

Modular strategies for spatial mapping of diverse cell type data of the mouse brain

⁴ Nicholas J. Tustison¹, Min Chen², Fae N. Kronman³, Jeffrey T. Duda², Clare Gamlin⁴, Mia
⁵ G. Tustison, Michael Kunst⁴, Rachel Dalley⁴, Staci Sorenson⁴, Quanxin Wang⁴, Lydia Ng⁴,
⁶ Yongsoo Kim³, and James C. Gee²

¹Department of Radiology and Medical Imaging, University of Virginia, Charlottesville, VA

²Department of Radiology, University of Pennsylvania, Philadelphia, PA

³Department of Neural and Behavioral Sciences, Penn State University, Hershey, PA

⁴Allen Institute for Brain Science, Seattle, WA

11 _____

12 Corresponding authors:

13

14 Nicholas J. Tustison, DSc

15 Department of Radiology and Medical Imaging

16 University of Virginia

17 ntustison@virginia.edu

J. C. G. BLOD

19 James C. Gee, PhD
20 Department of Radiology

20 Department of Radiol
21 University of Pennsyl

21 University of Pe

²³ **Abstract**

²⁴ Large-scale, international collaborative efforts by members of the BRAIN Initiative Cell
²⁵ Census Network (BICCN) consortium are aggregating the most comprehensive reference
²⁶ database to date for diverse cell type profiling of the mouse brain, which encompasses over
²⁷ 40 different multi-modal profiling techniques from more than 30 research groups. One central
²⁸ challenge for this integrative effort has been the need to map these unique datasets into
²⁹ common reference spaces such that the spatial, structural, and functional information from
³⁰ different cell types can be jointly analyzed. However, significant variation in the acquisition,
³¹ tissue processing, and imaging techniques across data types makes mapping such diverse
³² data a multifarious problem. Different data types exhibit unique tissue distortion and signal
³³ characteristics that precludes a single mapping strategy from being generally applicable across
³⁴ all cell type data. Tailored mapping approaches are often needed to address the unique barriers
³⁵ present in each modality. This work highlights modular atlas mapping strategies developed
³⁶ across separate BICCN studies using the Advanced Normalization Tools Ecosystem (ANTSX)
³⁷ to map spatial transcriptomic (MERFISH) and high-resolution morphology (fMOST) mouse
³⁸ brain data into the Allen Common Coordinate Framework (AllenCCFv3), and developmental
³⁹ (MRI and LSFM) data into the Developmental Common Coordinate Framework (DevCCF).
⁴⁰ We discuss common mapping strategies that can be shared across modalities and driven
⁴¹ by specific challenges from each data type. These mapping strategies include novel open-
⁴² source contributions that are made publicly available through ANTSX. These include 1) a
⁴³ velocity flow-based approach for continuously mapping developmental trajectories such as
⁴⁴ that characterizing the DevCCF and 2) an automated framework for determining structural
⁴⁵ morphology solely through the leveraging of publicly resources. Finally, we provide general
⁴⁶ guidance to aid investigators to tailor these strategies to address unique data challenges
⁴⁷ without the need to develop additional specialized software.

⁴⁸ 1 Introduction

⁴⁹ Over the past decade there have been significant advancements in mesoscopic single-cell
⁵⁰ analysis of the mouse brain. It is now possible to track single neurons in mouse brains¹,
⁵¹ observe whole brain developmental changes on a cellular level², associate brain regions
⁵² and tissues with their genetic composition³, and locally characterize neural connectivity⁴.
⁵³ Much of these scientific achievements have been made possible due to breakthroughs in high
⁵⁴ resolution cell profiling and imaging techniques that permit submicron, multi-modal, 3-D
⁵⁵ characterizations of whole mouse brains. Among these include advanced techniques such
⁵⁶ as micro-optical sectioning tomography⁶, tissue clearing^{1,7}, spatial transcriptomics⁹, and
⁵⁷ single-cell genomic profiling¹⁰, which have greatly expanded the resolution and specificity of
⁵⁸ single-cell measurements in the brain.

⁵⁹ Recent efforts by the National Institutes of Health's Brain Research Through Advancing
⁶⁰ Innovative Neurotechnologies (BRAIN) Initiative has pushed for large-scale, international
⁶¹ collaborative efforts to utilize these advanced single-cell techniques to create a comprehensive
⁶² reference database for high-resolution transcriptomic, epigenomic, structural and imaging
⁶³ data of the mouse brain. This consortium of laboratories and data centers, known as the
⁶⁴ BRAIN Initiative Cell Census Network (BICCN), has archived datasets encompassing over 40
⁶⁵ different multi-modal profiling techniques from more than 30 research groups, each providing
⁶⁶ unique characterizations of distinct cell types in the brain¹¹. Several of these modalities have
⁶⁷ been further developed into reference atlases to facilitate spatial alignment of individual
⁶⁸ brains and different data types into a common coordinate framework (CCF), thus allowing
⁶⁹ diverse single-cell information to be analyzed in an integrated manner. The most notable
⁷⁰ of these atlases is the Allen Mouse Brain Common Coordinate Framework (AllenCCFv3)¹²,
⁷¹ which serves as a primary target coordinate space for much of the work associated with the
⁷² BICCN. Other atlases include modality-specific atlases^{13–15}, and spatiotemporal atlases^{16,17}
⁷³ for the developing mouse brain.

⁷⁴ **1.1 Mouse brain mapping**

⁷⁵ The cross-modality associations that can be learned from mapping different cell type data
⁷⁶ into a CCF is critical for improving our understanding of the complex relationships between
⁷⁷ cellular structure, morphology, and genetics in the brain. However, finding an accurate
⁷⁸ mapping between each individual mouse brain and a CCF is a challenging and heterogeneous
⁷⁹ task. There is significant variance in the imaging protocols across different cell type data
⁸⁰ as well as different tissue processing and imaging methods which can potentially introduce
⁸¹ tissue distortion and signal differences^{18,19}. Certain modalities can have poor intensity
⁸² correspondence with the CCF, negatively impacting image alignment accuracy. Studies
⁸³ targeting specific regions or cell types can lead to missing anatomical correspondences. Other
⁸⁴ considerations include artifacts such as tissue distortion, holes, bubbles, folding, tears, and
⁸⁵ missing sections in the data that often require manual correction^{20–23}. Given the diversity
⁸⁶ of these challenges, it is unlikely any single mapping approach can be generally applicable
⁸⁷ across all cell type data. Diverse, and often specialized, strategies are needed to address the
⁸⁸ unique barriers present for mapping each modality.

⁸⁹ Existing solutions to address mapping cell type data into the AllenCCFv3 falls broadly into
⁹⁰ three main categories. The first consists of integrated processing platforms that directly
⁹¹ provide mapped data to the users. These include the Allen Brain Cell Atlas²⁴ for the Allen
⁹² Reference Atlas (ARA) and associated data, the Brain Architecture Portal²⁵ for combined ex
⁹³ vivo radiology and histology data, OpenBrainMap²⁶ for connectivity data, and the Image and
⁹⁴ Multi-Morphology Pipeline²⁷ for high resolution morphology data. These platforms provide
⁹⁵ users online access to pre-processed, multi-modal cell type data that are already mapped to
⁹⁶ the AllenCCFv3. The platforms are designed such that the data is interactively manipulated
⁹⁷ by users through integrated visualization software that allow users to spatially manipulate
⁹⁸ and explore each dataset within the mapped space. While highly convenient for investigators
⁹⁹ who are interested in studying the specific modalities provided by these platforms, these
¹⁰⁰ systems can be limited in flexibility, general applicability, and public availability. As a result,
¹⁰¹ investigators often find it difficult to apply the same mapping solutions to their own data.

¹⁰² The second category comprises specialized approaches specifically designed for mapping

103 one or more modalities into a CCF. These approaches use combinations of specialized
104 manual and automated processes that address specific challenges in each modality. Examples
105 include approaches for mapping histology^{28–30}, magnetic resonance imaging (MRI)³⁷, micro-
106 computed tomography (microCT)^{35,37}, light-sheet fluorescence microscopy (LSFM)^{34,36–39},
107 fluorescence micro-optical sectioning tomography (fMOST)^{15,40} and transcriptomic data^{41–43}.
108 As specialized approaches, these techniques tend to boast higher mapping accuracy, robustness,
109 and ease of use. Conversely, their specialized designs often rely on base assumptions regarding
110 the data type that can make them rigid and difficult to adapt for new modalities or unexpected
111 artifacts and distortions in the data. Adapting these specialize software tools to use with
112 new data can require significant development, validation time, and engineering expertise that
113 may not be readily available for all investigators.

114 The last category consists of modular mapping approaches constructed using general image
115 analysis toolkits, which are software packages that include modular image processing, segmen-
116 tation and registration tools that have been previously developed, and validated for multiple
117 application areas. Examples of such toolkits include elastix⁴⁴, Slicer3D⁴⁵, ANTsX⁴⁶, and
118 several others which have all been applied towards mouse brain spatial mapping. The main
119 challenge, in these mouse-specific study scenarios, is that tailored pipelines often need be
120 constructed from available software components. Investigators must therefore be familiar with
121 the these tools for formulating new or adapting existing pipelines. However, in comparison
122 to previously described specialized mapping approaches, these approaches are often easier
123 to create and prone to robustness, being typically constructed from pipeline components
124 which have been previously vetted in other contexts. In this work, we highlight such mapping
125 strategies designed using the ANTsX framework to map distinct mouse cell type data with
126 different characteristics into existing CCFs.

127 1.2 Advanced Normalization Tools (ANTsX)

128 The Advanced Normalization Tools Ecosystem (ANTsX) has been used in a number of
129 applications for mapping mouse brain data as part of core processing steps in various
130 workflows^{30,47–50}, particularly its pairwise, intensity-based image registration capabilities⁵¹ and

¹³¹ bias field correction⁵². Historically, ANTsX development is originally based on fundamental
¹³² approaches to image mapping^{53–55}, particularly in the human brain, which has resulted
¹³³ in core contributions to the field such as the widely-used Symmetric Normalization (SyN)
¹³⁴ algorithm⁵¹. Since its development, various independent platforms have been used to evaluate
¹³⁵ ANTsX image registration capabilities in the context of different application foci which
¹³⁶ include multi-site brain MRI data⁵⁶, pulmonary CT data⁵⁷, and most recently, multi-modal
¹³⁷ brain registration in the presence of tumors⁵⁸.

¹³⁸ Apart from its registration capabilities, ANTsX comprises additional functionality such
¹³⁹ as template generation⁵⁹, intensity-based segmentation⁶⁰, preprocessing^{52,61}, deep learning
¹⁴⁰ networks⁴⁶, and other utilities relevant to brain mapping (see Table 1). The use of the toolkit
¹⁴¹ has demonstrated high performance in multiple application areas (e.g., consensus labeling⁶²,
¹⁴² brain tumor segmentation⁶³, and cardiac motion estimation⁶⁴). Importantly, ANTsX is built
¹⁴³ on the Insight Toolkit (ITK)⁶⁵ deriving benefit from the open-source community of scientists
¹⁴⁴ and programmers as well as providing an important resource for algorithmic development,
¹⁴⁵ evaluation, and improvement.

¹⁴⁶ With respect to mouse cell type data, ANTsX provides a comprehensive toolset which serves
¹⁴⁷ as a basis for developing modular frameworks for mapping diverse image data into common
¹⁴⁸ coordinate frameworks (CCFs). Herein, we highlight its application for mapping data from
¹⁴⁹ separate BICCN projects focused on distinct data types: morphology data using fluorescence
¹⁵⁰ micro-optical sectioning tomography (fMOST), spatial transcriptomics from multiplexed error-
¹⁵¹ robust fluorescence in situ hybridization (MERFISH) data, and time-series developmental
¹⁵² data using light sheet fluorescence microscopy (LSFM) and magnetic resonance imaging
¹⁵³ (MRI). We describe both shared and targeted strategies developed to address the specific
¹⁵⁴ challenges of these modalities.

¹⁵⁵ 1.3 Novel ANTsX-based open-source contributions

¹⁵⁶ We introduce two novel inclusions to the ANTsX toolset that were developed as part of the
¹⁵⁷ MRI mapping and analysis pipeline for the Developmental Common Coordinate Framework
¹⁵⁸ (DevCCF). Consistent with previous ANTsX development, newly introduced capabilities

introduced below are available through ANTsX (specifically, via R and Python ANTsX packages), and illustrated through self-contained examples in the ANTsX tutorial (<https://tinyurl.com/antsxtutorial>) with a dedicated GitHub repository specific to this work (<https://github.com/ntustison/ANTsXMouseBrainMapping>). To complement standard preprocessing steps (e.g., bias correction, brain masking), additional mouse brain specific tools have also been introduced to the ANTsX ecosystem, such as section reconstruction and landmark-based alignment with corresponding processing scripts (<https://github.com/dontminchenit/CCFAlignmentToolkit>).

1.3.1 Continuously mapping the DevCCF developmental trajectory with a velocity flow model

Recently, the Developmental Common Coordinate Framework (DevCCF) was introduced to the mouse brain research community as a public resource¹⁶ comprising symmetric atlases of multi-modal image data and anatomical segmentations defined by developmental ontology. These templates sample the mouse embryonic days E11.5, E13.5, E15.5, E18.5 and postnatal days P4, P14, and P56. Modalities include LSFM and at least four MRI contrasts per developmental stage. Anatomical parcellations are also available for each time point and were generated from ANTsX-based mappings of gene expression and other cell type data. Additionally, the P56 template was integrated with the AllenCCFv3 to further enhance the practical utility of the DevCCF. These processes, specifically template generation and multi-modal image mapping, were performed using ANTsX functionality in the presence of image mapping difficulties such as missing data and tissue distortion.

Given the temporal gaps in the discrete set of developmental atlases, we also provide an open-source framework for inferring correspondence within the temporally continuous domain sampled by the existing set of embryonic and postnatal atlases of the DevCCF. This recently developed functionality permits the generation of a diffeomorphic velocity flow transformation model⁶⁶, influenced by previous work⁶⁷. The resulting time-parameterized velocity field spans the stages of the DevCCF where mappings between any two continuous time points within the span bounded by the E11.5 and P56 atlases are determined by numerical integration of

¹⁸⁷ the optimized velocity field.

¹⁸⁸ **1.3.2 Automated structural parcellations of the mouse brain**

¹⁸⁹ In contrast to the pipeline development in human data⁴⁶, limited tools exist yet to create
¹⁹⁰ adequate training data for automated parcellations of the mouse brain. In addition, mouse
¹⁹¹ brain data acquisition often has unique issues, such as lower data quality or sampling
¹⁹² anisotropy which limits its applicability to high resolution resources (e.g., AllenCCFv3,
¹⁹³ DevCCF), specifically with respect to the corresponding granular brain parcellations derived
¹⁹⁴ from numerous hours of expert annotation leveraging multi-modal imaging resources.

¹⁹⁵ Herein, we introduce a mouse brain parcellation pipeline for multi-modal MRI comprising two
¹⁹⁶ novel deep learning components: two-shot learning brain extraction from data augmentation of
¹⁹⁷ two ANTsX templates generated from two open datasets^{68,69} and single-shot brain parcellation
¹⁹⁸ derived from the AllenCCFv3 labelings mapped to the corresponding DevCCF P56 template.
¹⁹⁹ Although we anticipate that this pipeline will be beneficial to the research community, this work
²⁰⁰ demonstrates more generally how one can leverage ANTsX tools and other public resources
²⁰¹ for developing quantitative mouse brain morphological tools. Evaluation is performed on an
²⁰² independent open dataset⁷⁰ comprising longitudinal acquisitions of multiple specimens.

203 **2 Results**

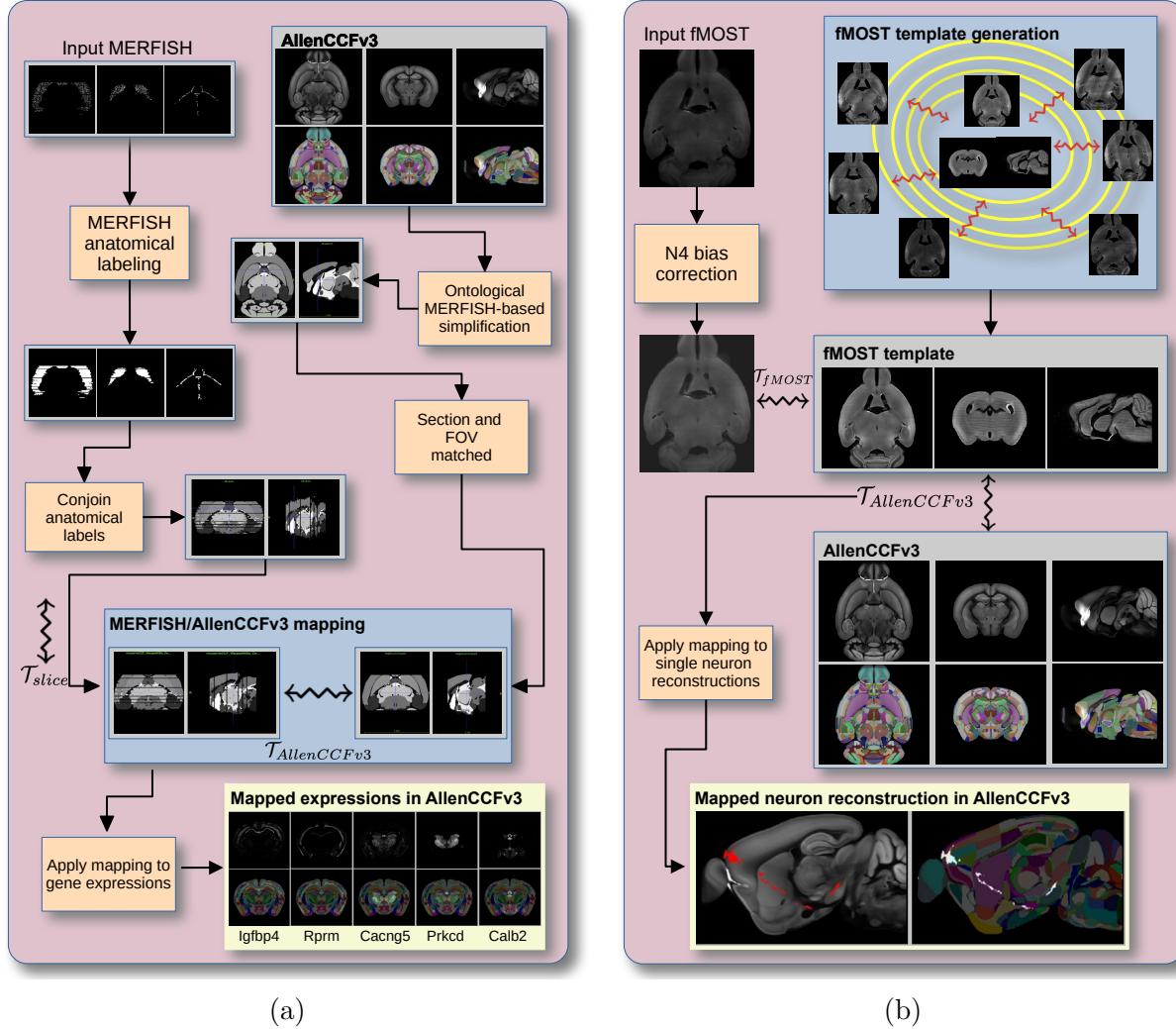


Figure 1: Diagram of the two ANTsX-based pipelines for mapping (a) MERFISH and (b)fMOST data into the space of AllenCCFv3. Each generates the requisite transforms, \mathcal{T} , to map individual images to the CCF.

204 **2.1 AllenCCFv3 brain image mapping**

205 **2.1.1 Mapping multiplexed error-robust fluorescence *in situ* hybridization
206 (MERFISH) data**

207 **Overview.** The ANTsX framework was used to develop a pipeline for mapping multiplexed
208 error-robust fluorescence *in situ* hybridization (MERFISH) spatial transcriptomic mouse

209 data onto the AllenCCFv3 (see Figure 1(a)). This pipeline, used recently in creating
210 a high-resolution transcriptomic atlas of the mouse brain⁵⁰, performs mappings by first
211 generating anatomical labels from tissue related gene expressions in the MERFISH data, and
212 then spatially matching these labels to corresponding anatomical tissue parcellations in the
213 AllenCCFv3. The pipeline consists of MERFISH data specific preprocessing which includes
214 section reconstruction, mapping corresponding anatomical labels between AllenCCFv3 and the
215 spatial transcriptomic maps of the MERFISH data, and matching MERFISH sections to the
216 atlas space. Following preprocessing, two main alignment steps were performed: 1) 3-D global
217 affine mapping and section matching of the AllenCCFv3 into the MERFISH data and 2) 2-D
218 global and deformable mapping between each MERFISH section and matched AllenCCFv3
219 section. Mappings learned via each step in the pipeline are preserved and concatenated to
220 provide point-to-point correspondence between the original MERFISH data and AllenCCFv3,
221 thus allowing individual gene expressions to be transferred into the AllenCCFv3.

222 **Data.** MERFISH mouse brain data was acquired using a previously detailed procedure⁵⁰.
223 Briefly, a brain of C57BL/6 mouse was dissected according to standard procedures and placed
224 into an optimal cutting temperature (OCT) compound (Sakura FineTek 4583) in which it
225 was stored at -80°C. The fresh frozen brain was sectioned at 10 μm on Leica 3050 S cryostats
226 at intervals of 200 μm to evenly cover the brain. A set of 500 genes were imaged that had been
227 carefully chosen to distinguish the ~5200 clusters of our existing RNAseq taxonomy. For
228 staining the tissue with MERFISH probes, a modified version of instructions provided by the
229 manufacturer was used⁵⁰. Raw MERSCOPE data were decoded using Vizgen software (v231).
230 Cells were segmented based on DAPI and PolyT staining using Cellpose^{71,72}. Segmentation
231 was performed on a median z-plane (fourth out of seven) and cell borders were propagated to
232 z-planes above and below. To assign cluster identity to each cell in the MERFISH dataset,
233 we mapped the MERFISH cells to the scRNA-seq reference taxonomy.

234 **Evaluation.** Alignment of the MERFISH data into the AllenCCFv3 was qualitatively assessed
235 by an expert anatomist at each iteration of the registration using known correspondence of
236 gene markers and their associations with the AllenCCFv3. As previously reported⁵⁰, further
237 assessment of the alignment showed that, of the 554 terminal regions (gray matter only)

²³⁸ in the AllenCCFv3, only seven small subregions were missed from the MERFISH dataset:
²³⁹ frontal pole, layer 1 (FRP1), FRP2/3, FRP5; accessory olfactory bulb, glomerular layer
²⁴⁰ (AOBgl); accessory olfactory bulb, granular layer (AOBgr); accessory olfactory bulb, mitral
²⁴¹ layer (AOBmi); and accessory supraoptic group (ASO).

²⁴² 2.1.2 Mapping fluorescence micro-optical sectioning tomography (fMOST) data

²⁴³ **Overview.** We developed a pipeline for mapping fluorescence micro-optical sectioning
²⁴⁴ tomography (fMOST) mouse brain images into the AllenCCFv3 (see Figure 1(b)). The
²⁴⁵ pipeline is adapted from previously developed frameworks for human brain mapping⁵⁹, and
²⁴⁶ uses a modality specific (fMOST) average atlas to assist in the image registration and
²⁴⁷ mapping. This approach has been well validated in human studies^{73–75}, and successfully
²⁴⁸ used in other mouse data^{12,15,34}. Briefly, we construct an intensity- and shape-based average
²⁴⁹ fMOST atlas using 30 fMOST images to serve as an intermediate registration target for
²⁵⁰ mapping fMOST images from individual specimens into the AllenCCFv3. Preprocessing
²⁵¹ steps include downsampling to match the $25\mu m$ isotropic AllenCCFv3, acquisition-based
²⁵² stripe artifact removal, and inhomogeneity correction⁵². Preprocessing also includes a single
²⁵³ annotation-driven registration to establish a canonical mapping between the fMOST atlas and
²⁵⁴ the AllenCCFv3. This step allows us to align expert determined landmarks to accurately map
²⁵⁵ structures with large morphological differences between the modalities, which are difficult to
²⁵⁶ address using standard approaches. Once this canonical mapping is established, standard
²⁵⁷ intensity-based registration is used to align each new fMOST image to the fMOST specific
²⁵⁸ atlas. This mapping is concatenated with the canonical fMOST atlas-to-AllenCCFv3 mapping
²⁵⁹ to further map each individual brain into the latter without the need to generate additional
²⁶⁰ landmarks. Transformations learned through this mapping can be applied to single neuron
²⁶¹ reconstructions from the fMOST images to evaluate neuronal distributions across different
²⁶² specimens into the AllenCCFv3 for the purpose of cell census analyses.

²⁶³ **Data.** The high-throughput and high-resolution fluorescence micro-optical sectioning tomog-
²⁶⁴ raphy (fMOST)^{76,77} platform was used to image 55 mouse brains containing gene-defined
²⁶⁵ neuron populations, with sparse transgenic expression^{78,79}. In short, the fMOST imaging

platform results in 3-D images with voxel sizes of $0.35 \times 0.35 \times 1.0 \mu\text{m}^3$ and is a two-channel imaging system where the green channel displays the green fluorescent protein (GFP) labeled neuron morphology and the red channel is used to visualize the counterstained propidium iodide cytoarchitecture. The spatial normalizations described in this work were performed using the red channel, which offered higher tissue contrast for alignment, although other approaches are possible including multi-channel registration.

Evaluation. Evaluation of the canonical fMOST atlas to Allen CCFv3 mapping was performed via quantitative comparison at each step of the registration and qualitative assessment of structural correspondence after alignment by an expert anatomist. Dice values were generated for the following structures: whole brain, 0.99; fimbria, 0.91; habenular commissure, 0.63; posterior choroid plexus, 0.93; anterior choroid plexus, 0.96; optic chiasm, 0.77; caudate putamen, 0.97. Similar qualitative assessment was performed for each fMOST specimen including the corresponding neuron reconstruction data.

2.2 Continuously mapping the DevCCF developmental trajectory with a velocity flow model

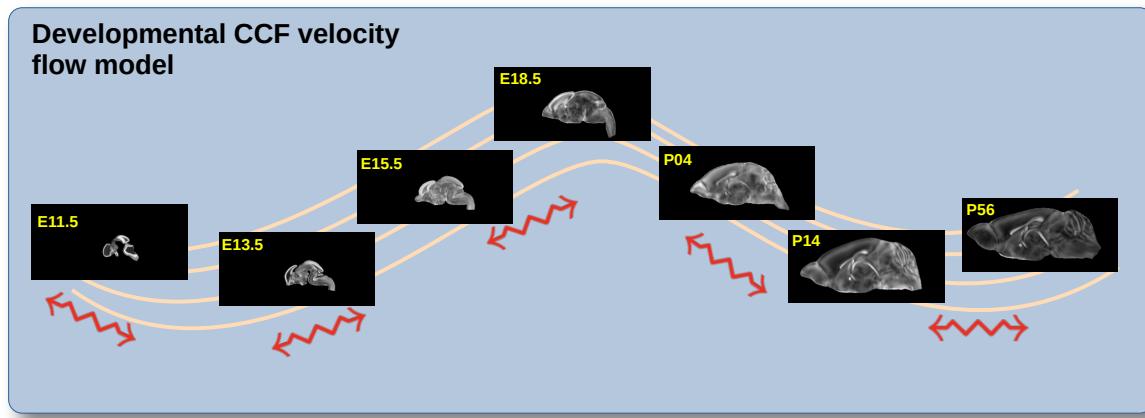


Figure 2: The spatial transformation between any two time points within the continuous DevCCF longitudinal developmental trajectory is available through the use of ANTsX functionality for generating a velocity flow model.

The DevCCF is an openly accessible resource for the mouse brain research community⁸⁰. It

282 consists of multi-modal MRI and LSFM symmetric ANTsX-generated templates⁵⁹ sampling
283 the mouse brain developmental trajectory, specifically the embryonic (E) and postnatal (P)
284 days E11.5, E13.5, E15.5, E18.5 P4, P14, and P56. Each template space includes structural
285 labels defined by a developmental ontology. Its utility is also enhanced by a coordinated
286 construction with AllenCCFv3. Although this work represents a significant contribution, the
287 gaps between time points potentially limit its applicability which could be addressed through
288 the development of the ability to map not only between time points but also within and
289 across time points.

290 To continuously generate transformations between the different stages of the DevCCF atlases,
291 we developed a general velocity flow model approach which we apply to DevCCF-derived
292 data. We also introduce this functionality into both the ANTsR and ANTsPy packages (for
293 the latter, see `ants.fit_time_varying_transform_to_point_sets(...)`) for potential
294 application to this and other analogous scenarios (e.g., modeling the cardiac and respiratory
295 cycles). ANTsX, being built on top of ITK, uses an ITK image data structure for the 4-D
296 velocity field where each voxel contains the x , y , z components of the field at that point.

297 2.2.1 Data

298 Labeled annotations are available as part of the original DevCCF and reside in the space
299 of each developmental template which range in resolution from $31.5 - 50\mu\text{m}$. Across all
300 atlases, the total number of labeled regions exceeds 2500. From these labels, a common set
301 of 26 labels (13 per hemisphere) across all atlases were used for optimization and evaluation.
302 These simplified regions include: terminal hypothalamus, subpallium, pallium, peduncular
303 hypothalamus, prosomere, prosomere, prosomere, midbrain, prepontine hindbrain, pontine
304 hindbrain, pontomedullary hindbrain, medullary hindbrain, and tracts (see Figure 3).

305 Prior to velocity field optimization, all data were rigidly transformed to DevCCF P56 using
306 the centroids of the common label sets. In order to determine the landmark correspondence
307 across DevCCF stages, the multi-metric capabilities of `ants.registration(...)` were used.
308 Instead of performing intensity-based pairwise registration directly on these multi-label
309 images, each label was used to construct a separate fixed and moving image pair resulting in

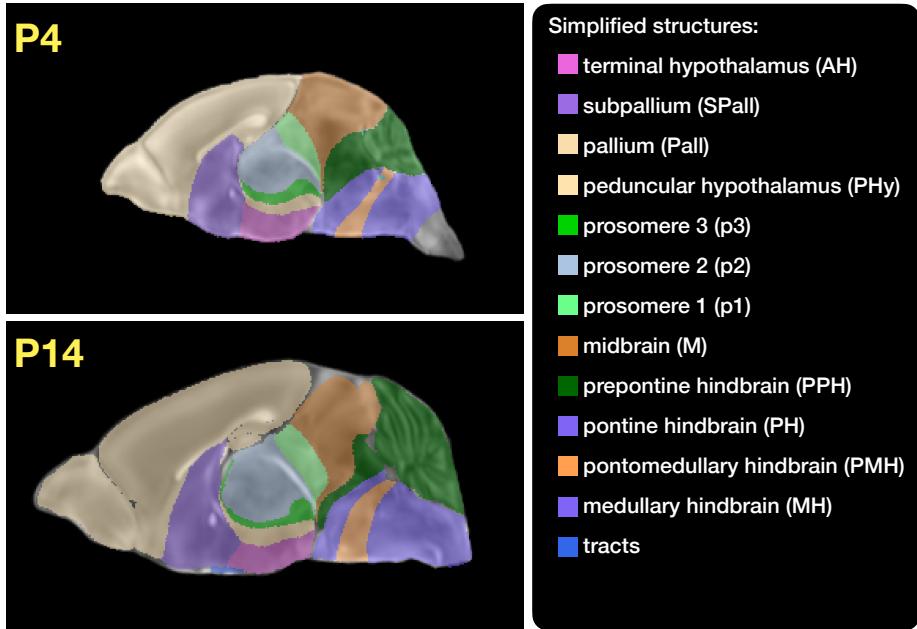


Figure 3: Annotated regions representing common labels across developmental stages which are illustrated for both P4 and P14.

310 a multi-metric registration optimization scenario involving 24 binary image pairs (each label
 311 weighted equally) for optimizing diffeomorphic correspondence between neighboring time
 312 point atlases using the mean squares metric and the symmetric normalization transform⁵¹.

313 To generate the set of common point sets across all seven developmental atlases, the label
 314 boundaries and whole regions were sampled in the P56 atlas and then propagated to each atlas
 315 using the transformations derived from the pairwise registrations. We selected a sampling
 316 rate of 10% for the contour points and 1% for the regional points for a total number of points
 317 being per atlas being 173303 ($N_{contour} = 98151$ and $N_{region} = 75152$). Regional boundary
 318 points were weighted twice as those of non-boundary points during optimization.

319 **2.2.2 Velocity field optimization**

320 The velocity field was optimized using the input composed of the seven corresponding point
 321 sets and their associated weight values, the selected number of integration points for the
 322 velocity field ($N = 11$), and the parameters defining the geometry of the spatial dimensions
 323 of the velocity field. Thus, the optimized velocity field described here is of size [256, 182, 360]

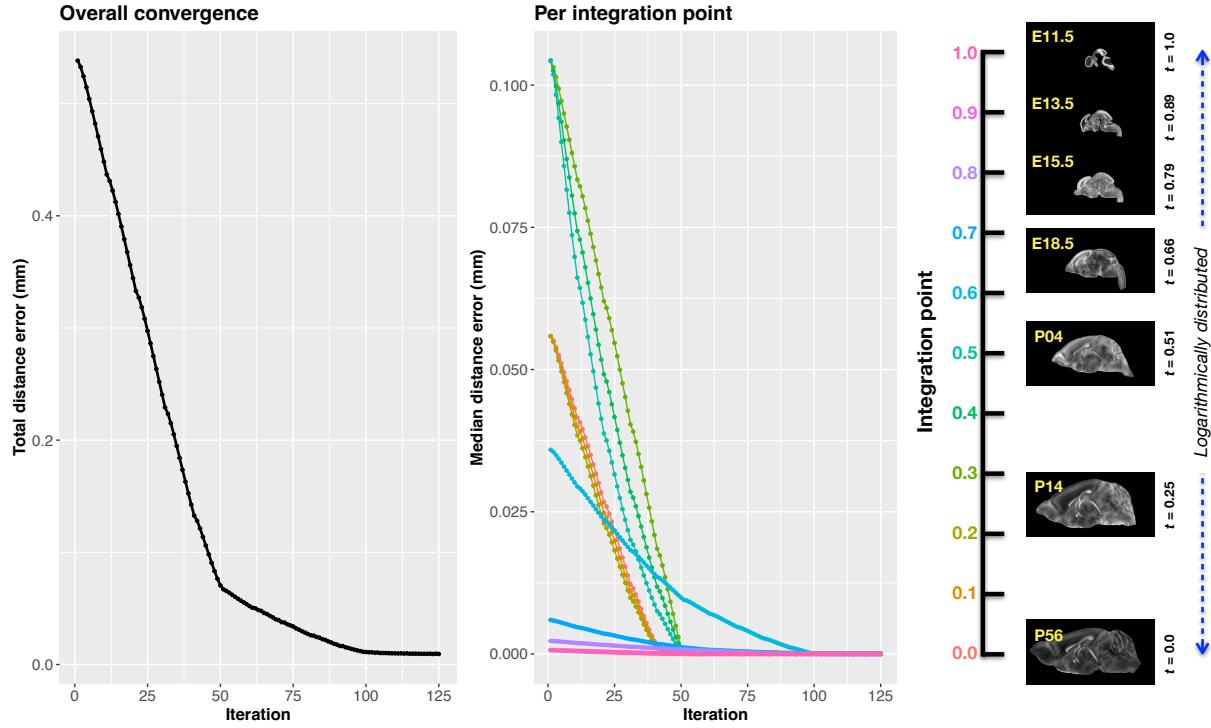


Figure 4: Convergence of the optimization of the velocity field for describing the transformation through the developmental stages from E11.5 through P56. Integration points in diagram on the right are color-coordinated with the center plot and placed in relation to the logarithmically situated temporal placement of the individual DevCCF atlases.

324 (50 μ m isotropic) \times 11 integration points for a total compressed size of a little over 2 GB.
 325 This choice represented weighing the trade-off between tractability, portability, and accuracy.
 326 However, all data and code to reproduce the results described are available in the dedicated
 327 GitHub repository.
 328 The normalized time point scalar value for each atlas/point-set in the temporal domains [0, 1]
 329 was also defined. Given the increasingly larger gaps in the postnatal time point sampling, we
 330 made two adjustments. Based on known mouse brain development, we used 28 days for the
 331 P56 data. We then computed the log transform of the adjusted set of time points prior to
 332 normalization between 0 and 1 (see the right side of Figure 4). This log transform, as part of
 333 the temporal normalization, significantly improves the temporal spacing of data.
 334 The maximum number of iterations was set to 200 with each iteration taking approximately
 335 six minutes on a 2020 iMac (processor, 3.6 GHz 10-Core Intel Core i9; memory, 64 GB 2667
 336 MHz DDR4) At each iteration we looped over the 11 integration points. At each integration

point, the velocity field estimate was updated by warping the two immediately adjacent
 point sets to the integration time point and determining the regularized displacement field
 between the two warped point sets. As with any gradient-based descent algorithm, this
 field was multiplied by a small step size ($\delta = 0.2$) before adding to the current velocity field.
 Convergence is determined by the average displacement error over each of the integration
 points. As can be seen in the left panel of Figure 4, convergence occurred around 125
 iterations when the average displacement error over all integration points is minimized. The
 median displacement error at each of the integration points also trends towards zero but at
 different rates.

346 2.2.3 The velocity flow transformation model

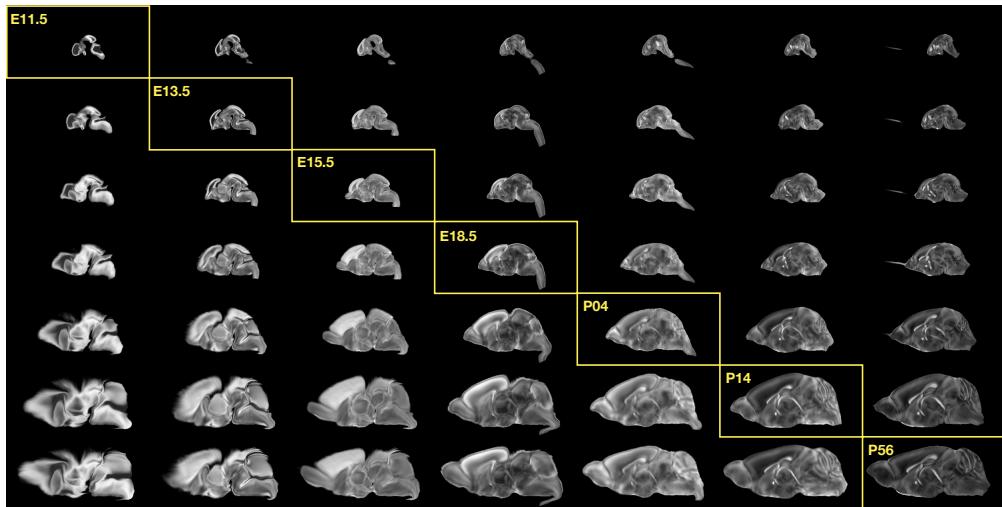


Figure 5: Mid-sagittal visualization of the effects of the transformation model in warping every developmental stage to the time point of every other developmental stage. The original images are located along the diagonal. Columns correspond to the warped original image whereas the rows represent the reference space to which each image is warped.

Once optimized, the resulting velocity field can be used to generate the deformable transform
 between any two continuous points within the time interval bounded by E11.5 and P56.
 As a demonstration, in Figure 5, we transform each atlas to the space of every other atlas
 using the DevCCF transform model. Additionally, one can use this transformation model to
 construct virtual templates in the temporal gaps of the DevCCF. Given an arbitrarily chosen
 time point within the normalized time point interval, the existing adjacent DevCCF atlases

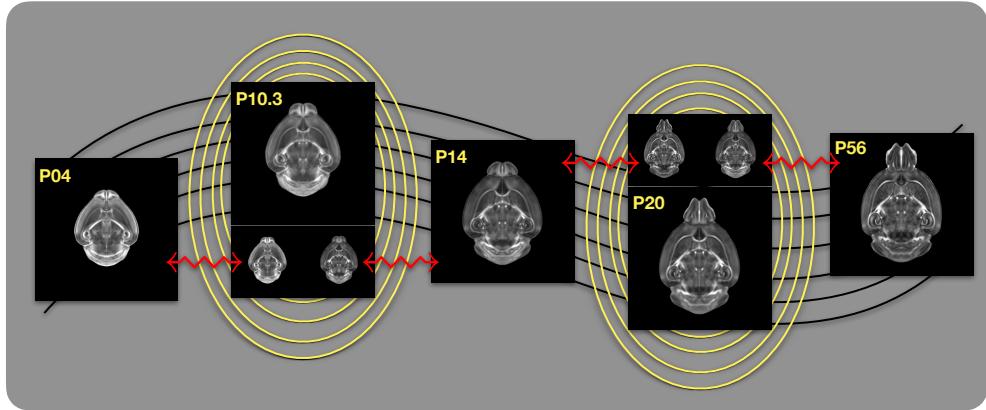


Figure 6: Illustration of the use of the velocity flow model for creating virtual templates at continuous time points not represented in one of the existing DevCCF time points. For example, FA templates at time point P10.3 and P20 can be generated by warping the existing temporally adjacent developmental templates to the target time point and using those images in the ANTsX template building process.

353 on either chronological side can be warped to the desired time point. A subsequent call to
 354 one of the ANTsX template building functions then permits the construction of the template
 355 at that time point. In Figure 6, we illustrate the use of the DevCCF velocity flow model for
 356 generating two such virtual templates for two arbitrary time points. Note that both of these
 357 usage examples can be found in the GitHub repository previously given.

358 2.3 Automated structural parcellations of the mouse brain

359 Brain parcellation strategies for the mouse brain are pivotal for understanding the complex
 360 organization and function of murine nervous system⁸¹. By dividing the brain into distinct
 361 regions based on anatomical, physiological, or functional characteristics, researchers can
 362 investigate specific areas in isolation and identify their roles in various behaviors and processes.
 363 For example, such parcellation schemes can help elucidate the spatial distribution of gene
 364 expression patterns⁸² as well as identify functional regions involved in specific cognitive
 365 tasks⁸³.

366 Although deep learning techniques have been used to develop useful parcellation tools for
 367 human brain research (e.g., SynthSeg⁸⁴, ANTsXNet⁴⁶), analogous development for the mouse
 368 brain is limited. In addition, mouse data is often characterized by unique imaging issues

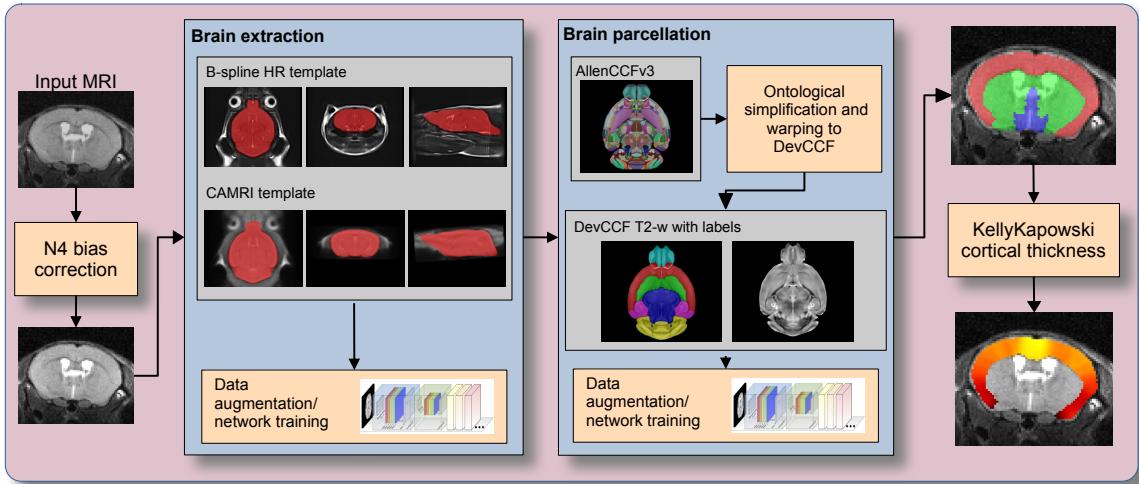


Figure 7: The mouse brain cortical parcellation pipeline integrating two deep learning components for brain extraction and brain parcellation prior to estimating cortical labels. Both deep learning networks rely heavily on aggressive data augmentation on templates built from open data and provide an outline for further refinement and creating alternative parcellations for tailored research objectives. Possible applications include voxelwise cortical thickness measurements.

such as extreme anisotropic sampling which are often in sharp contrast to the high resolution template-based resources available within the community, e.g., AllenCCFv3 and DevCCF. We demonstrate how one can use the ANTsX tools to develop a complete mouse brain structural morphology pipeline as illustrated in Figure 7 and detailed below.

2.3.1 Few-shot mouse brain extraction network

In order to create a generalized mouse brain extraction network, we built whole-head templates from two publicly available datasets. The Center for Animal MRI (CAMRI) dataset⁶⁸ from the University of North Carolina at Chapel Hill consists of 16 T2-w MRI volumes of voxel resolution $0.16 \times 0.16 \times 0.16 mm^3$. The second high-resolution dataset⁶⁹ comprises 88 specimens each with three spatially aligned canonical views with in-plane resolution of $0.08 \times 0.08 mm^2$ with a slice thickness of $0.5 mm$. These three orthogonal views were used to reconstruct a single high-resolution volume per subject using a B-spline fitting algorithm available in ANTsX⁸⁵.

From these two datasets, two ANTsX templates⁵⁹ were generated. Bias field simulation,

383 intensity histogram warping, noise simulation, random translation and warping, and random
384 anisotropic resampling in the three canonical directions were used for data augmentation
385 in training an initial T2-w brain extraction network. This network was posted and the
386 corresponding functionality was immediately made available within ANTsXNet, similar to
387 our previous contributions to the community.

388 User interest led to a GitHub inquiry regarding possible study-specific improvements (<https://github.com/ANTsX/ANTsPyNet/issues/133>). This interaction led to the offering of a
389 user-made third template and extracted brain mask generated from T2-w ex-vivo data with
390 isotropic spacing of 0.08 mm in each voxel dimension. This third template, in conjunction
391 with the other two, were used with the same aggressive data augmentation to refine the
392 network weights which were subsequently posted and made available through ANTsPyNet
393 using the function `antspynet.mouse_brain_extraction(...)`.

395 2.3.2 Single-shot mouse brain parcellation network

396 AllenCCFv3 and its hierarchical ontological labeling, along with the DevCCF, provides the
397 necessary data for developing a tailored structural parcellation network for multi-modal
398 imaging. The `allensdk` Python library permits the creation of any gross parcellation based
399 on the AllenCCFv3 ontology. Specifically, using `allensdk` we coalesced the labels to the
400 following six major structures: cerebral cortex, cerebral nuclei, brain stem, cerebellum, main
401 olfactory bulb, and hippocampal formation. This labeling was mapped to the P56 component
402 of the DevCCF for use with the T2-w template component.

403 The T2-w P56 DevCCF and labelings, in conjunction with the data augmentation
404 described previously for brain extraction, were used to train the proposed brain
405 parcellation network. This is available in ANTsXNet (e.g. in ANTsPyNet using
406 `antspynet.mouse_brain_parcellation(...)`). Note that other brain parcellation net-
407 works have also been trained using alternative regions and parcellation schemes and are
408 available in the same ANTsXNet functionality. One usage note is that the data augmentation
409 used to train the network permits a learned interpolation in 0.08 mm isotropic space. Since
410 the training data is isotropic and data augmentation includes downsampling in the canonical

411 directions, each of the two networks learns mouse brain-specific interpolation such that
 412 one can perform prediction on thick-sliced images, as, for example, in these evaluation
 413 data, and return isotropic probability and thickness maps (a choice available to the user).
 414 This permits robust cortical thickness estimation even in the case of anisotropic data (see
 415 `antspynet.mouse_cortical_thickness(...)`).

416 **2.3.3 Evaluation**

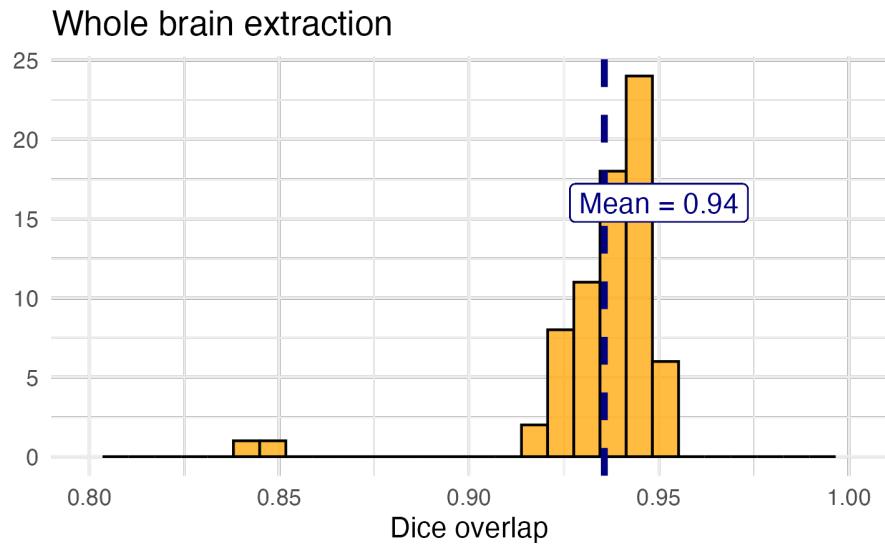
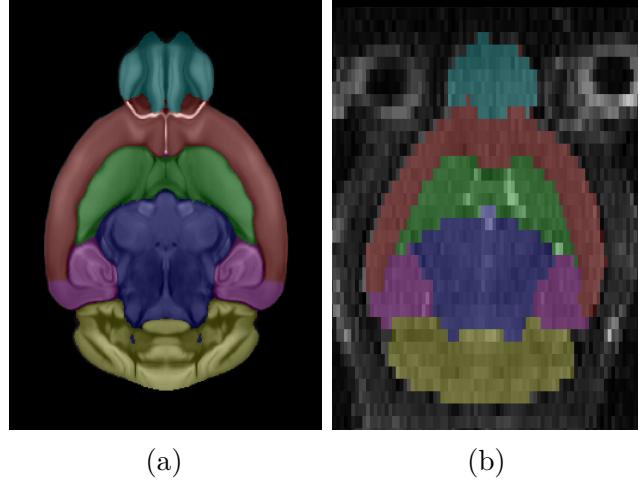


Figure 8: Evaluation of the ANTsX mouse brain extraction on an independent, publicly available dataset consisting of 12 specimens \times 7 time points = 84 total images. Dice overlap comparisons with the user-generated brain masks provide good agreement with the automated results from the brain extraction network.

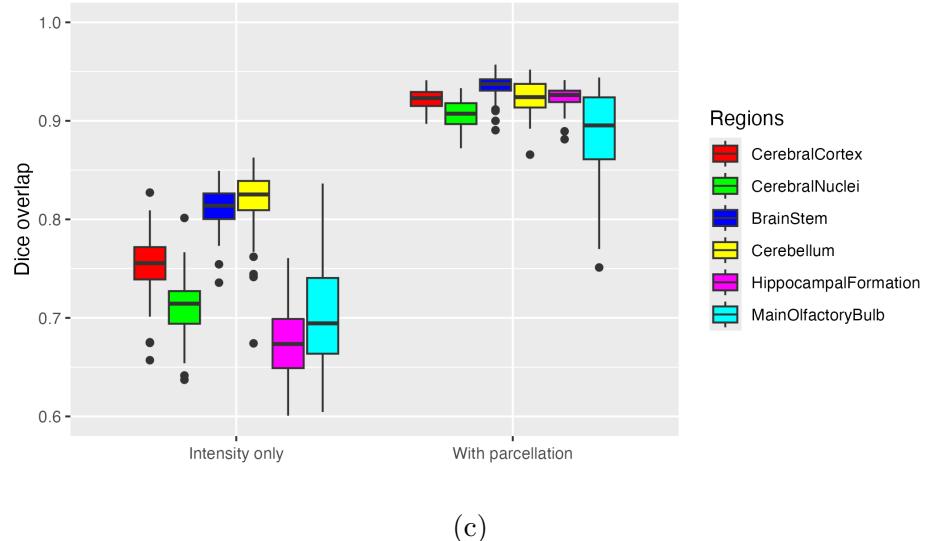
417 For evaluation, we used an additional publicly available dataset⁷⁰ that is completely indepen-
 418 dent from the data used in training the brain extraction and parcellation networks. Data
 419 includes 12 specimens each imaged at seven time points (Day 0, Day 3, Week 1, Week 4,
 420 Week 8, Week 20) with in-house-generated brain masks for a total of 84 images. Spacing is
 421 anistropic with an in-plane resolution of $0.1 \times 0.1 mm^2$ and a slice thickness of $0.5 mm$.
 422 Figure 8 summarizes the whole brain overlap between the provided segmentations for all
 423 84 images and the results of applying the proposed network. Also, since mapping to the
 424 AllenCCFv3 atlas is crucial for many mouse studies, we demonstrate the utility of the second



(a)

(b)

Normalization to AllenCCFv3



(c)

Figure 9: Evaluation of the ANTsX mouse brain parcellation on the same dataset. (a) T2-w DevCCF P56 with the described parcellation consisting of the cerebral cortex, nuclei, brain stem, cerebellum, main olfactory bulb, and hippocampal formation. (b) Sample subject (NR5 Day 0) with the proposed deep learning-based segmentation. (c) Dice overlap for comparing the regional alignments between registration using intensity information only and using intensity with the given parcellation scheme.

425 network by leveraging the labeled regions to perform anatomically-explicit alignment using
426 ANTsX multi-component registration instead of intensity-only registration. For these data,
427 the whole brain extraction demonstrates excellent performance across the large age range.
428 And although the intensity-only image registration provides adequate alignment, intensity
429 with the regional parcellations significantly improves those measures.

430 **3 Discussion**

431 The diverse mouse brain cell type profiles gathered through BICCN and associated efforts
432 provides a rich multi-modal resource to the research community. However, despite significant
433 progress, optimized leveraging of these valuable resources is ongoing. A central component
434 to data integration is accurately mapping novel cell type data into CCFs for subsequent
435 processing and analysis. To meet these needs, tools for mapping mouse cell type data must be
436 both generally accessible to a wide audience of investigators, and capable of handling distinct
437 challenges unique to each data type. In this work, we described modular ANTsX-based
438 pipelines developed to address the needs of three BICCN projects that cover distinct cell
439 type data, including spatial transcriptomic, morphological, and developmental data. We
440 highlighted how a modular toolbox like ANTsX can be tailored to address problems unique
441 to each modality through leveraging a variety of ready-to-use powerful tools that have been
442 previously validated in multiple application scenarios.

443 Our MERFISH pipeline provides an example of how to map high-resolution spatial tran-
444 scriptomic data into the AllenCCFv3. While the techniques employed for mapping the
445 sectioned data can be generally applicable to spatially transform other serial histology images,
446 much of the pipeline was designed to specifically address known alignment challenges in the
447 MERFISH data. Thus pipeline shows how general ANTsX tools can be adapted to target
448 highly specialized problems in mouse cell type data.

449 In contrast to the MERFISH pipeline, our fMOST pipeline was designed to be a more general
450 solution that can be employed in other modalities. The pipeline primarily uses previously
451 developed ANTsX preprocessing and atlasing tools to map fMOST data into the AllenCCFv3.
452 The key component of the pipeline is the use of a fMOST-specific average shape and intensity
453 atlas to most effectively address image registration in this context. The mapping between
454 the fMOST atlas is generated once and reused for each new fMOST image. Lastly, ANTsX
455 provides point set transformation tools to allow the mappings found through the pipeline to
456 be directly applied to associated single-cell reconstructions from the fMOST data to study
457 neuronal morphology.

458 The pipeline for continuously mapping the DevCCF data is also available in ANTsX and is
459 generally applicable for spatio-temporal mapping. With specific application to the DevCCF,
460 despite the significant expansion of available developmental age templates beyond what existed
461 previously, there are still temporal gaps in the DevCCF which can be potentially sampled
462 by future research efforts. However, pioneering work involving time-varying diffeomorphic
463 transformations allow us to continuously situate the existing templates within a velocity
464 flow model. This allows one to determine the diffeomorphic transformation from any one
465 temporal location to any other temporal location within the time span defined by the temporal
466 limits of the DevCCF. This functionality is built on multiple ITK components including the
467 B-spline scattered data approximation technique for field regularization and velocity field
468 integration. This velocity field model permits intra-template comparison and the construction
469 of virtual templates where a template can be estimated at any continuous time point within
470 the temporal domain. This novel application can potentially enhance our understanding of
471 intermediate developmental stages.

472 We also presented a mouse brain morphological pipeline for brain extraction and brain
473 parcellation using single-shot and few-shot learning with aggressive data augmentation. This
474 approach attempts to circumvent (or at least minimize) the typical requirement of large
475 training datasets as with the human ANTsX pipeline analog. However, even given our initial
476 success on independent data, we anticipate that refinements will be necessary. Given that the
477 ANTsX toolkit is a dynamic effort undergoing continual improvement, we manually correct
478 cases that fail and use them for future training and refinement of network weights as we have
479 done for our human-based networks. And, as demonstrated, we welcome contributions from
480 the community for improving these approaches which, generally, provide a way to bootstrap
481 training data for manual refinement and future generation of more accurate deep learning
482 networks in the absence of other applicable tools.

483 The ANTsX ecosystem is a powerful framework that has demonstrated applicability to
484 diverse cell type data in the mouse brain. This is further evidenced by the many software
485 packages that use various ANTsX components in their own mouse-specific workflows. The
486 extensive functionality of ANTsX makes it possible to create complete processing pipelines

⁴⁸⁷ without requiring the integration of multiple packages or lengthy software development. These
⁴⁸⁸ open-source components not only perform well but are available across multiple platforms
⁴⁸⁹ which facilitates the construction of tailored pipelines for individual study solutions. These
⁴⁹⁰ components are also supported by years of development not only by the ANTsX development
⁴⁹¹ team but by the larger ITK community.

492 **4 Methods**

493 The following methods are all available as part of the ANTsX ecosystem with analogous
494 elements existing in both ANTsR (ANTs in R) and ANTsPy (ANTs in Python) with an
495 ANTs/ITK C++ core. However, most of the development for the work described below was
496 performed using ANTsPy. For equivalent calls in ANTsR, please see the ANTsX tutorial at
497 <https://tinyurl.com/antsxtutorial>.

498 **4.1 General ANTsX utilities**

499 Although they focus on distinct data types, the three pipelines presented share common
500 components that are generally applicable when mapping mouse cell type data. These include,
501 addressing intensity biases and noise in the data, image registration to solve the mapping,
502 creating custom templates and atlases from the data, and visualization of the results. Table
503 1 provides a brief summary of key general functionalities in ANTsX for addressing these
504 challenges.

505 **4.1.1 Preprocessing: bias field correction and denoising**

506 Bias field correction and image denoising are standard preprocessing steps in improving
507 overall image quality in mouse brain images. The bias field, a gradual spatial intensity
508 variation in images, can arise from various sources such as magnetic field inhomogeneity or
509 acquisition artifacts, leading to distortions that can compromise the quality of brain images.
510 Correcting for bias fields ensures a more uniform and consistent representation of brain
511 structures, enabling more accurate quantitative analysis. Additionally, brain images are
512 often susceptible to various forms of noise, which can obscure subtle features and affect the
513 precision of measurements. Denoising techniques help mitigate the impact of noise, enhancing
514 the signal-to-noise ratio and improving the overall image quality. The well-known N4 bias
515 field correction algorithm⁵² has its origins in the ANTs toolkit which was implemented and
516 introduced into the ITK toolkit, i.e. `ants.n4_bias_field_correction(...)`. Similarly,

Table 1: Sampling of ANTsX functionality

<i>ANTsPy: Preprocessing</i>	
bias field correction	<code>n4_bias_field_correction(...)</code>
image denoising	<code>denoise_image(...)</code>
<i>ANTsPy: Registration</i>	
image registration	<code>registration(...)</code>
image transformation	<code>apply_transforms(...)</code>
template generation	<code>build_template(...)</code>
landmark registration	<code>fit_transform_to_paired_points(...)</code>
time-varying landmark reg.	<code>fit_time_varying_transform_to_point_sets(...)</code>
integrate velocity field	<code>integrate_velocity_field(...)</code>
invert displacement field	<code>invert_displacement_field(...)</code>
<i>ANTsPy: Segmentation</i>	
MRF-based segmentation	<code>atropos(...)</code>
Joint label fusion	<code>joint_label_fusion(...)</code>
diffeomorphic thickness	<code>kelly_kapowski(...)</code>
<i>ANTsPy: Miscellaneous</i>	
Regional intensity statistics	<code>label_stats(...)</code>
Regional shape measures	<code>label_geometry_measures(...)</code>
B-spline approximation	<code>fit_bspline_object_to_scattered_data(...)</code>
Visualize images and overlays	<code>plot(...)</code>
<i>ANTsPyNet: Mouse-specific</i>	
brain extraction	<code>mouse_brain_extraction(...modality="t2"...)</code>
brain parcellation	<code>mouse_brain_parcellation(...)</code>
cortical thickness	<code>mouse_cortical_thickness(...)</code>
super resolution	<code>mouse_histology_super_resolution(...)</code>

ANTsX provides state-of-the-art functionality for processing biomedical image data. Such tools, including deep learning networks, support a variety of mapping-related tasks. A more comprehensive listing of ANTsX tools with self-contained R and Python examples is provided as a gist page on GitHub (<https://tinyurl.com/antsxtutorial>).

517 ANTsX contains an implementation of a well-performing patch-based denoising technique⁶¹
518 and is also available as an image filter to the ITK community, `ants.denoise_image(...)`.

519 4.1.2 Image registration

520 The ANTs registration toolkit is a complex framework permitting highly tailored solutions
521 to pairwise image registration scenarios⁸⁶. It includes innovative transformation models
522 for biological modeling^{51,67} and has proven capable of excellent performance^{56,87}. Vari-
523 ous parameter sets targeting specific applications have been packaged with the different
524 ANTsX packages, specifically ANTs, ANTsPy, and ANTsR⁴⁶. In ANTsPy, the function
525 `ants.registration(...)` is used to register a pair of images or a pair of image sets where
526 `type_of_transform` is a user-specified option that invokes a specific parameter set. For exam-
527 ple `type_of_transform='antsRegistrationSyNQuick[s]'` encapsulates an oft-used param-
528 eter set for quick registration whereas `type_of_transform='antsRegistrationSyN[s]'` is
529 a more aggressive alternative. Transforming images using the derived transforms is performed
530 via the `ants.apply_transforms(...)` function.

531 Initially, linear optimization is initialized with center of (intensity) mass alignment typically
532 followed by optimization of both rigid and affine transforms using the mutual information
533 similarity metric. This is followed by diffeomorphic deformable alignment using symmetric
534 normalization (SyN) with Gaussian⁵¹ or B-spline regularization⁶⁷ where the forward transform
535 is invertible and differentiable. The similarity metric employed at this latter stage is typically
536 either neighborhood cross-correlation or mutual information. Note that these parameter
537 sets are robust to input image type (e.g., light sheet fluorescence microscopy, Nissl staining,
538 and the various MRI modalities) and are adaptable to mouse image geometry and scaling.
539 Further details can be found in the various documentation sources for these ANTsX packages.

540 4.1.3 Template generation

541 ANTsX provides functionality for constructing templates from a set (or multi-modal sets) of
542 input images as originally described⁵⁹ and recently used to create the DevCCF templates¹⁶.

543 An initial template estimate is constructed from an existing subject image or a voxelwise
544 average derived from a rigid pre-alignment of the image population. Pairwise registration
545 between each subject and the current template estimate is performed using the Symmetric
546 Normalization (SyN) algorithm⁵¹. The template estimate is updated by warping all subjects
547 to the space of the template, performing a voxelwise average, and then performing a “shape
548 update” of this latter image by warping it by the average inverse deformation, thus yielding
549 a mean image of the population in terms of both intensity and shape. The corresponding
550 ANTsPy function is `ants.build_template(...)`.

551 4.1.4 Visualization

552 To complement the well-known visualization capabilities of R and Python, e.g., `ggplot2`
553 and `matplotlib`, respectively, image-specific visualization capabilities are available in the
554 `ants.plot(...)` function (Python). These are capable of illustrating multiple slices in
555 different orientations with other image overlays and label images.

556 4.2 Mapping fMOST data to AllenCCFv3

557 4.2.1 Preprocessing

- 558 • *Downsampling*. The first challenge when mapping fMOST images into the AllenCCFv3
559 is addressing the resolution scale of the data. Native fMOST data from an individual
560 specimen can range in the order of terabytes, which leads to two main problems. First,
561 volumetric registration methods (particularly those estimating local deformation) have
562 high computational complexity and typically cannot operate on such high-resolution
563 data under reasonable memory and runtime constraints. Second, the resolution of
564 the AllenCCFv3 atlas is much lower than the fMOST data, thus the mapping process
565 will cause much of the high-resolution information in the fMOST images to be lost
566 regardless. Thus, we perform a cubic B-spline downsampling of the fMOST data to
567 reduce the resolution of each image to match the isotropic $25 \mu\text{m}$ voxel resolution of the
568 AllenCCFv3 intensity atlas using `ants.resample_image(...)`. An important detail

569 to note is that while the fMOST images and atlas are downsampled, the mapping
570 learned during the registration is assumed to be continuous. Thus, after establishing
571 the mapping to the AllenCCFv3, we can interpolate the learned mapping and apply it
572 directly to the high-resolution native data directly to transform any spatially aligned
573 data (such as the single-cell neuron reconstructions) into the AllenCCFv3.

- 574 • *Stripe artifact removal.* Repetitive pattern artifacts are a common challenge in fMOST
575 imaging where inhomogeneity during the cutting and imaging of different sections can
576 leave stripes of hyper- and hypo-intensity across the image. These stripe artifacts
577 can be latched onto by the registration algorithm as unintended features that are
578 then misregistered to non-analogous structures in the AllenCCFv3. We address these
579 artifacts by fitting a 3-D bandstop (notch) filter to target the frequency of the stripe
580 patterns and removing them prior to the image registration.
- 581 • *Inhomogeneity correction.* Regional intensity inhomogeneity can also occur within
582 and between sections in fMOST imaging due to staining or lighting irregularity during
583 acquisition. Similar to stripe artifacts, intensity gradients due to inhomogeneity
584 can be misconstrued as features during the mapping and result in matching of non-
585 corresponding structures. Our pipeline addresses these intensity inhomogeneities using
586 N4 bias field correction⁵², `ants.n4_bias_field_correction(...)`.

587 4.2.2 Steps for spatial normalization to AllenCCFv3

- 588 1. *Average fMOST atlas as an intermediate target.* Due to the preparation of the mouse
589 brain for fMOST imaging, the resulting structure in the mouse brain has several large
590 morphological deviations from the AllenCCFv3 atlas. Most notable of these is an
591 enlargement of the ventricles, and compression of cortical structures. In addition,
592 there is poor intensity correspondence for the same anatomic features due to intensity
593 dissimilarity between imaging modalities. We have found that standard intensity-base
594 registration is insufficient to capture the significant deformations required to map these
595 structures correctly into the AllenCCFv3. We address this challenge in ANTsX by using
596 explicitly corresponding parcellations of the brain, ventricles and surrounding structures

597 to directly recover these large morphological differences. However, generating these
598 parcellations for each individual mouse brain is a labor-intensive task. Our solution
599 is to create an average atlas whose mapping to AllenCCFv3 encapsulates these large
600 morphological differences to serve as an intermediate registration point. This has the
601 advantage of only needing to generate one set of corresponding annotations which is
602 used to register between the two atlas spaces. New images are first aligned to the
603 fMOST average atlas, which shares common intensity and morphological features and
604 thus can be achieved through standard intensity-based registration.

605 2. *Average fMOST atlas construction.* An intensity and shape-based contralaterally
606 symmetric average of the fMOST image data is constructed from 30 images and their
607 contralateral flipped versions. We ran three iterations of the atlas construction using the
608 default settings. Additional iterations (up to six) were evaluated and showed minimal
609 changes to the final atlas construction, suggesting a convergence of the algorithm.

610 3. *fMOST atlas to AllenCCFv3 alignment.* Alignment between the fMOST average atlas
611 and AllenCCFv3 was performed using a one-time annotation-driven approach. Label-
612 to-label registration is used to align 7 corresponding annotations in both atlases in the
613 following: 1) brain mask/ventricles, 2) caudate/putamen, 3) fimbria, 4) posterior choroid
614 plexus, 5) optic chiasm, 6) anterior choroid plexus, and 7) habenular commissure. The
615 alignments were performed sequentially, with the largest, most relevant structures being
616 aligned first using coarse registration parameters, followed by other structures using
617 finer parameters. This coarse-to-fine approach allows us to address large morphological
618 differences (such as brain shape and ventricle expansion) at the start of registration
619 and then progressively refine the mapping using the smaller structures. The overall
620 ordering of these structures was determined manually by an expert anatomist, where
621 anatomical misregistration after each step of the registration was evaluated and used to
622 determine which structure should be used in the subsequent iteration to best improve
623 the alignment. The transformation from this one-time expert-guided alignment is
624 preserved and used as the canonical fMOST atlas to AllenCCFv3 mapping in the
625 pipeline.

- 626 4. *Alignment of individual fMOST mouse brains.* The canonical transformation between
627 the fMOST atlas and AllenCCFv3 greatly simplifies the registration of new individual
628 fMOST mouse brains into the AllenCCFv3. Each new image is first registered into the
629 fMOST average atlas, which shares intensity, modality, and morphological characteris-
630 tics. This allows us to leverage standard, intensity-based registration functionality⁸⁶
631 available in ANTsX to perform this alignment. Transformations are then concate-
632 nated to the original fMOST image to move it into the AllenCCFv3 space using
633 `ants.apply_transforms(...)`.
- 634 5. *Transformation of single cell neurons.* A key feature of fMOST imaging is the ability
635 to reconstruct and examine whole-brain single neuron projections⁷⁹. Spatial mapping
636 of these neurons from individual brains into the AllenCCFv3 allows investigators to
637 study different neuron types within the same space and characterize their morphology
638 with respect to their transcriptomics. Mappings found between the fMOST image
639 and the AllenCCFv3 using our pipeline can be applied in this way to fMOST neuron
640 reconstruction point set data using `ants.apply_transforms_to_points(..)`.

641 4.3 Mapping MERFISH data to AllenCCFv3

642 4.3.1 Preprocessing

- 643 • *Initial volume reconstruction.* Alignment of MERFISH data into a 3-D atlas space
644 requires an estimation of anatomical structure within the data. For each section,
645 this anatomic reference image was created by aggregating the number of detected
646 genetic markers (across all probes) within each pixel of a $10 \times 10 \mu\text{m}^2$ grid to match
647 the resolution of the $10 \mu\text{m}$ AllenCCFv3 atlas. These reference image sections are then
648 coarsely reoriented and aligned across sections using manual annotations of the most
649 dorsal and ventral points of the midline. The procedure produces an anatomic image
650 stack that serves as an initialization for further global mappings into the AllenCCFv3.
- 651 • *Anatomical correspondence labeling.* Mapping the MERFISH data into the AllenCCFv3
652 requires us to establish correspondence between the anatomy depicted in the MERFISH

and AllenCCFv3 data. Intensity-based features in MERFISH data are not sufficiently apparent to establish this correspondence, so we need to generate instead corresponding anatomical labelings of both images with which to drive registration. These labels are already available as part of the AllenCCFv3; thus, the main challenge is deriving analogous labels from the spatial transcriptomic maps of the MERFISH data. Toward this end, we assigned each cell from the scRNA-seq dataset to one of the following major regions: cerebellum, CTXsp, hindbrain, HPF, hypothalamus, isocortex, LSX, midbrain, OLF, PAL, sAMY, STRd, STRv, thalamus and hindbrain. A label map of each section was generated for each region by aggregating the cells assigned to that region within a $10 \times 10\mu m^2$ grid. The same approach was used to generate more fine grained region specific landmarks (i.e., cortical layers, habenula, IC). Unlike the broad labels which cover large swaths of the section these regions are highly specific to certain parts of the section. Once cells in the MERFISH data are labeled, morphological dilation is used to provide full regional labels for alignment into the AllenCCFv3.

- *Section matching.* Since the MERFISH data is acquired as sections, its 3-D orientation may not be fully accounted for during the volume reconstruction step, due to the particular cutting angle. This can lead to obliqueness artifacts in the section where certain structures can appear to be larger or smaller, or missing outright from the section. To address this, we first use a global alignment to match the orientations of the MERFISH sections to the atlas space. In our pipeline, this section matching is performed in the reverse direction by performing a global affine transformation of the AllenCCFv3 into the MERFISH data space, and then resampling digital sections from the AllenCCFv3 to match each MERFISH section. This approach limits the overall transformation and thus resampling that is applied to the MERFISH data, and, since the AllenCCFv3 is densely sampled, it also reduces in-plane artifacts that result from missing sections or undefined spacing in the MERFISH data.

679 **4.3.2 2.5D deformable, landmark-driven alignment to AllenCCFv3**

680 After global alignment of the AllenCCFv3 into the MERFISH dataset, 2D per-section
681 deformable refinements are used to address local differences between the MERFISH sections
682 and the resampled AllenCCFv3 sections. Nine registrations were performed in sequence using
683 a single label at each iteration in the following order: 1) brain mask, 2) isocortex (layer
684 2+3), 3) isocortex (layer 5), 4) isocortex (layer 6), 5) striatum, 6) medial habenula, 7) lateral
685 habenula, 8) thalamus, and 9) hippocampus. This ordering was determined empirically by an
686 expert anatomist who prioritized which structure to use in each iteration by evaluating the
687 anatomical alignment from the previous iteration. Global and local mappings are then all
688 concatenated (with appropriate inversions) to create the final mapping between the MERFISH
689 data and AllenCCFv3. This mapping is then used to provide a point-to-point correspondence
690 between the original MERFISH coordinate space and the AllenCCFv3 space, thus allowing
691 mapping of individual genes and cell types located in the MERFISH data to be directly
692 mapped into the AllenCCFv3.

693 **4.4 DevCCF velocity flow transformation model**

694 Given multiple, linearly or non-linearly ordered point sets where individual points across the
695 sets are in one-to-one correspondence, we developed an approach for generating a velocity
696 flow transformation model to describe a time-varying diffeomorphic mapping as a variant of
697 the landmark matching solution. Integration of the resulting velocity field can then be used
698 to describe the displacement between any two time points within this time-parameterized
699 domain. Regularization of the sparse correspondence between point sets is performed using a
700 generalized B-spline scattered data approximation technique⁸⁵, also created by the ANTsX
701 developers and contributed to ITK.

702 **4.4.1 Velocity field optimization**

703 To apply this methodology to the developmental templates¹⁶, we coalesced the manual
704 annotations of the developmental templates into 26 common anatomical regions (see Figure

705 3). We then used these regions to generate invertible transformations between successive time
706 points. Specifically each label was used to create a pair of single region images resulting in 26
707 pairs of “source” and “target” images. The multiple image pairs were simultaneously used to
708 iteratively estimate a diffeomorphic pairwise transform. Given the seven atlases E11.5, E13.5,
709 E15.5, E18.5, P4, P14, and P56, this resulted in 6 sets of transforms between successive time
710 points. Approximately 10^6 points were randomly sampled labelwise in the P56 template
711 space and propagated to each successive atlas providing the point sets for constructing the
712 velocity flow model. Approximately 125 iterations resulted in a steady convergence based on
713 the average Euclidean norm between transformed point sets. Ten integration points were
714 used and point sets were distributed along the temporal dimension using a log transform for
715 a more evenly spaced sampling. For additional information a help menu is available for the
716 ANTsPy function `ants.fit_time_varying_transform_to_point_sets(...)`.

717 4.5 ANTsXNet mouse brain applications

718 4.5.1 General notes regarding deep learning training

719 All network-based approaches described below were implemented and organized in the
720 ANTsXNet libraries comprising Python (ANTsPyNet) and R (ANTsRNet) analogs using the
721 Keras/Tensorflow libraries available as open-source in ANTsX GitHub repositories. For the
722 various applications, both share the identically trained weights for mutual reproducibility.
723 For all GPU training, we used Python scripts for creating custom batch generators which
724 we maintain in a separate GitHub repository for public availability (<https://github.com/ntustison/ANTsXNetTraining>). These scripts provide details such as batch size, choice of
725 loss function, and network parameters. In terms of GPU hardware, all training was done on
726 a DGX (GPUs: 4X Tesla V100, system memory: 256 GB LRDIMM DDR4).

728 Data augmentation is crucial for generalizability and accuracy of the trained networks.
729 Intensity-based data augmentation consisted of randomly added noise (i.e., Gaussian, shot,
730 salt-and-pepper), simulated bias fields based on N4 bias field modeling, and histogram warping
731 for mimicking well-known MRI intensity nonlinearities^{46,88}. These augmentation techniques

732 are available in ANTsXNet (only ANTsPyNet versions are listed with ANTsRNet versions
733 available) and include:

- 734 • image noise: `ants.add_noise_to_image(...)`,
- 735 • simulated bias field: `antspynet.simulate_bias_field(...)`, and
- 736 • nonlinear intensity warping: `antspynet.histogram_warp_image_intensities(...)`.

737 Shape-based data augmentation used both random linear and nonlinear deformations in
738 addition to anisotropic resampling in the three canonical orientations to mimic frequently
739 used acquisition protocols for mice brains:

- 740 • random spatial warping: `antspynet.randomly_transform_image_data(...)` and
- 741 • anisotropic resampling: `ants.resample_image(...)`.

742 **4.5.2 Brain extraction**

743 Similar to human neuroimage processing, brain extraction is a crucial preprocessing step for
744 accurate brain mapping. We developed similar functionality for T2-weighted mouse brains.

745 This network uses a conventional U-net architecture⁸⁹ and, in ANTsPyNet, this functionality is
746 available in the program `antspynet.mouse_brain_extraction(...)`. For the two-shot T2-
747 weighted brain extraction network, two brain templates were generated along with their masks.

748 One of the templates was generated from orthogonal multi-plane, high resolution data⁶⁹ which
749 were combined to synthesize isotropic volumetric data using the B-spline fitting algorithm⁸⁵.

750 This algorithm is encapsulated in `ants.fit_bspline_object_to_scattered_data(...)`
751 where the input is the set of voxel intensity values and each associated physical location.

752 Since each point can be assigned a confidence weight, we use the normalized gradient value
753 to more heavily weight edge regions. Although both template/mask pairs are available in the
754 GitHub repository associated with this work, the synthesized volumetric B-spline T2-weighted
755 pair is available within ANTsXNet through the calls:

- 756 • template: `antspynet.get_antsxnet_data("bsplineT2MouseTemplate")` and
757 • mask: `antspynet.get_antsxnet_data("bsplineT2MouseTemplateBrainMask")`.

758 **4.5.3 Brain parcellation**

759 The T2-weighted brain parcellation network is also based on a 3-D U-net architecture and
760 the T2-w DevCCF P56 template component with extensive data augmentation, as described
761 previously. Intensity differences between the template and any brain extracted input image
762 are minimized through the use of the rank intensity transform (`ants.rank_intensity(...)`).
763 Shape differences are reduced by the additional preprocessing step of warping the brain
764 extracted input image to the template. Additional input channels include the prior probability
765 images created from the template parcellation. These images are also available through the
766 ANTsXNet `get_antsxnet_data(...)` interface.

⁷⁶⁷ **Data availability**

⁷⁶⁸ All data and software used in this work are publicly available. The DevCCF atlas is
⁷⁶⁹ available at <https://kimlab.io/brain-map/DevCCF/>. ANTsPy, ANTsR, ANTsPyNet, and
⁷⁷⁰ ANTsRNet are available through GitHub at the ANTsX Ecosystem (<https://github.com/>
⁷⁷¹ [ANTsX](#)). Training scripts for all deep learning functionality in ANTsXNet can also be found on
⁷⁷² GitHub (<https://github.com/ntustison/ANTsXNetTraining>). A GitHub repository specifically
⁷⁷³ pertaining to the AllenCCFv3 mapping is available at <https://github.com/dontminchenit/>
⁷⁷⁴ [CCFAAlignmentToolkit](#). For the other two contributions contained in this work, the longitudinal
⁷⁷⁵ DevCCF mapping and mouse cortical thickness pipeline, we refer the interested reader to
⁷⁷⁶ <https://github.com/ntustison/ANTsXMouseBrainMapping>.

777 Acknowledgments

778 Support for the research reported in this work includes funding from the National Institute
779 of Biomedical Imaging and Bioengineering (R01-EB031722) and National Institute of Mental
780 Health (RF1-MH124605 and U24-MH114827).

781 We also acknowledge the data contribution of Dr. Adam Raikes (GitHub @araikes) of the
782 Center for Innovation in Brain Science at the University of Arizona for refining the weights
783 of the mouse brain extraction network.

⁷⁸⁴ **Author contributions**

⁷⁸⁵ N.T., M.C., and J.G. wrote the main manuscript text and figures. M.C., M.K., R.D., S.S.,
⁷⁸⁶ Q.W., L.G., J.D., C.G., and J.G. developed the Allen registration pipelines. N.T. and F.K.
⁷⁸⁷ developed the time-varying velocity transformation model for the DevCCF. N.T. and M.T.
⁷⁸⁸ developed the brain parcellation and cortical thickness methodology. All authors reviewed
⁷⁸⁹ the manuscript.

790 **References**

- 791 1. Keller, P. J. & Ahrens, M. B. Visualizing whole-brain activity and development at the single-cell level using light-sheet microscopy. *Neuron* **85**, 462–83 (2015).
- 792 2. La Manno, G. *et al.* Molecular architecture of the developing mouse brain. *Nature* **596**, 92–96 (2021).
- 793 3. Wen, L. *et al.* Single-cell technologies: From research to application. *Innovation (Camb)* **3**, 100342 (2022).
- 794 4. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–14 (2014).
- 795 5. Gong, H. *et al.* Continuously tracing brain-wide long-distance axonal projections in mice at a one-micron voxel resolution. *Neuroimage* **74**, 87–98 (2013).
- 796 6. Li, A. *et al.* Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain. *Science* **330**, 1404–8 (2010).
- 797 7. Ueda, H. R. *et al.* Tissue clearing and its applications in neuroscience. *Nat Rev Neurosci* **21**, 61–79 (2020).
- 798 8. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
- 799 9. Burgess, D. J. Spatial transcriptomics coming of age. *Nat Rev Genet* **20**, 317 (2019).
- 800 10. Hardwick, S. A. *et al.* Single-nuclei isoform RNA sequencing unlocks barcoded exon connectivity in frozen brain tissue. *Nature biotechnology* **40**, 1082–1092 (2022).
- 801 11. Hawrylycz, M. *et al.* A guide to the BRAIN initiative cell census network data ecosystem. *PLoS biology* **21**, e3002133 (2023).
- 802 12. Wang, Q. *et al.* The allen mouse brain common coordinate framework: A 3D reference atlas. *Cell* **181**, 936–953.e20 (2020).
- 803 13. Perens, J. *et al.* An optimized mouse brain atlas for automated mapping and quantification of neuronal activity using iDISCO+ and light sheet fluorescence microscopy. *Neuroinformatics* **19**, 433–446 (2021).
- 804 14. Ma, Y. *et al.* A three-dimensional digital atlas database of the adult C57BL/6J mouse brain by magnetic resonance microscopy. *Neuroscience* **135**, 1203–1215 (2005).

- 805 15. Qu, L. *et al.* Cross-modal coherent registration of whole mouse brains. *Nature Methods* **19**, 111–118 (2022).
- 806 16. Kronman, F. N. *et al.* [Developmental mouse brain common coordinate framework](#). *Nat Commun* **15**, 9072 (2024).
- 807 17. Chuang, N. *et al.* An MRI-based atlas and database of the developing mouse brain. *Neuroimage* **54**, 80–89 (2011).
- 808 18. Dries, R. *et al.* Advances in spatial transcriptomic data analysis. *Genome research* **31**, 1706–1718 (2021).
- 809 19. Ricci, P. *et al.* Removing striping artifacts in light-sheet fluorescence microscopy: A review. *Progress in biophysics and molecular biology* **168**, 52–65 (2022).
- 810 20. Agarwal, N., Xu, X. & Gopi, M. Robust registration of mouse brain slices with severe histological artifacts. in *Proceedings of the tenth indian conference on computer vision, graphics and image processing* 1–8 (2016).
- 811 21. Agarwal, N., Xu, X. & Gopi, M. Automatic detection of histological artifacts in mouse brain slice images. in *Medical computer vision and bayesian and graphical models for biomedical imaging: MICCAI 2016 international workshops, MCV and BAMBI, athens, greece, october 21, 2016, revised selected papers* 8 105–115 (Springer, 2017).
- 812 22. Tward, D. *et al.* 3d mapping of serial histology sections with anomalies using a novel robust deformable registration algorithm. in *International workshop on multimodal brain image analysis* 162–173 (Springer, 2019).
- 813 23. Cahill, L. S. *et al.* Preparation of fixed mouse brains for MRI. *Neuroimage* **60**, 933–939 (2012).
- 814 24. Sunkin, S. M. *et al.* Allen brain atlas: An integrated spatio-temporal portal for exploring the central nervous system. *Nucleic acids research* **41**, D996–D1008 (2012).
- 815 25. Kim, Y. *et al.* Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell* **171**, 456–469 (2017).
- 816 26. Fürth, D. *et al.* [An interactive framework for whole-brain maps at cellular resolution](#). *Nat Neurosci* **21**, 139–149 (2018).

- 817 27. Li, Y. *et al.* mBrainAligner-web: A web server for cross-modal coherent registration
of whole mouse brains. *Bioinformatics* **38**, 4654–4655 (2022).
- 818 28. Puchades, M. A., Csucs, G., Ledergerber, D., Leergaard, T. B. & Bjaalie, J. G. Spatial
registration of serial microscopic brain images to three-dimensional reference atlases
with the QuickNII tool. *PloS one* **14**, e0216796 (2019).
- 819 29. Eastwood, B. S. *et al.* Whole mouse brain reconstruction and registration to a
reference atlas with standard histochemical processing of coronal sections. *Journal of
Comparative Neurology* **527**, 2170–2178 (2019).
- 820 30. Ni, H. *et al.* A robust image registration interface for large volume brain atlas. *Sci
Rep* **10**, 2139 (2020).
- 821 31. Pallast, N. *et al.* Processing pipeline for atlas-based imaging data analysis of structural
and functional mouse brain MRI (AIDAmri). *Front Neuroinform* **13**, 42 (2019).
- 822 32. Celestine, M., Nadkarni, N. A., Garin, C. M., Bougacha, S. & Dhenain, M. **Sammba-
MRI: A library for processing SmAll-MaMmal BrAin MRI data in python**. *Front
Neuroinform* **14**, 24 (2020).
- 823 33. Ioanas, H.-I., Marks, M., Zerbi, V., Yanik, M. F. & Rudin, M. An optimized registration
workflow and standard geometric space for small animal brain imaging. *Neuroimage*
241, 118386 (2021).
- 824 34. Perens, J. *et al.* Multimodal 3D mouse brain atlas framework with the skull-derived
coordinate system. *Neuroinformatics* **21**, 269–286 (2023).
- 825 35. Aggarwal, M., Zhang, J., Miller, M. I., Sidman, R. L. & Mori, S. Magnetic resonance
imaging and micro-computed tomography combined atlas of developing and adult
mouse brains for stereotaxic surgery. *Neuroscience* **162**, 1339–1350 (2009).
- 826 36. Goubran, M. *et al.* Multimodal image registration and connectivity analysis for
integration of connectomic data from microscopy to MRI. *Nat Commun* **10**, 5504
(2019).
- 827 37. Chandrashekhar, V. *et al.* CloudReg: Automatic terabyte-scale cross-modal brain
volume registration. *Nature methods* **18**, 845–846 (2021).

- 828 38. Jin, M. *et al.* SMART: An open-source extension of WholeBrain for intact mouse
brain registration and segmentation. *eNeuro* **9**, (2022).
- 829 39. Negwer, M. *et al.* FriendlyClearMap: An optimized toolkit for mouse brain mapping
and analysis. *Gigascience* **12**, (2022).
- 830 40. Lin, W. *et al.* Whole-brain mapping of histaminergic projections in mouse brain.
Proceedings of the National Academy of Sciences **120**, e2216231120 (2023).
- 831 41. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by
MERFISH. *Nature* **598**, 137–143 (2021).
- 832 42. Shi, H. *et al.* Spatial atlas of the mouse central nervous system at molecular resolution.
Nature **622**, 552–561 (2023).
- 833 43. Zhang, Y. *et al.* Reference-based cell type matching of in situ image-based spatial
transcriptomics data on primary visual cortex of mouse brain. *Scientific Reports* **13**,
9567 (2023).
- 834 44. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluim, J. P. W. Elastix: A
toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging* **29**,
196–205 (2010).
- 835 45. Fedorov, A. *et al.* 3D slicer as an image computing platform for the quantitative
imaging network. *Magnetic resonance imaging* **30**, 1323–1341 (2012).
- 836 46. Tustison, N. J. *et al.* The ANTsX ecosystem for quantitative biological and medical
imaging. *Sci Rep* **11**, 9068 (2021).
- 837 47. Pagani, M., Damiano, M., Galbusera, A., Tsafaris, S. A. & Gozzi, A. Semi-automated
registration-based anatomical labelling, voxel based morphometry and cortical thick-
ness mapping of the mouse brain. *Journal of neuroscience methods* **267**, 62–73
(2016).
- 838 48. Anderson, R. J. *et al.* Small animal multivariate brain analysis (SAMBA) - a high
throughput pipeline with a validation framework. *Neuroinformatics* **17**, 451–472
(2019).
- 839 49. Allan Johnson, G. *et al.* Whole mouse brain connectomics. *Journal of Comparative
Neurology* **527**, 2146–2157 (2019).

- 840 50. Yao, Z. *et al.* A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **624**, 317–332 (2023).
- 841 51. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* **12**, 26–41 (2008).
- 842 52. Tustison, N. J. *et al.* N4ITK: Improved N3 bias correction. *IEEE Trans Med Imaging* **29**, 1310–20 (2010).
- 843 53. Bajcsy, R. & Broit, C. Matching of deformed images. in *Sixth International Conference on Pattern Recognition (ICPR'82)* 351–353 (1982).
- 844 54. Bajcsy, R. & Kovacic, S. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing* **46**, 1–21 (1989).
- 845 55. Gee, J. C., Reivich, M. & Bajcsy, R. Elastically deforming 3D atlas to match anatomical brain images. *J Comput Assist Tomogr* **17**, 225–36 (1993).
- 846 56. Klein, A. *et al.* Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* **46**, 786–802 (2009).
- 847 57. Murphy, K. *et al.* Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge. *IEEE Trans Med Imaging* **30**, 1901–20 (2011).
- 848 58. Baheti, B. *et al.* The brain tumor sequence registration challenge: Establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients. (2021).
- 849 59. Avants, B. B. *et al.* The optimal template effect in hippocampus studies of diseased populations. *Neuroimage* **49**, 2457–66 (2010).
- 850 60. Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A. & Gee, J. C. An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics* **9**, 381–400 (2011).
- 851 61. Manjón, J. V., Coupé, P., Martí-Bonmatí, L., Collins, D. L. & Robles, M. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J Magn Reson Imaging* **31**, 192–203 (2010).

- 852 62. Wang, H. *et al.* Multi-atlas segmentation with joint label fusion. *IEEE Trans Pattern Anal Mach Intell* **35**, 611–23 (2013).
- 853 63. Tustison, N. J. *et al.* Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR. *Neuroinformatics* (2014) doi:10.1007/s12021-014-9245-2.
- 854 64. Tustison, N. J., Yang, Y. & Salerno, M. Advanced normalization tools for cardiac motion correction. in *Statistical atlases and computational models of the heart - imaging and modelling challenges* (eds. Camara, O. et al.) vol. 8896 3–12 (Springer International Publishing, 2015).
- 855 65. McCormick, M., Liu, X., Jomier, J., Marion, C. & Ibanez, L. ITK: Enabling reproducible research and open science. *Front Neuroinform* **8**, 13 (2014).
- 856 66. Beg, M. F., Miller, M. I., Trouvé, A. & Younes, L. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision* **61**, 139–157 (2005).
- 857 67. Tustison, N. J. & Avants, B. B. Explicit B-spline regularization in diffeomorphic image registration. *Front Neuroinform* **7**, 39 (2013).
- 858 68. Hsu, L.-M. *et al.* CAMRI mouse brain MRI data.
- 859 69. Reshetnikov, V. *et al.* High-resolution MRI data of brain C57BL/6 and BTBR mice in three different anatomical views.
- 860 70. Rahman, N., Xu, K., Budde, M. D., Brown, A. & Baron, C. A. A longitudinal microstructural MRI dataset in healthy C57Bl/6 mice at 9.4 tesla. *Sci Data* **10**, 94 (2023).
- 861 71. Liu, J. *et al.* Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing. *Life Sci Alliance* **6**, (2023).
- 862 72. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: A generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100–106 (2021).
- 863 73. Jia, H., Yap, P.-T., Wu, G., Wang, Q. & Shen, D. Intermediate templates guided groupwise registration of diffusion tensor images. *NeuroImage* **54**, 928–939 (2011).

- 864 74. Tang, S., Fan, Y., Wu, G., Kim, M. & Shen, D. RABBIT: Rapid alignment of brains
by building intermediate templates. *NeuroImage* **47**, 1277–1287 (2009).
- 865 75. Dewey, B. E., Carass, A., Blitz, A. M. & Prince, J. L. Efficient multi-atlas registration
using an intermediate template image. in *Proceedings of SPIE—the international society
for optical engineering* vol. 10137 (NIH Public Access, 2017).
- 866 76. Gong, H. *et al.* High-throughput dual-colour precision imaging for brain-wide connec-
tome with cytoarchitectonic landmarks at the cellular level. *Nat Commun* **7**, 12142
(2016).
- 867 77. Wang, J. *et al.* Divergent projection patterns revealed by reconstruction of individual
neurons in orbitofrontal cortex. *Neurosci Bull* **37**, 461–477 (2021).
- 868 78. Rotolo, T., Smallwood, P. M., Williams, J. & Nathans, J. Genetically-directed, cell
type-specific sparse labeling for the analysis of neuronal morphology. *PLoS One* **3**,
e4099 (2008).
- 869 79. Peng, H. *et al.* Morphological diversity of single neurons in molecularly defined cell
types. *Nature* **598**, 174–181 (2021).
- 870 80. Kronman, F. A. *et al.* Developmental mouse brain common coordinate framework.
bioRxiv (2023) doi:[10.1101/2023.09.14.557789](https://doi.org/10.1101/2023.09.14.557789).
- 871 81. Chon, U., Vanselow, D. J., Cheng, K. C. & Kim, Y. Enhanced and unified anatomical
labeling for a common mouse brain atlas. *Nat Commun* **10**, 5067 (2019).
- 872 82. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcript-
omics. *Nat Neurosci* **19**, 335–46 (2016).
- 873 83. Bergmann, E., Gofman, X., Kavushansky, A. & Kahn, I. Individual variability in
functional connectivity architecture of the mouse brain. *Commun Biol* **3**, 738 (2020).
- 874 84. Billot, B. *et al.* SynthSeg: Segmentation of brain MRI scans of any contrast and
resolution without retraining. *Med Image Anal* **86**, 102789 (2023).
- 875 85. Tustison, N. J. & Amini, A. A. Biventricular myocardial strains via nonrigid registration
of anatomical NURBS model [corrected]. *IEEE Trans Med Imaging* **25**, 94–112 (2006).
- 876 86. Avants, B. B. *et al.* The Insight ToolKit image registration framework. *Front Neu-
roinform* **8**, 44 (2014).

- 877 87. Avants, B. B. *et al.* A reproducible evaluation of ANTs similarity metric performance
in brain image registration. *Neuroimage* **54**, 2033–44 (2011).
- 878 88. Nyúl, L. G., Udupa, J. K. & Zhang, X. New variants of a method of MRI scale
standardization. *IEEE Trans Med Imaging* **19**, 143–50 (2000).
- 879 89. Falk, T. *et al.* U-net: Deep learning for cell counting, detection, and morphometry.
Nat Methods **16**, 67–70 (2019).