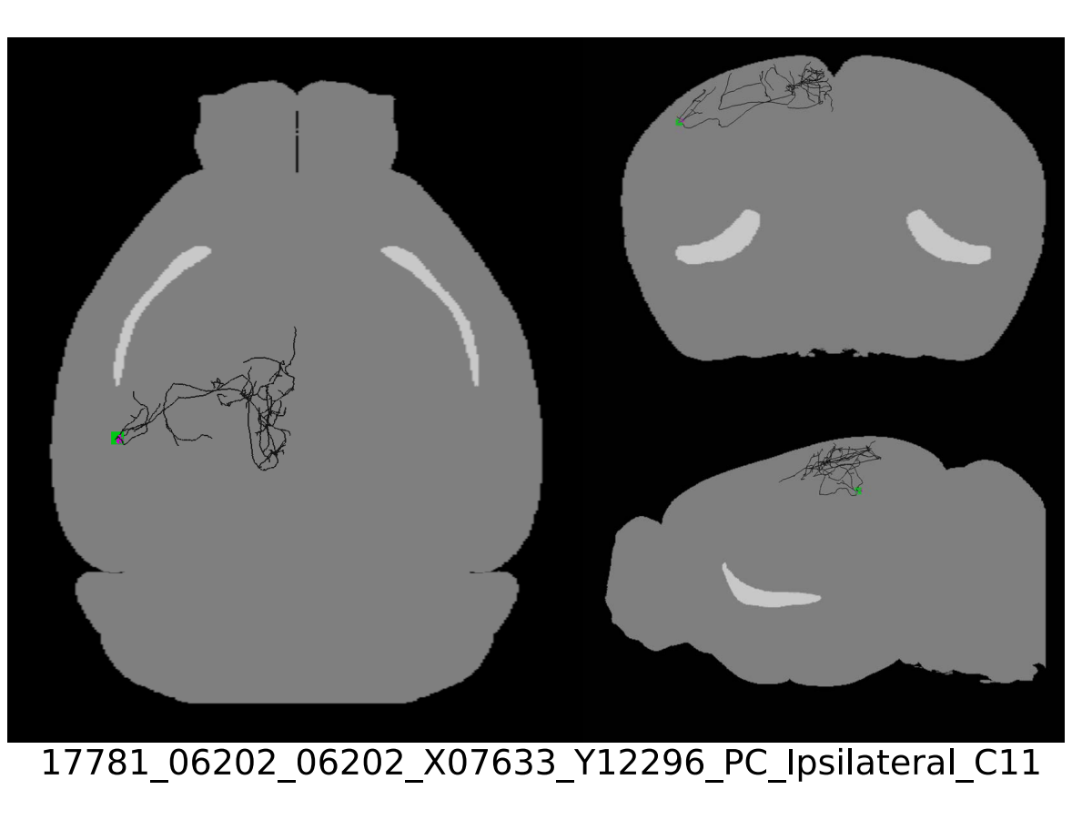

236174_05738_05837_X11712_Y06818_PC_Ipsilateral_C10

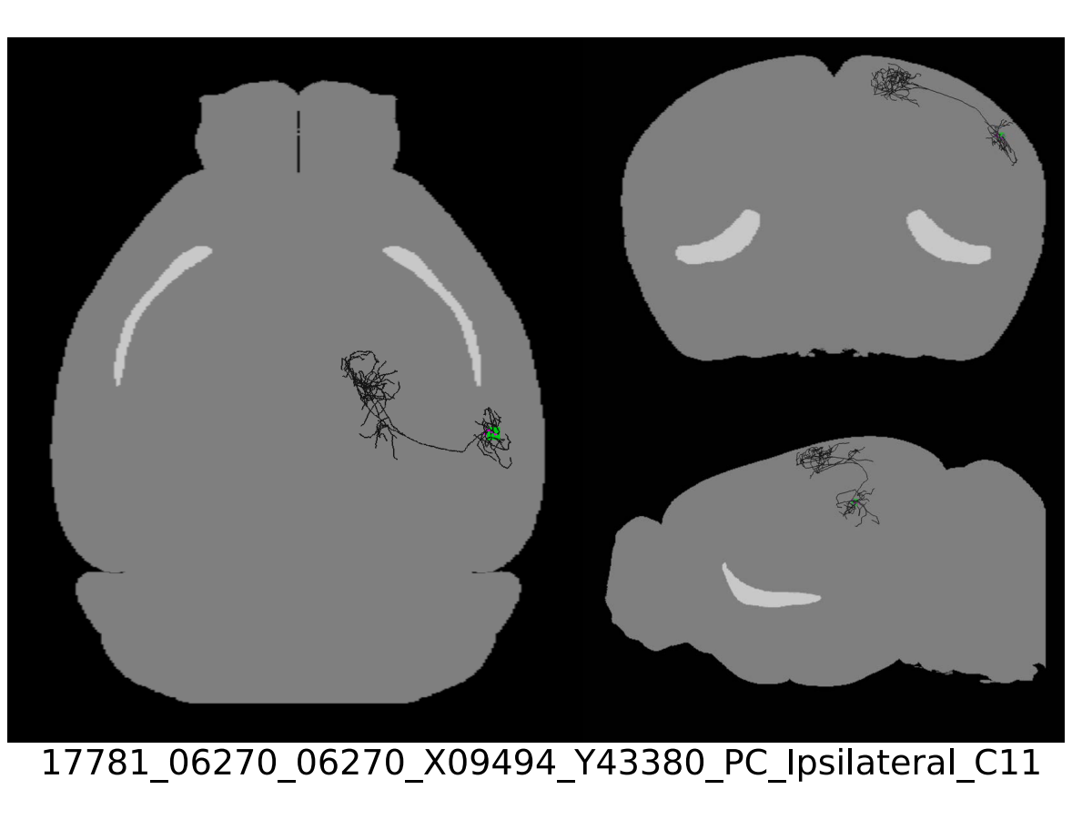

236174_05670_05870_X14073_Y29439_PC_Ipsilateral_C10


236174_05840_06040_X15240_Y29741_PC_Ipsilateral_C10


17781_05894_05894_X11460_Y10079_PC_Ipsilateral_C11

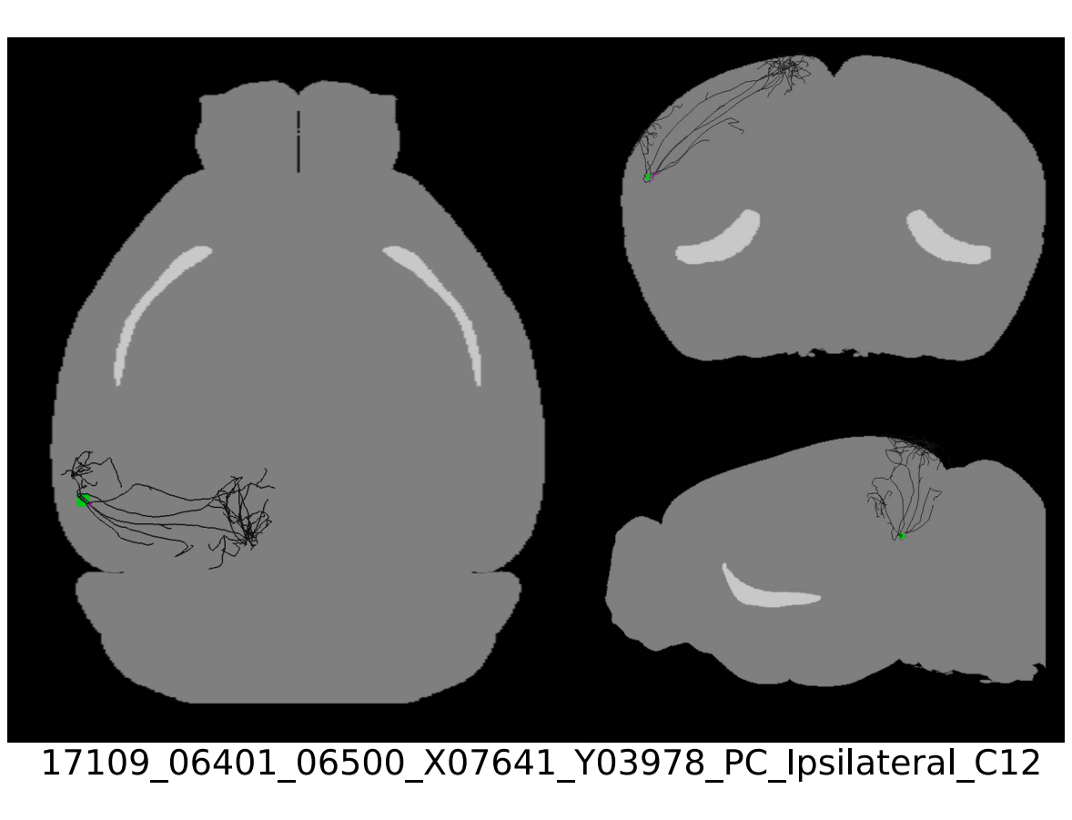

17781_06028_06028_X09772_Y42888_PC_Ipsilateral_C11

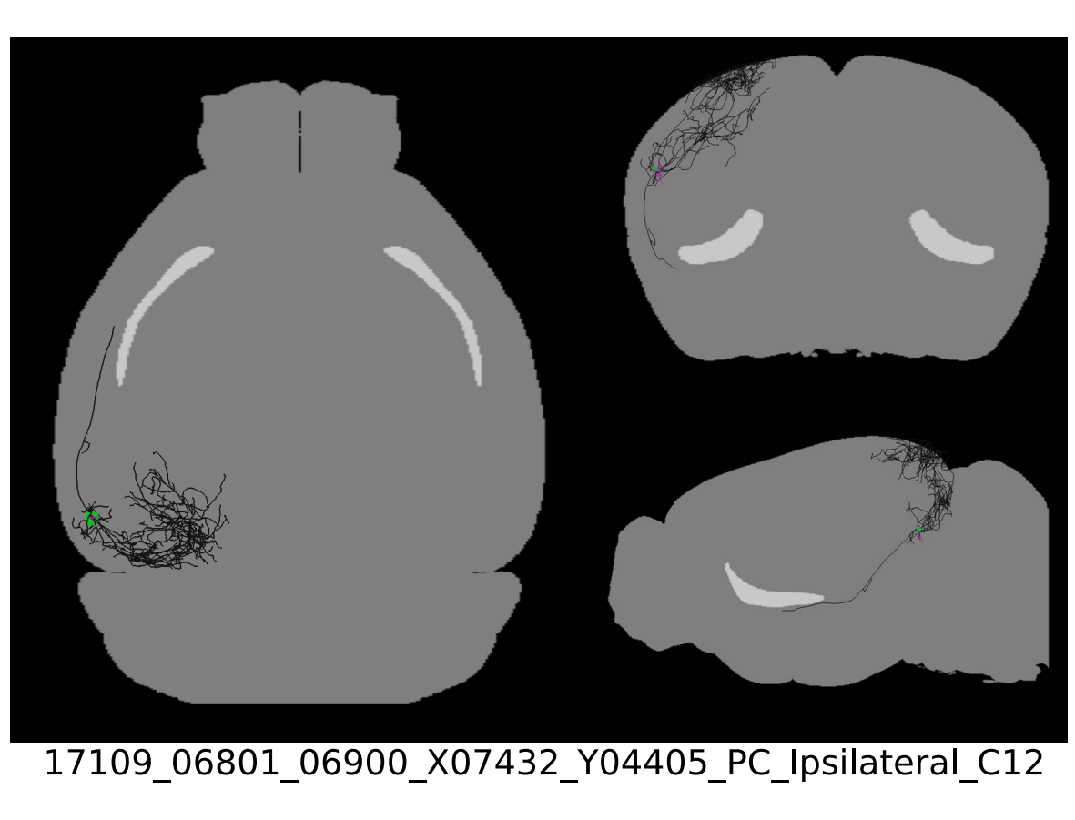

17781_06202_06202_X07633_Y12296_PC_Ipsilateral_C11

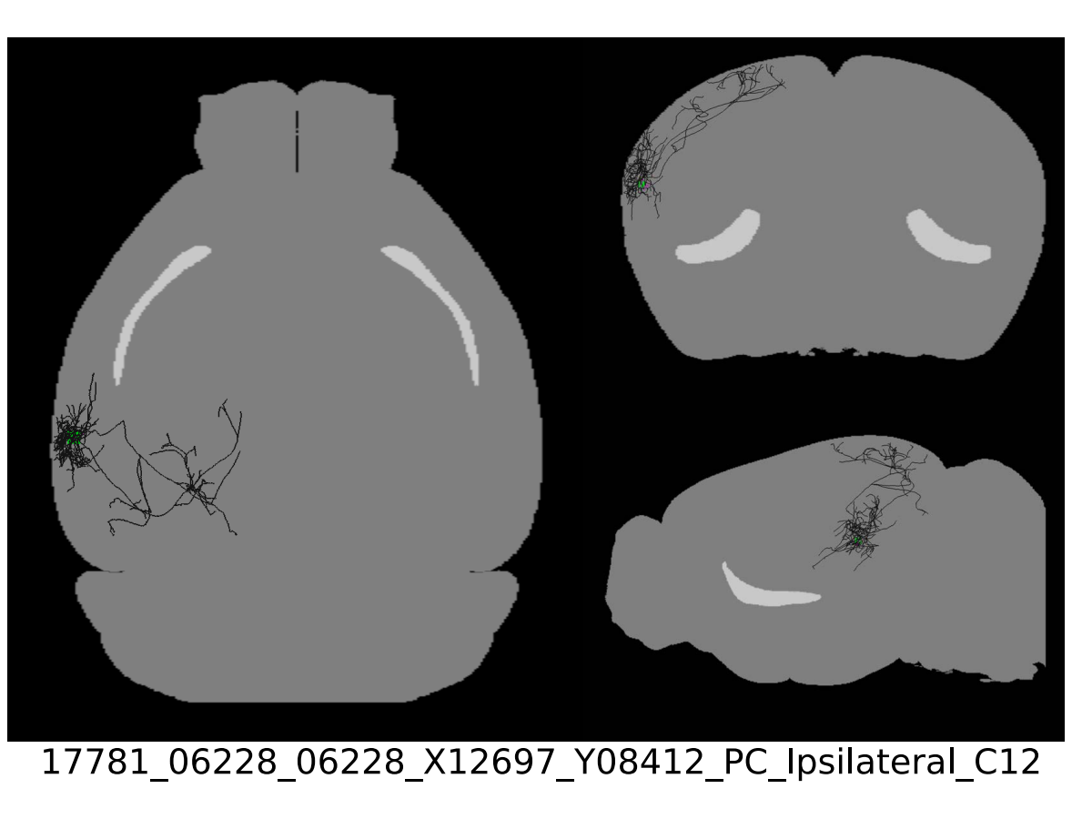

17781_06270_06270_X09494_Y43380_PC_Ipsilateral_C11

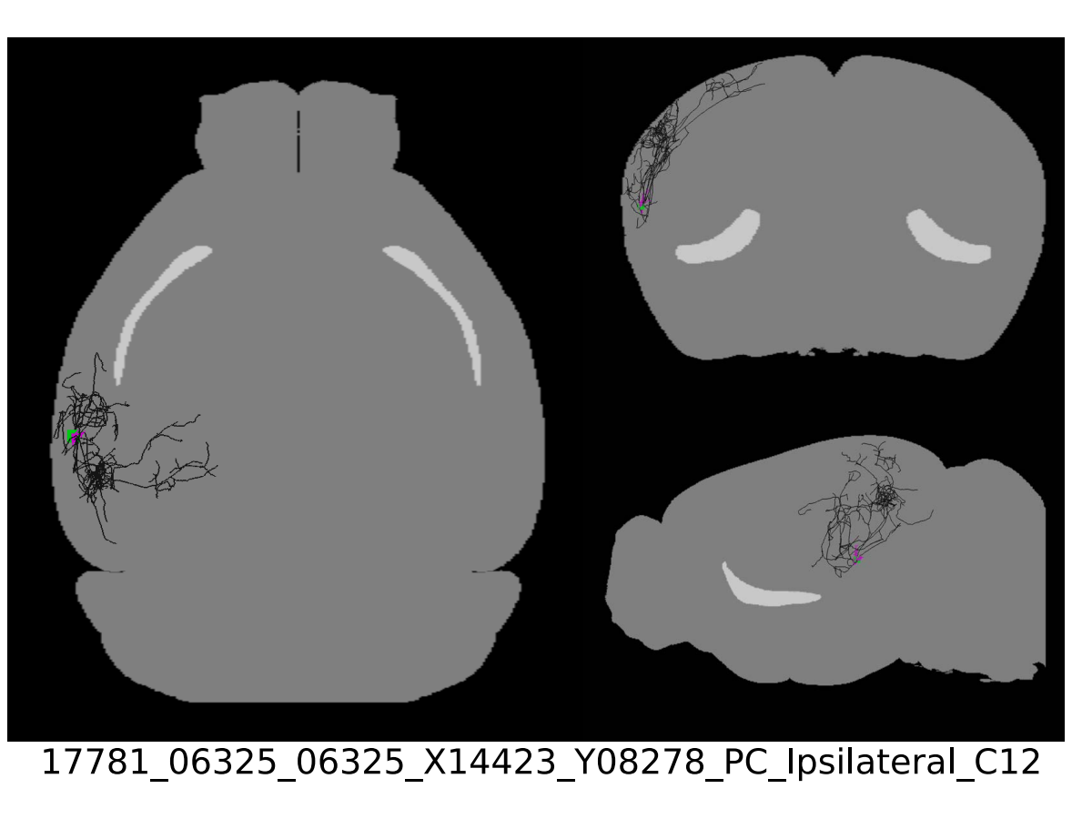

17781_06287_06287_X08230_Y11645_PC_Ipsilateral_C11


17109_06201_06300_X04328_Y06753_PC_Ipsilateral_C12

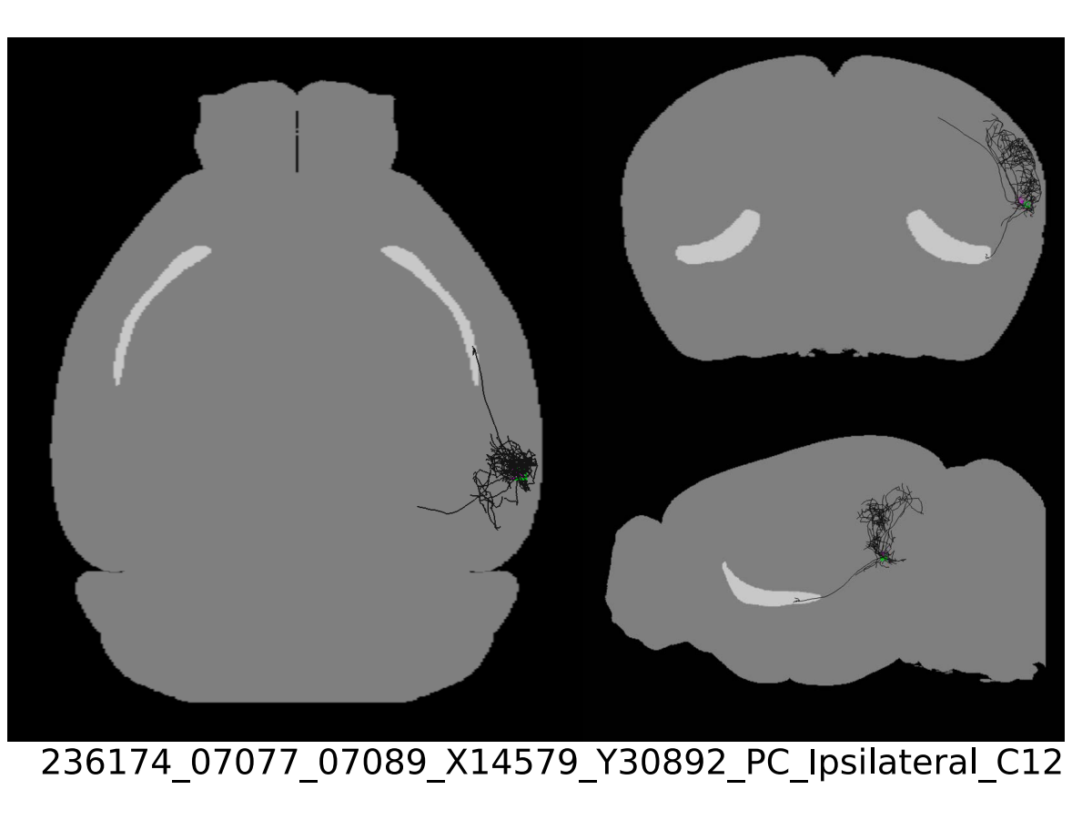

17109_06401_06500_X07641_Y03978_PC_Ipsilateral_C12

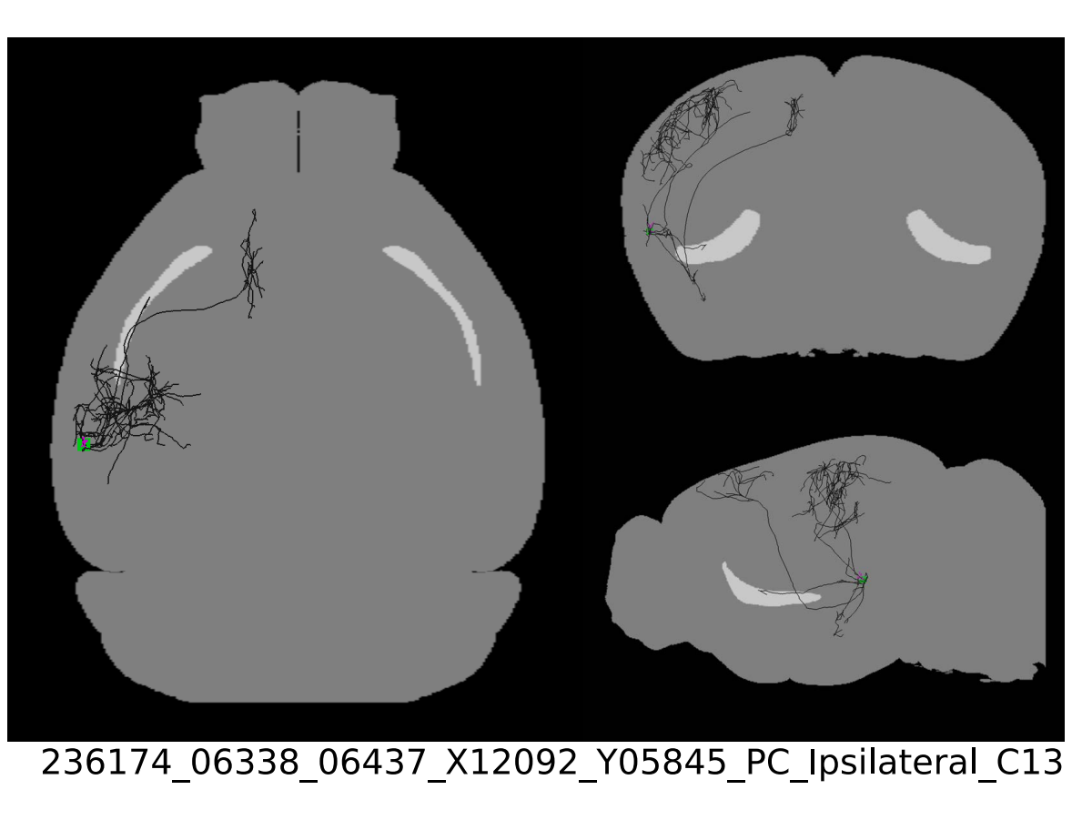

17109_06801_06900_X07432_Y04405_PC_Ipsilateral_C12

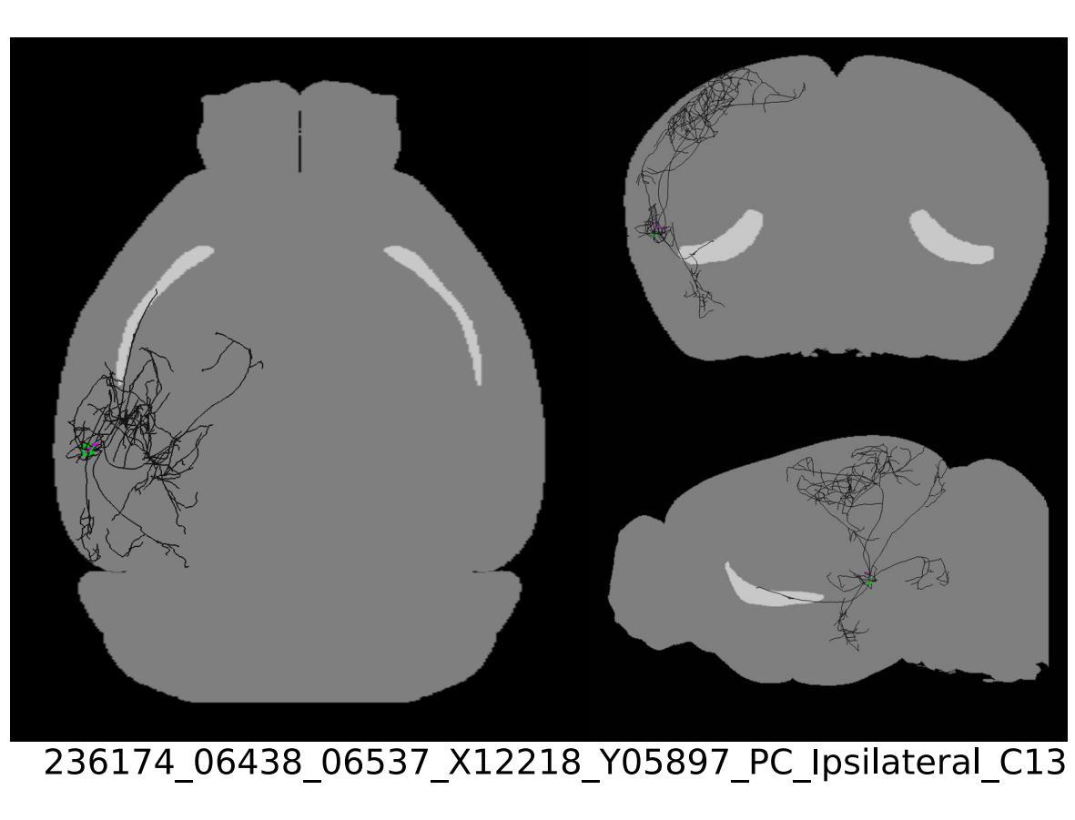

17781_06228_06228_X12697_Y08412_PC_Ipsilateral_C12

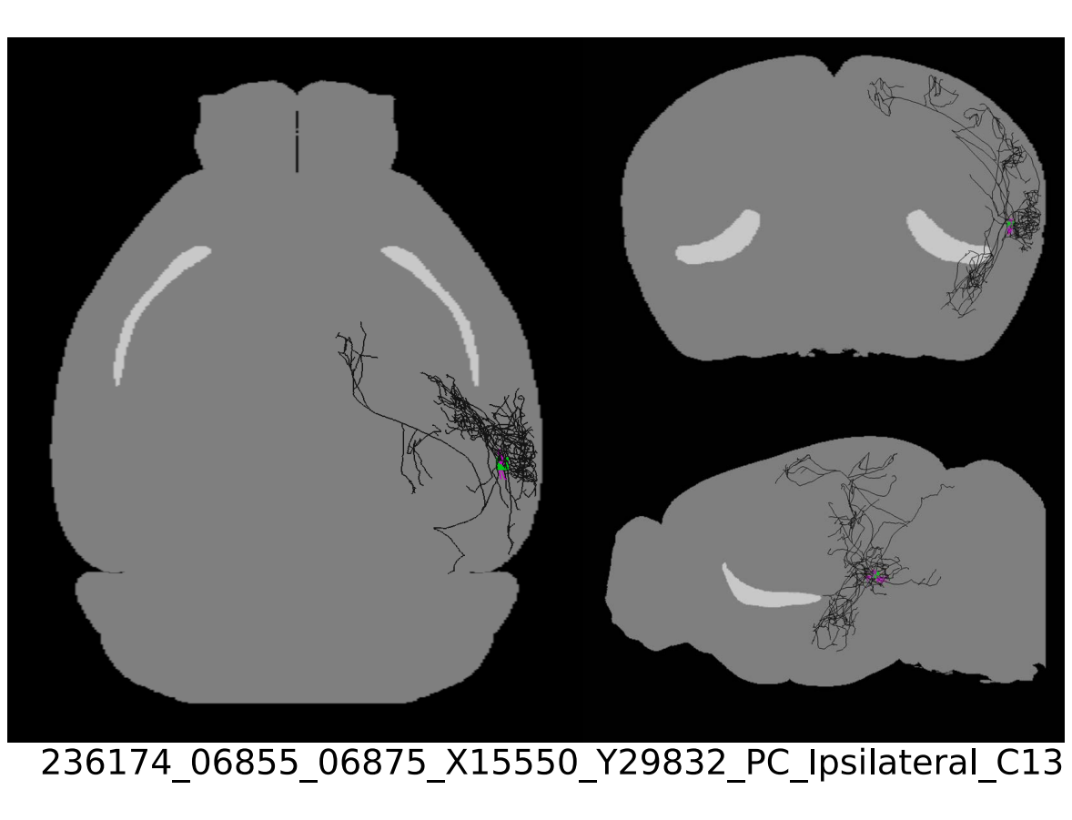

17781_06325_06325_X14423_Y08278_PC_Ipsilateral_C12

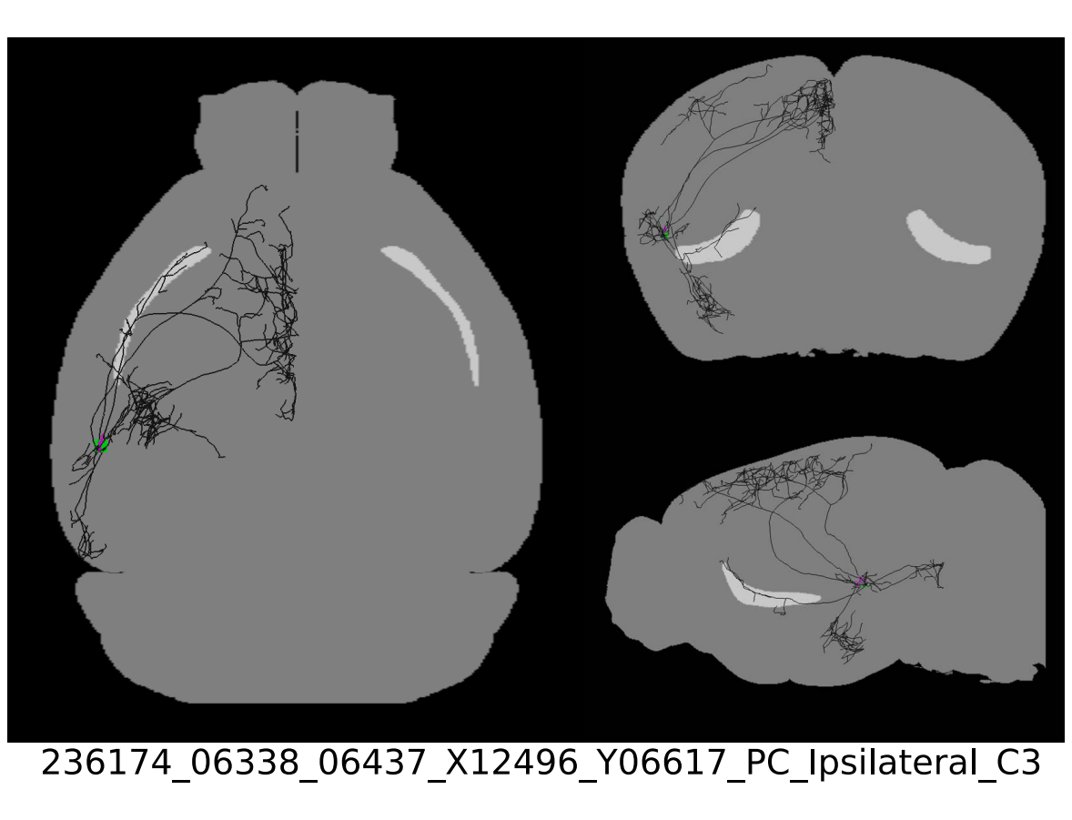

17781_06866_06866_X08954_Y41812_PC_Ipsilateral_C12

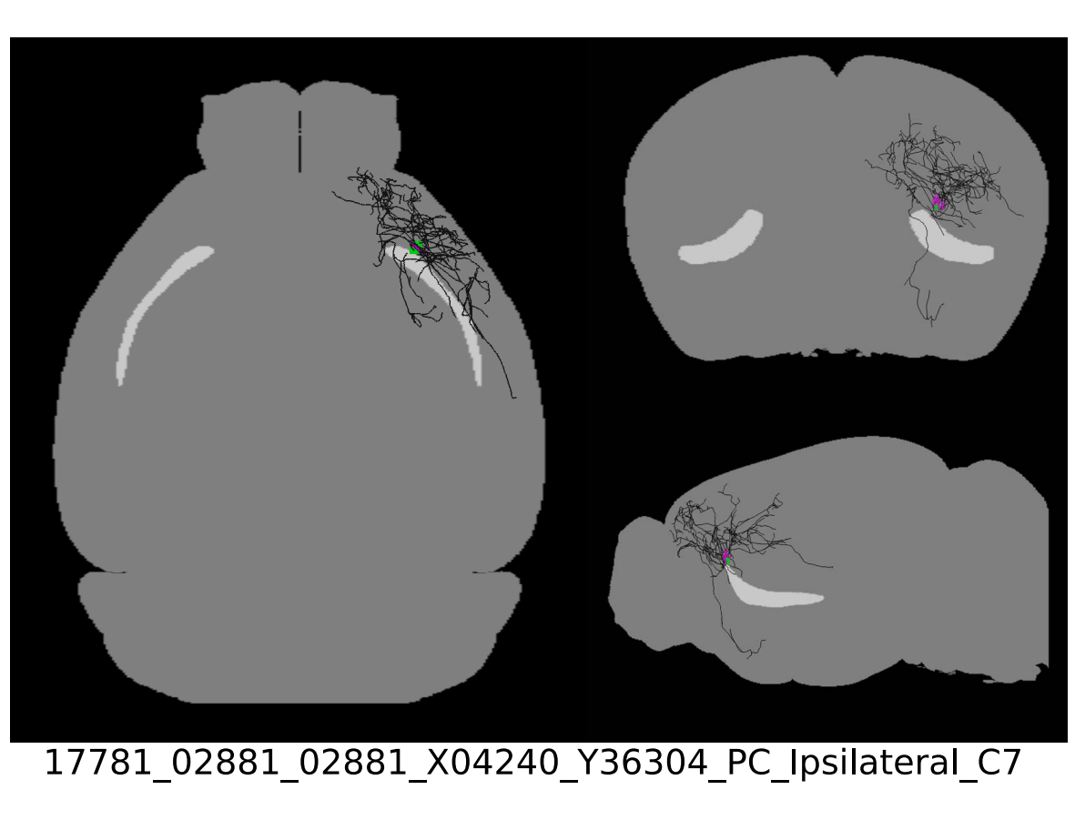

236174_07077_07089_X14579_Y30892_PC_Ipsilateral_C12

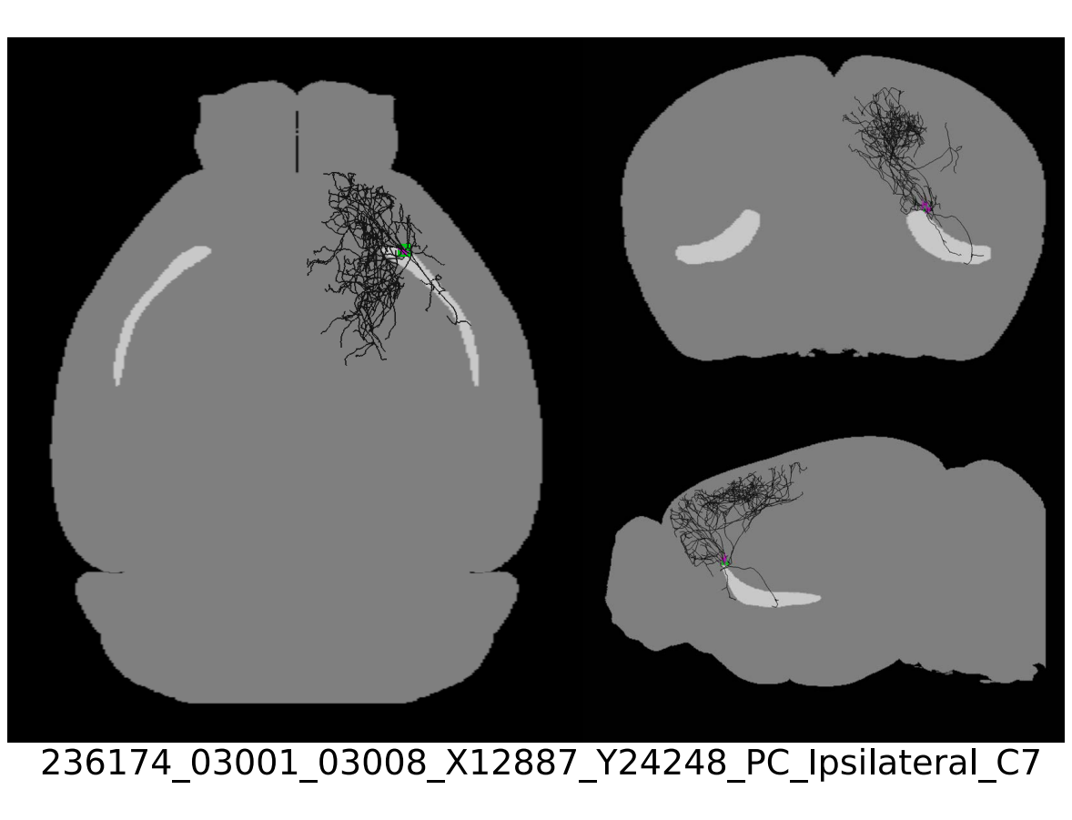

236174_06338_06437_X12092_Y05845_PC_Ipsilateral_C13

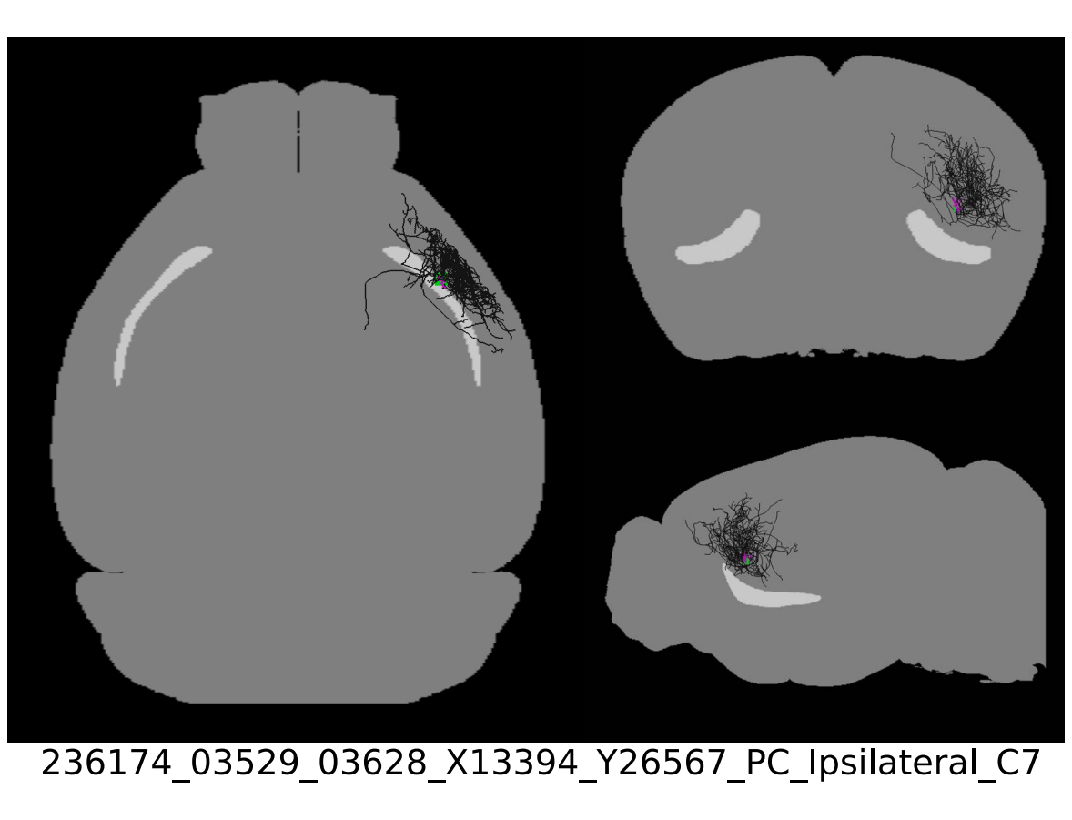

236174_06438_06537_X12218_Y05897_PC_Ipsilateral_C13

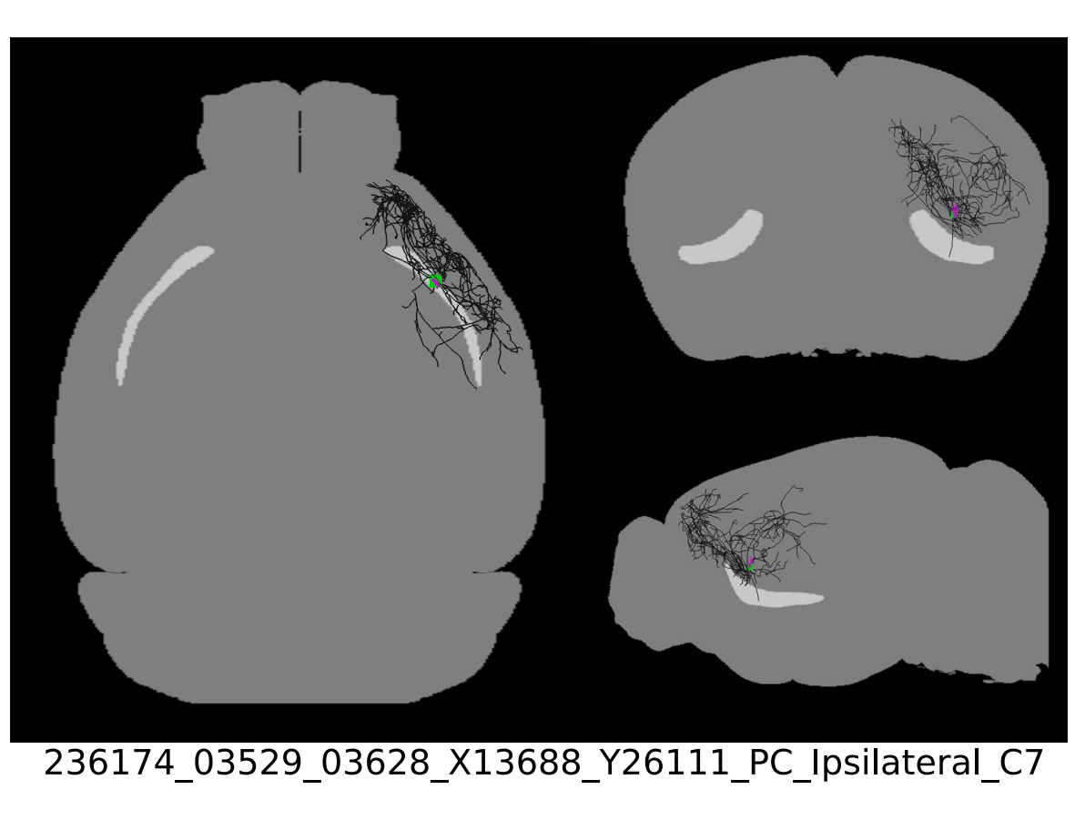

236174_06855_06875_X15550_Y29832_PC_Ipsilateral_C13

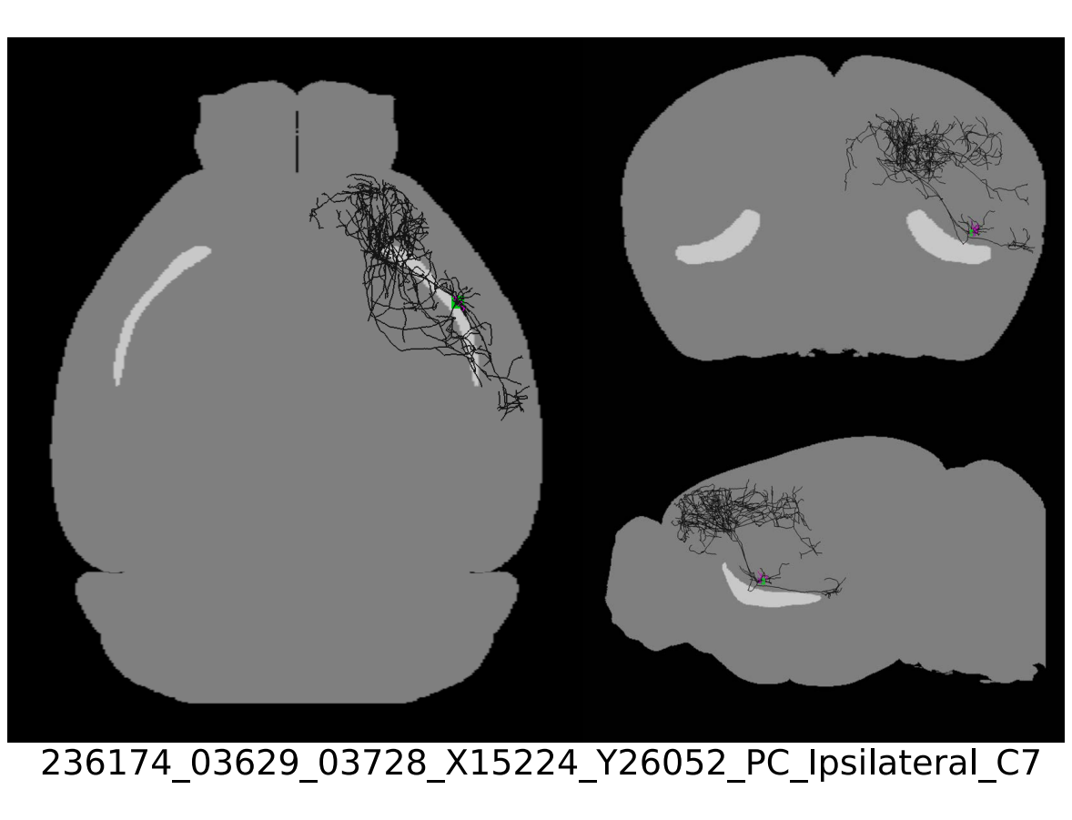

236174_06338_06437_X12496_Y06617_PC_Ipsilateral_C3

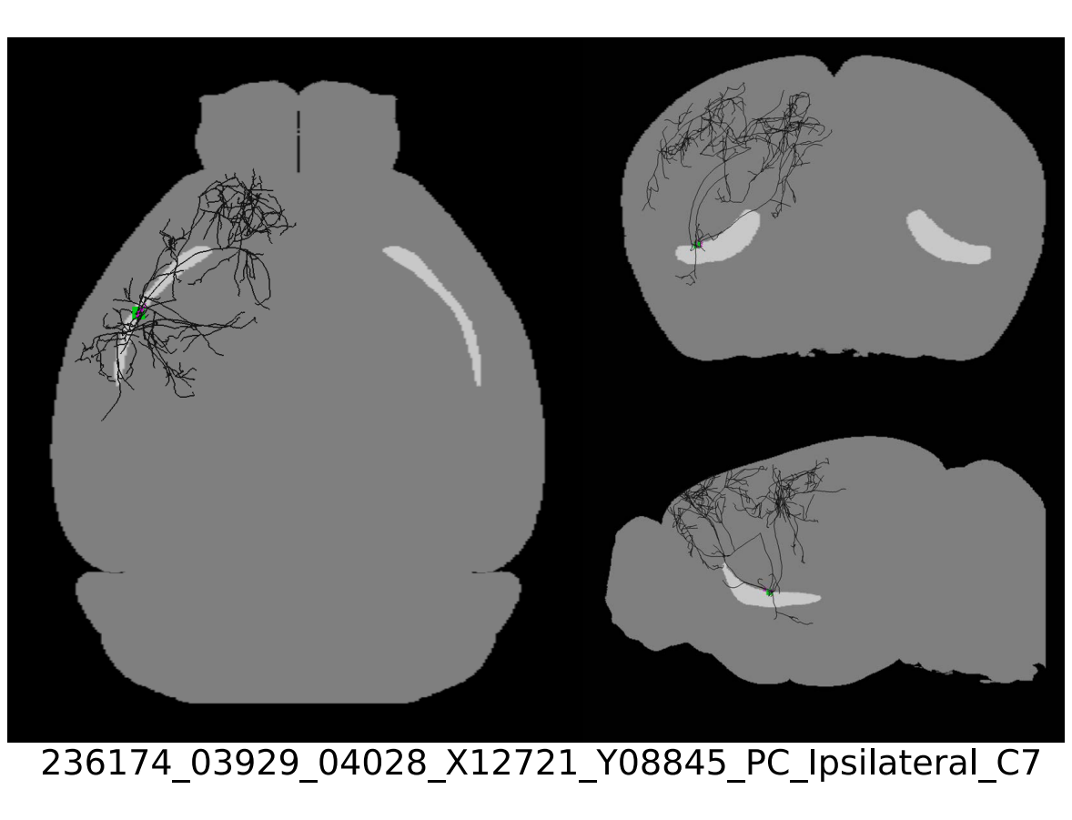

17781_02881_02881_X04240_Y36304_PC_Ipsilateral_C7

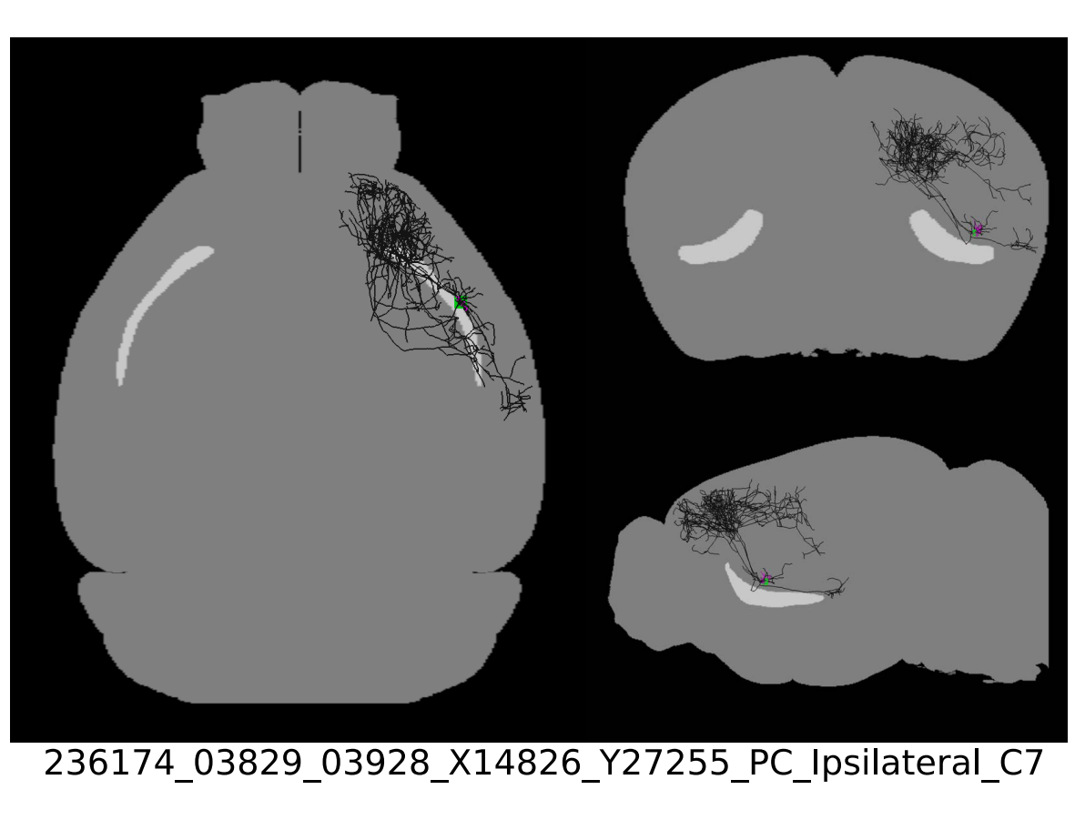

236174_03001_03008_X12887_Y24248_PC_Ipsilateral_C7


236174_03529_03628_X13394_Y26567_PC_Ipsilateral_C7

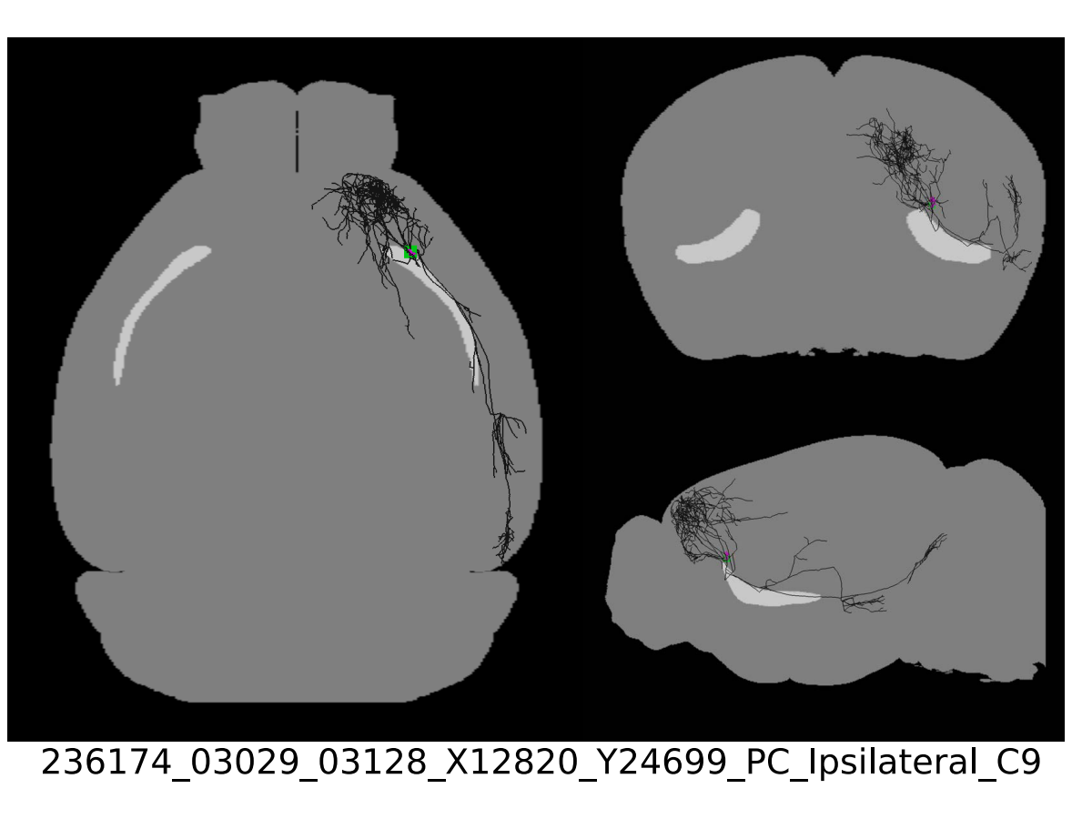

236174_03529_03628_X13688_Y26111_PC_Ipsilateral_C7

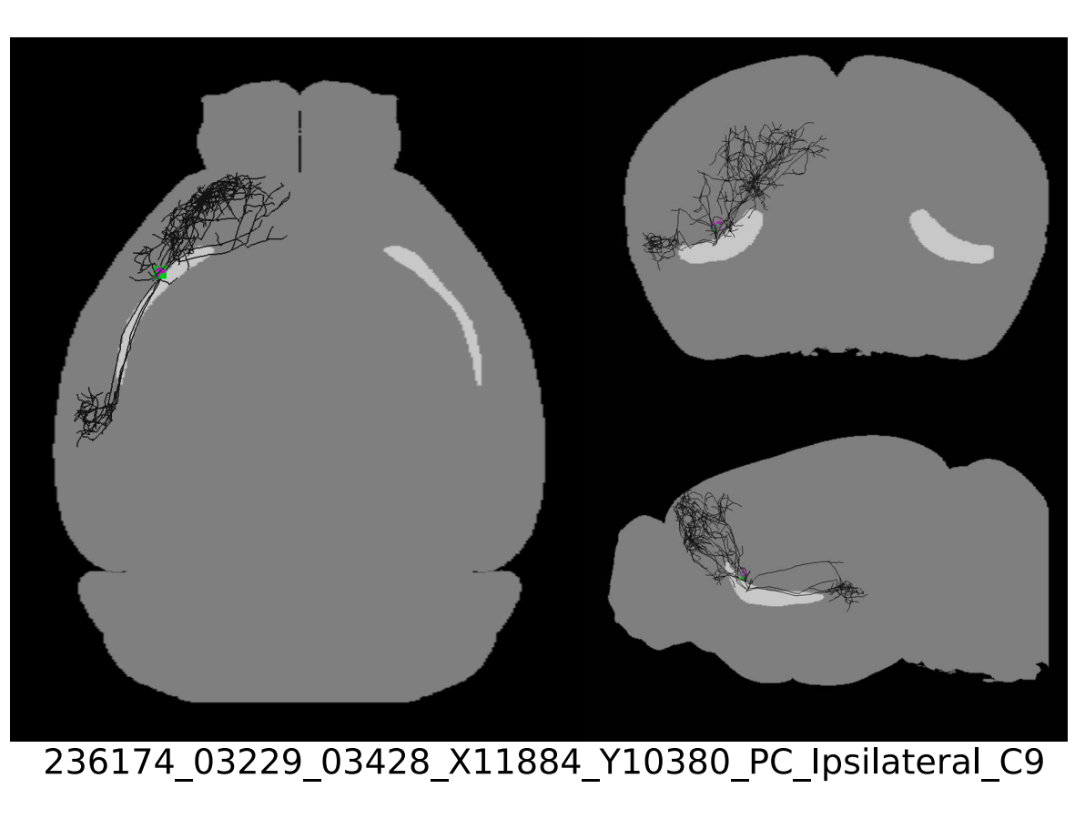

236174_03629_03728_X15224_Y26052_PC_Ipsilateral_C7

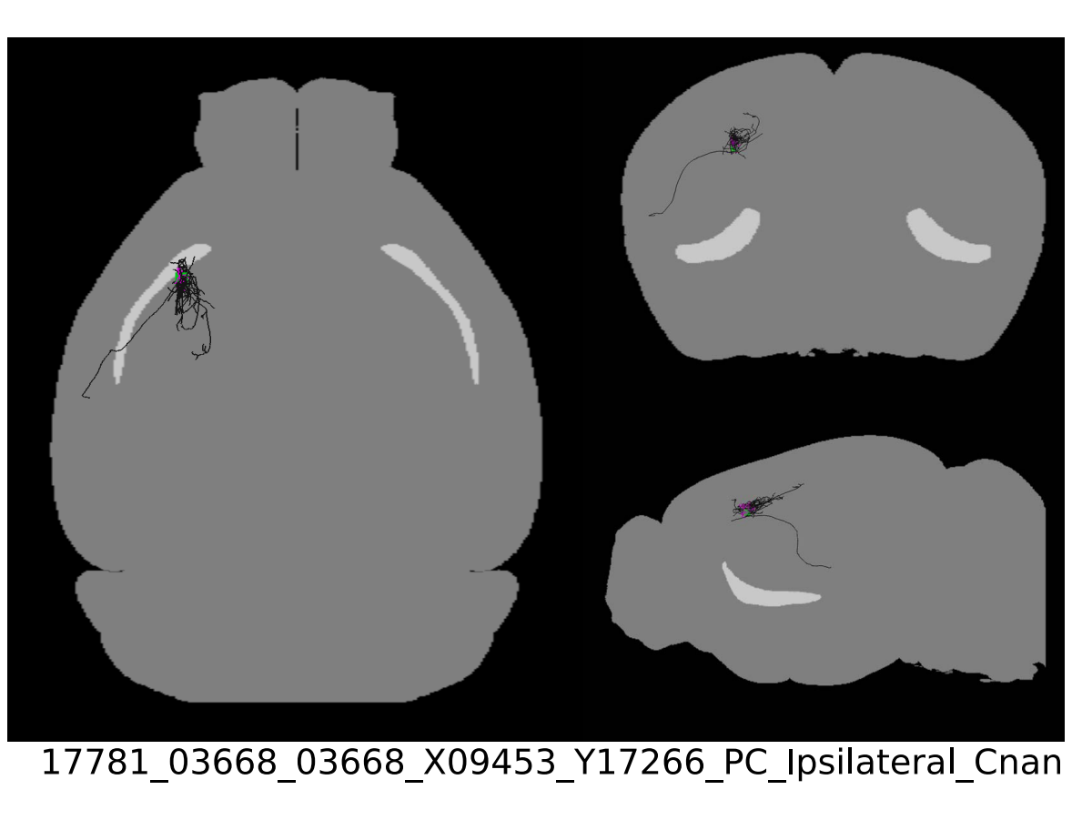

236174_03929_04028_X12721_Y08845_PC_Ipsilateral_C7

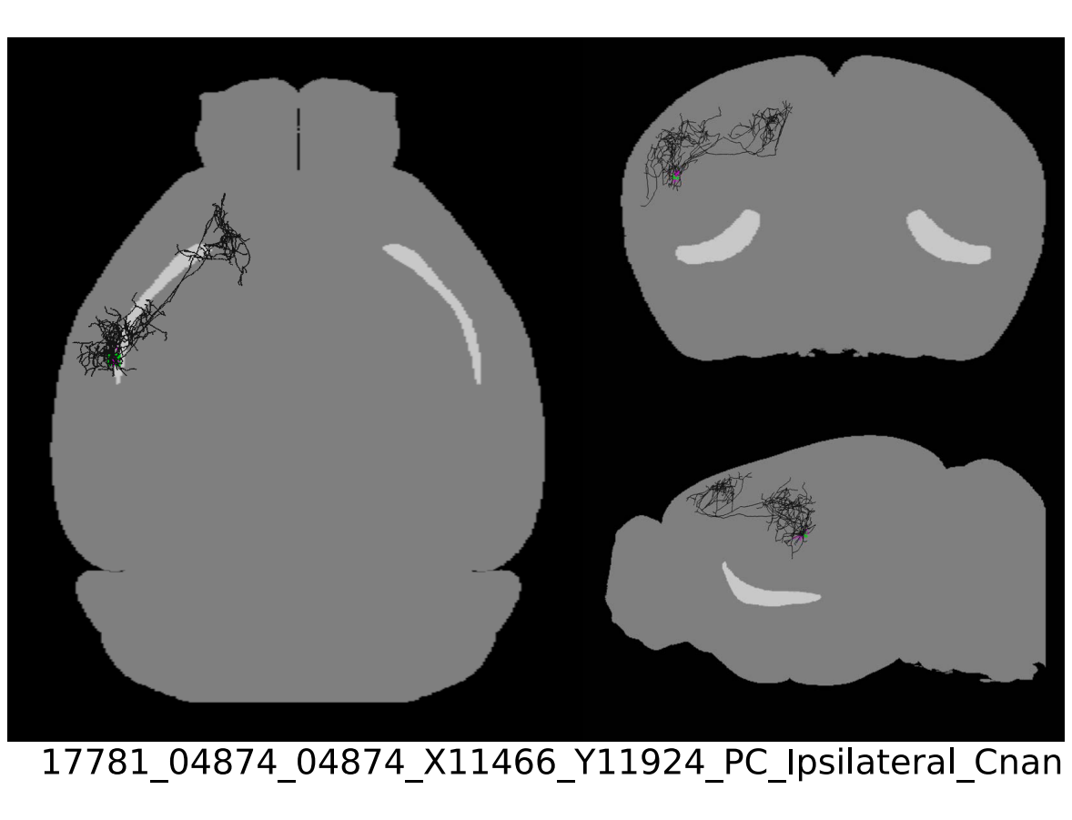

236174_03829_03928_X14826_Y27255_PC_Ipsilateral_C7

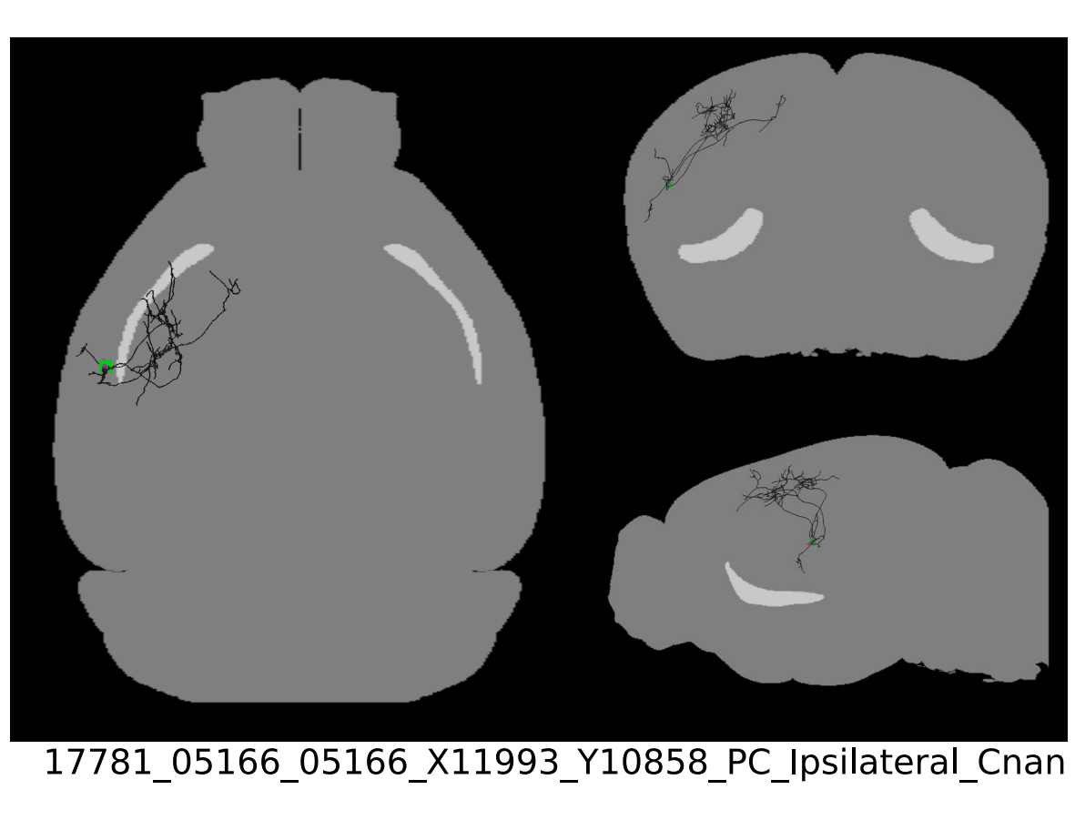

17781_03139_03139_X20033_Y17506_PC_Ipsilateral_C9

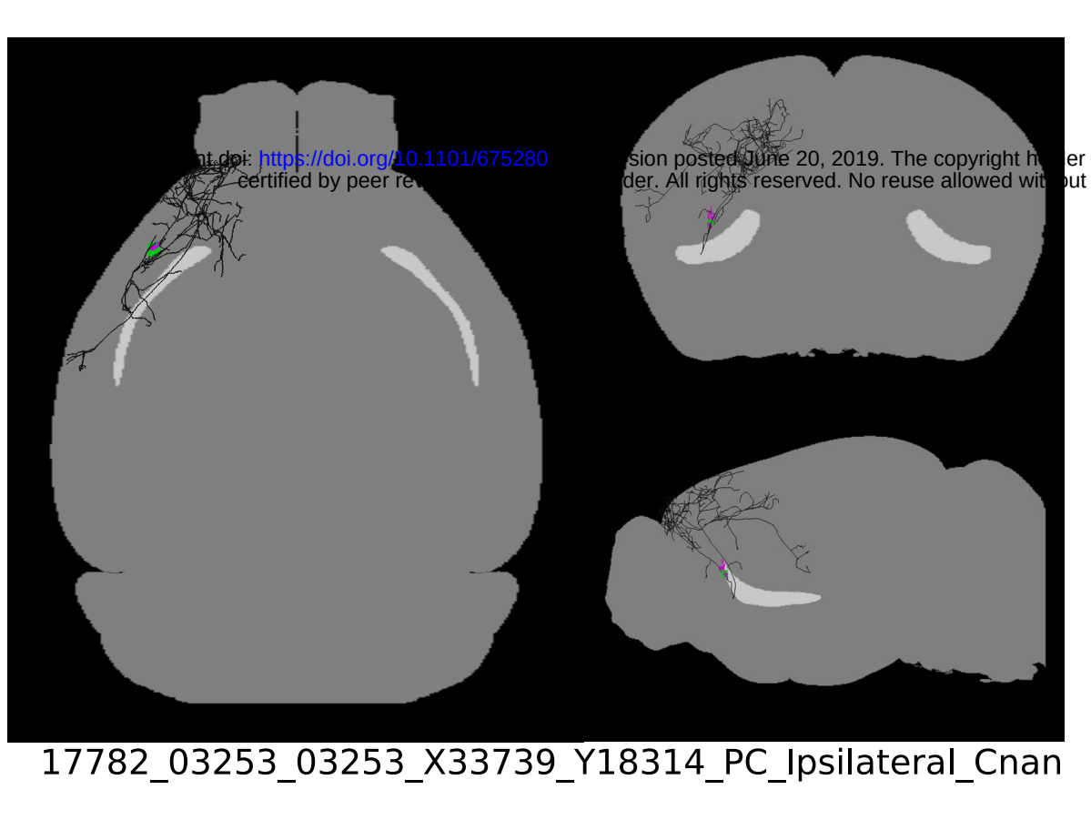

236174_03029_03128_X12820_Y24699_PC_Ipsilateral_C9

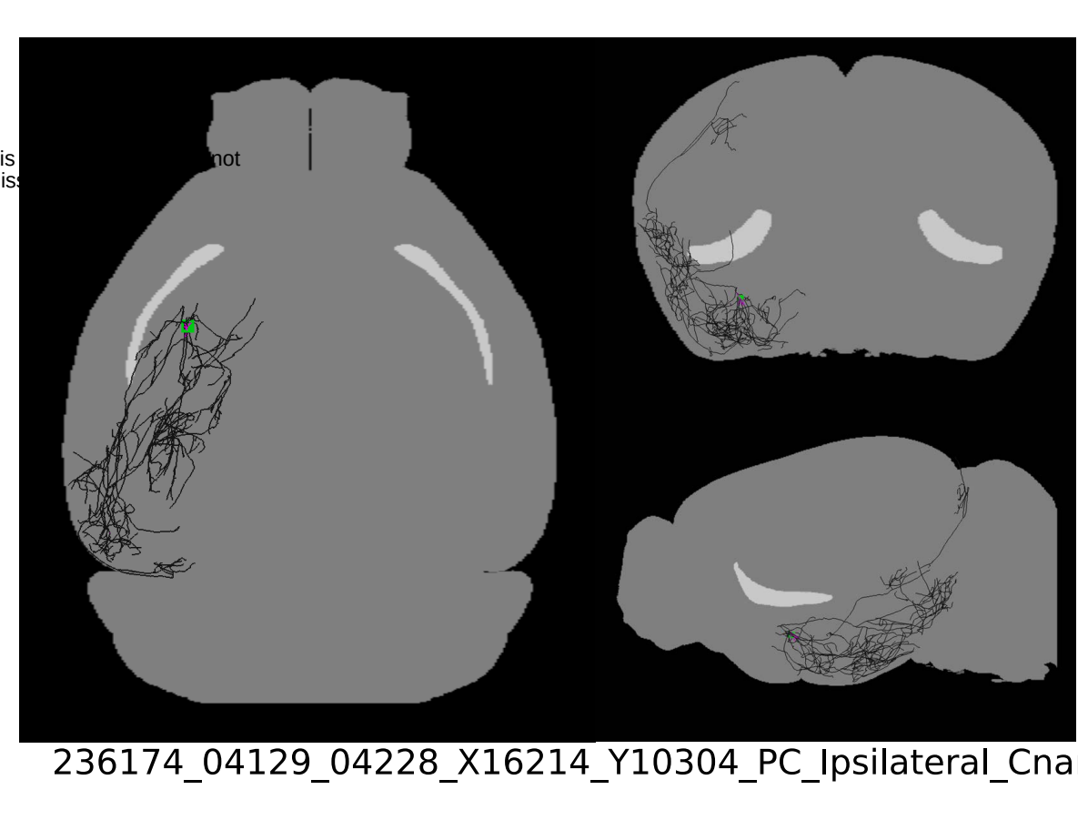

236174_03229_03428_X11884_Y10380_PC_Ipsilateral_C9


17781_03668_03668_X09453_Y17266_PC_Ipsilateral_Cnan


17781_04874_04874_X11466_Y11924_PC_Ipsilateral_Cnan


17781_05166_05166_X11993_Y10858_PC_Ipsilateral_Cnan


17782_03253_03253_X33739_Y18314_PC_Ipsilateral_Cnan


236174_04129_04228_X16214_Y10304_PC_Ipsilateral_Cnan

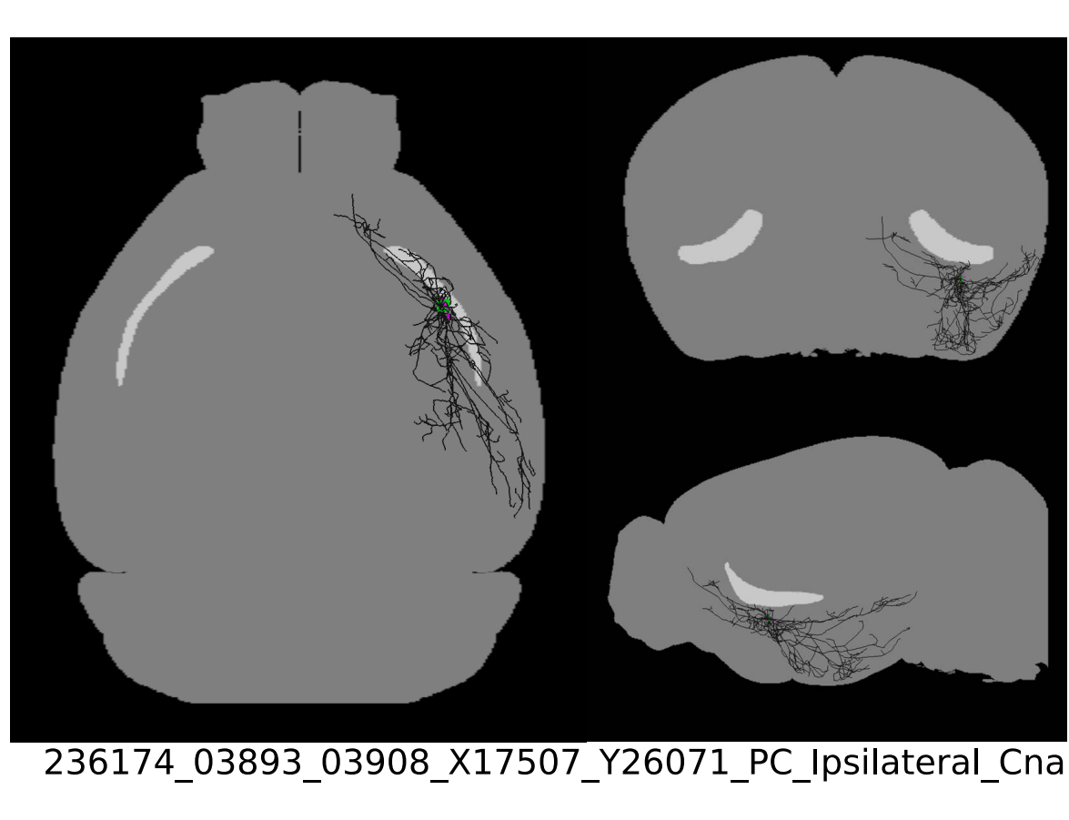
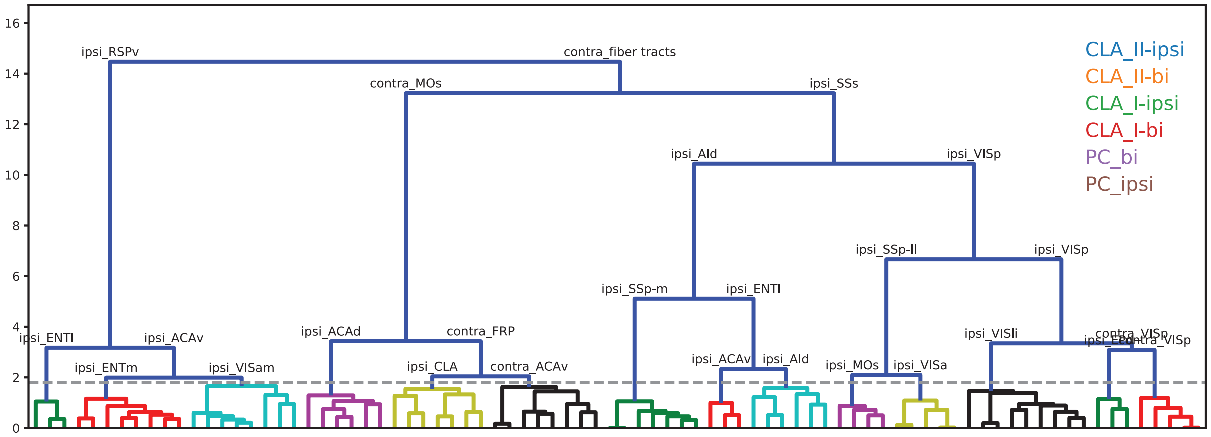
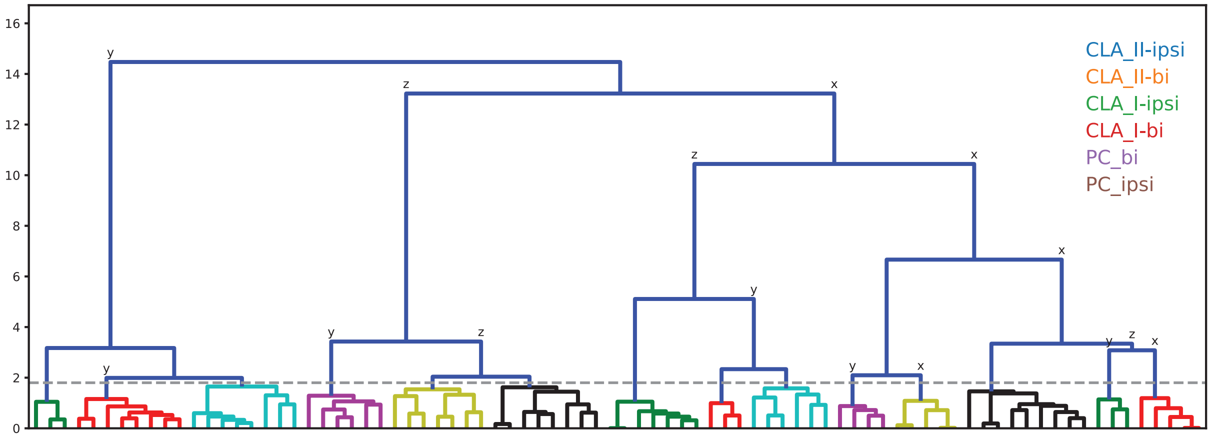

236174_03893_03908_X17507_Y26071_PC_Ipsilateral_Cnan

Figure S6

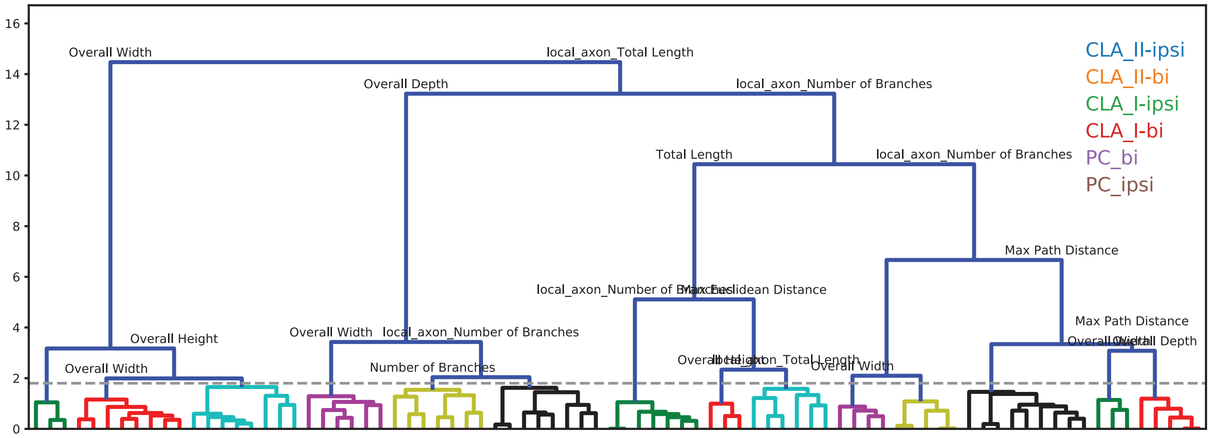
Projection regions



Soma coordinates



Axon morphological features



Dendrite morphological features

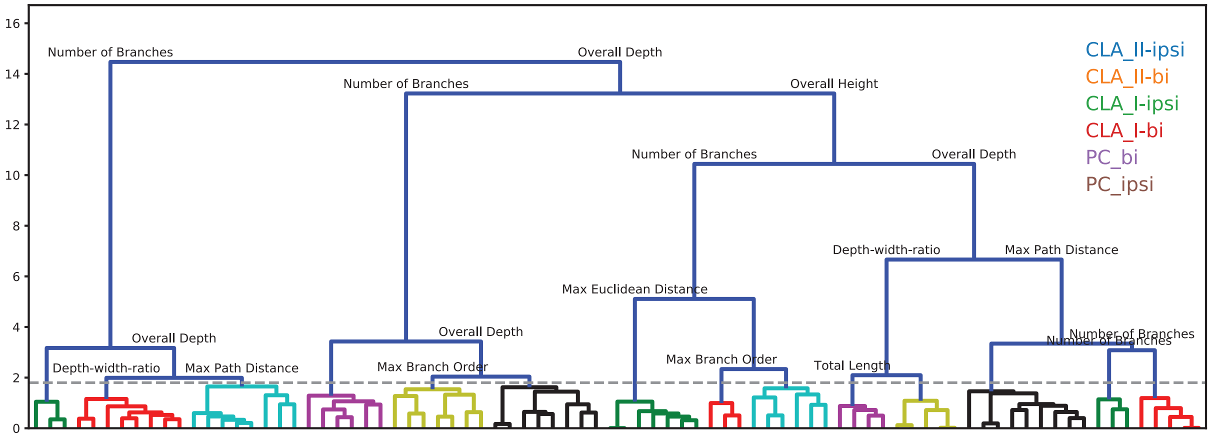
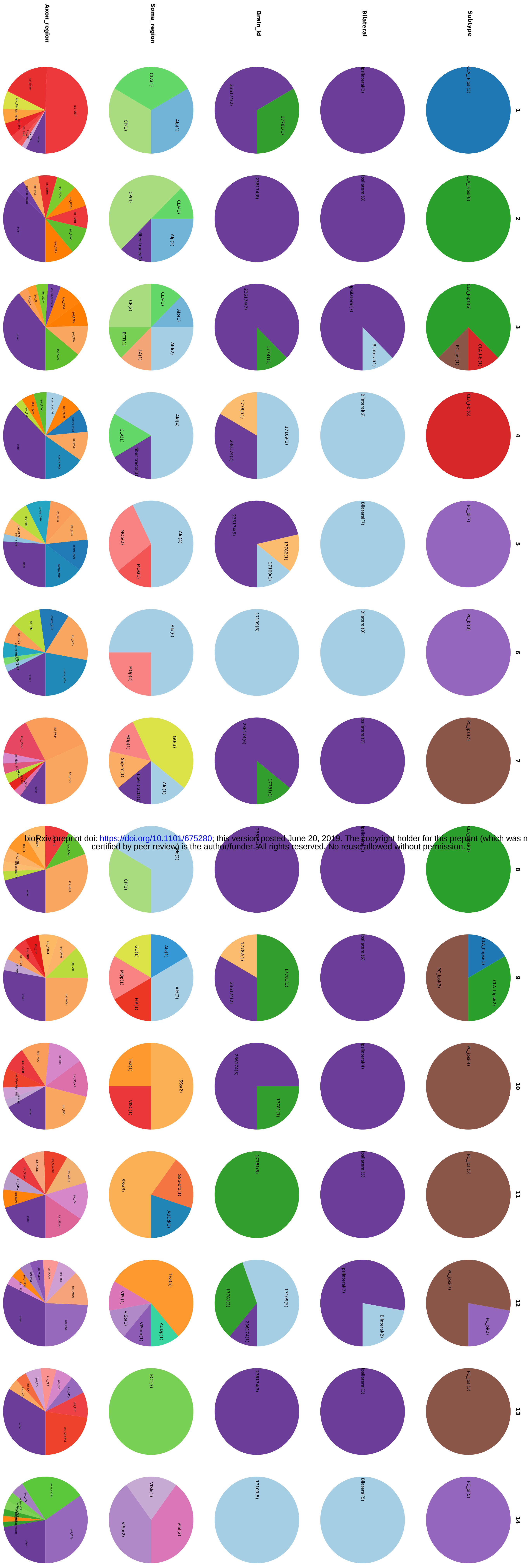


Figure S7


Axon_region
Soma_region
Brain_id
Bilateral
Subtype
1
2
3
4
5
6
7
8
9
10
11
12
13
14