

1

## 2 Modular strategies for spatial mapping of 3 diverse cell type data of the mouse brain

4 Nicholas J. Tustison<sup>1</sup>, Min Chen<sup>2</sup>, Fae N. Kronman<sup>3</sup>, Jeffrey T. Duda<sup>2</sup>, Clare Gamlin<sup>4</sup>, Mia  
5 G. Tustison, Michael Kunst<sup>4</sup>, Rachel Dalley<sup>4</sup>, Staci Sorenson<sup>4</sup>, Quanxin Wang<sup>4</sup>, Lydia Ng<sup>4</sup>,  
6 Yongsoo Kim<sup>3</sup>, and James C. Gee<sup>2</sup>

7 <sup>1</sup>Department of Radiology and Medical Imaging, University of Virginia, Charlottesville, VA

8 <sup>2</sup>Department of Radiology, University of Pennsylvania, Philadelphia, PA

9 <sup>3</sup>Department of Neural and Behavioral Sciences, Penn State University, Hershey, PA

10 <sup>4</sup>Allen 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](mailto:ntustison@virginia.edu)

18

19 James C. Gee, PhD

20 Department of Radiology

21 University of Pennsylvania

22 [gee@upenn.edu](mailto:gee@upenn.edu)

<sup>23</sup> **Abstract**

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

## <sup>48</sup> 1 Introduction

<sup>49</sup> Over the past decade there have been significant advancements in mesoscopic single-cell analysis of the mouse brain. It is now possible to track single neurons in mouse brains<sup>1</sup>, observe whole brain developmental changes on a cellular level<sup>2</sup>, associate brain regions and tissues with their genetic composition<sup>3</sup>, and locally characterize neural connectivity<sup>4</sup>. Much of these scientific achievements have been made possible due to breakthroughs in high resolution cell profiling and imaging techniques that permit submicron, multi-modal, 3-D characterizations of whole mouse brains. Among these include advanced techniques such as micro-optical sectioning tomography<sup>6</sup>, tissue clearing<sup>1,7</sup>, spatial transcriptomics<sup>9</sup>, and single-cell genomic profiling<sup>10</sup>, which have greatly expanded the resolution and specificity of single-cell measurements in the brain.

<sup>59</sup> Recent efforts by the National Institutes of Health’s Brain Research Through Advancing Innovative Neurotechnologies (BRAIN) Initiative has pushed for large-scale, international collaborative efforts to utilize these advanced single-cell techniques to create a comprehensive reference database for high-resolution transcriptomic, epigenomic, structural and imaging data of the mouse brain. This consortium of laboratories and data centers, known as the BRAIN Initiative Cell Census Network (BICCN), has archived datasets encompassing over 40 different multi-modal profiling techniques from more than 30 research groups, each providing unique characterizations of distinct cell types in the brain<sup>11</sup>. Several of these modalities have been further developed into reference atlases to facilitate spatial alignment of individual brains and different data types into a common coordinate framework (CCF), thus allowing diverse single-cell information to be analyzed in an integrated manner. The most notable of these atlases is the Allen Mouse Brain Common Coordinate Framework (AllenCCFv3)<sup>12</sup>, which serves as a primary target coordinate space for much of the work associated with the BICCN. Other atlases include modality-specific atlases<sup>13–15</sup>, and spatiotemporal atlases<sup>16,17</sup> for the developing mouse brain.

<sup>74</sup> **1.1 Mouse brain mapping**

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

<sup>89</sup> Existing solutions to address mapping cell type data into the AllenCCFv3 falls broadly into  
<sup>90</sup> three main categories. The first consists of integrated processing platforms that directly  
<sup>91</sup> provide mapped data to the users. These include the Allen Brain Cell Atlas<sup>24</sup> for the Allen  
<sup>92</sup> Reference Atlas (ARA) and associated data, the Brain Architecture Portal<sup>25</sup> for combined ex  
<sup>93</sup> vivo radiology and histology data, OpenBrainMap<sup>26</sup> for connectivity data, and the Image and  
<sup>94</sup> Multi-Morphology Pipeline<sup>27</sup> for high resolution morphology data. These platforms provide  
<sup>95</sup> users online access to pre-processed, multi-modal cell type data that are already mapped to  
<sup>96</sup> the AllenCCFv3. The platforms are designed such that the data is interactively manipulated  
<sup>97</sup> by users through integrated visualization software that allow users to spatially manipulate  
<sup>98</sup> and explore each dataset within the mapped space. While highly convenient for investigators  
<sup>99</sup> who are interested in studying the specific modalities provided by these platforms, these  
<sup>100</sup> systems can be limited in flexibility, general applicability, and public availability. As a  
<sup>101</sup> result, investigators often find it difficult to apply the same mapping solutions to their own  
<sup>102</sup> data.

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

115 The last category consist of modular mapping approaches constructed using general im-  
116 age analysis toolkits, which are software packages that include modular image processing,  
117 segmentation and registration tools that have been previously developed, and validated for  
118 multiple application areas. Examples of such toolkits include elastix<sup>44</sup>, Slicer3D<sup>45</sup>, ANTsX<sup>46</sup>,  
119 and several others which have all been applied towards mouse brain spatial mapping. The  
120 main challenge, in these mouse-specific study scenarios, is that tailored pipelines often need  
121 be constructed from available software components. Investigators must therefore be familiar  
122 with the these tools for formulating new or adapting existing pipelines. However, in com-  
123 parison to previously described specialized mapping approaches, these approaches are often  
124 easier to create and prone to robustness, being typically constructed from pipeline compo-  
125 nents which have been previously vetted in other contexts. In this work, we highlight such  
126 mapping strategies designed using the ANTsX framework to map distinct mouse cell type  
127 data with different characteristics into existing CCFs.

## 128 1.2 Advanced Normalization Tools (ANTsX)

129 The Advanced Normalization Tools Ecosystem (ANTsX) has been used in a number of  
130 applications for mapping mouse brain data as part of core processing steps in various

<sup>131</sup> workflows<sup>30,47–50</sup>, particularly its pairwise, intensity-based image registration capabilities<sup>51</sup>  
<sup>132</sup> and bias field correction<sup>52</sup>. Historically, ANTsX development is originally based on fun-  
<sup>133</sup> damental approaches to image mapping<sup>53–55</sup>, particularly in the human brain, which has  
<sup>134</sup> resulted in core contributions to the field such as the widely-used Symmetric Normalization  
<sup>135</sup> (SyN) algorithm<sup>51</sup>. Since its development, various independent platforms have been used  
<sup>136</sup> to evaluate ANTsX image registration capabilities in the context of different application  
<sup>137</sup> foci which include multi-site brain MRI data<sup>56</sup>, pulmonary CT data<sup>57</sup>, and most recently,  
<sup>138</sup> multi-modal brain registration in the presence of tumors<sup>58</sup>.

<sup>139</sup> Apart from its registration capabilities, ANTsX comprises additional functionality such  
<sup>140</sup> as template generation<sup>59</sup>, intensity-based segmentation<sup>60</sup>, preprocessing<sup>52,61</sup>, deep learning  
<sup>141</sup> networks<sup>46</sup>, and other utilities relevant to brain mapping (see Table 1). The use of the toolkit  
<sup>142</sup> has demonstrated high performance in multiple application areas (e.g., consensus labeling<sup>62</sup>,  
<sup>143</sup> brain tumor segmentation<sup>63</sup>, and cardiac motion estimation<sup>64</sup>). Importantly, ANTsX is built  
<sup>144</sup> on the Insight Toolkit (ITK)<sup>65</sup> deriving benefit from the open-source community of scientists  
<sup>145</sup> and programmers as well as providing an important resource for algorithmic development,  
<sup>146</sup> evaluation, and improvement.

<sup>147</sup> With respect to mouse cell type data, ANTsX provides a comprehensive toolset which serves  
<sup>148</sup> as a basis for developing modular frameworks for mapping diverse image data into com-  
<sup>149</sup> mon coordinate frameworks (CCFs). Herein, we highlight its application for mapping data  
<sup>150</sup> from separate BICCN projects focused on distinct data types: morphology data using flu-  
<sup>151</sup> orescence micro-optical sectioning tomography (fMOST), spatial transcriptomics from mul-  
<sup>152</sup> tiplexed error-robust fluorescence in situ hybridization (MERFISH) data, and time-series  
<sup>153</sup> developmental data using light sheet fluorescence microscopy (LSFM) and magnetic reso-  
<sup>154</sup> nance imaging (MRI). We describe both shared and targeted strategies developed to address  
<sup>155</sup> the specific challenges of these modalities.

### <sup>156</sup> 1.3 Novel ANTsX-based open-source contributions

<sup>157</sup> We introduce two novel inclusions to the ANTsX toolset that were developed as part of  
<sup>158</sup> the MRI mapping and analysis pipeline for the Developmental Common Coordinate Frame-

<sup>159</sup> work (DevCCF). Consistent with previous ANTsX development, newly introduced capa-  
<sup>160</sup> bilities introduced below are available through ANTsX (specifically, via R and Python  
<sup>161</sup> ANTsX packages), and illustrated through self-contained examples in the ANTsX tuto-  
<sup>162</sup> rial (<https://tinyurl.com/antsxtutorial>) with a dedicated GitHub repository specific to this  
<sup>163</sup> work (<https://github.com/ntustison/ANTsXMouseBrainMapping>). To complement stan-  
<sup>164</sup> dard preprocessing steps (e.g., bias correction, brain masking), additional mouse brain spe-  
<sup>165</sup> cific tools have also been introduced to the ANTsX ecosystem, such as section reconstruction  
<sup>166</sup> and landmark-based alignment with corresponding processing scripts (<https://github.com/>  
<sup>167</sup> [dontminchenit/CCFAAlignmentToolkit](https://github.com/dontminchenit/CCFAAlignmentToolkit)).

<sup>168</sup> **1.3.1 Continuously mapping the DevCCF developmental trajectory with a ve-**  
<sup>169</sup> **locity flow model**

<sup>170</sup> Recently, the Developmental Common Coordinate Framework (DevCCF) was introduced to  
<sup>171</sup> the mouse brain research community as a public resource<sup>16</sup> comprising symmetric atlases of  
<sup>172</sup> multi-modal image data and anatomical segmentations defined by developmental ontology.  
<sup>173</sup> These templates sample the mouse embryonic days E11.5, E13.5, E15.5, E18.5 and postnatal  
<sup>174</sup> days P4, P14, and P56. Modalities include LSFM and at least four MRI contrasts per  
<sup>175</sup> developmental stage. Anatomical parcellations are also available for each time point and  
<sup>176</sup> were generated from ANTsX-based mappings of gene expression and other cell type data.  
<sup>177</sup> Additionally, the P56 template was integrated with the AllenCCFv3 to further enhance the  
<sup>178</sup> practical utility of the DevCCF. These processes, specifically template generation and multi-  
<sup>179</sup> modal image mapping, were performed using ANTsX functionality in the presence of image  
<sup>180</sup> mapping difficulties such as missing data and tissue distortion.

<sup>181</sup> Given the temporal gaps in the discrete set of developmental atlases, we also provide an  
<sup>182</sup> open-source framework for inferring correspondence within the temporally continuous do-  
<sup>183</sup> main sampled by the existing set of embryonic and postnatal atlases of the DevCCF. This  
<sup>184</sup> recently developed functionality permits the generation of a diffeomorphic velocity flow trans-  
<sup>185</sup> formation model<sup>66</sup>, influenced by previous work<sup>67</sup>. The resulting time-parameterized veloc-  
<sup>186</sup> ity field spans the stages of the DevCCF where mappings between any two continuous time

<sup>187</sup> points within the span bounded by the E11.5 and P56 atlases are determined by numerical  
<sup>188</sup> integration of the optimized velocity field.

<sup>189</sup> **1.3.2 Automated structural parcellations of the mouse brain**

<sup>190</sup> In contrast to the pipeline development in human data<sup>46</sup>, limited tools exist yet to cre-  
<sup>191</sup> ate adequate training data for automated parcellations of the mouse brain. In addition,  
<sup>192</sup> mouse brain data acquisition often has unique issues, such as lower data quality or sampling  
<sup>193</sup> anisotropy which limits its applicability to high resolution resources (e.g., AllenCCFv3, De-  
<sup>194</sup> vCCF), specifically with respect to the corresponding granular brain parcellations derived  
<sup>195</sup> from numerous hours of expert annotation leveraging multi-modal imaging resources.

<sup>196</sup> Herein, we introduce a mouse brain parcellation pipeline for multi-modal MRI comprising  
<sup>197</sup> two novel deep learning components: two-shot learning brain extraction from data augmen-  
<sup>198</sup> tation of two ANTsX templates generated from two open datasets<sup>68,69</sup> and single-shot brain  
<sup>199</sup> parcellation derived from the AllenCCFv3 labelings mapped to the corresponding DevCCF  
<sup>200</sup> P56 template.

<sup>201</sup> Although we anticipate that this pipeline will be beneficial to the research community, this  
<sup>202</sup> work demonstrates more generally how one can leverage ANTsX tools and other public  
<sup>203</sup> resources for developing quantitative mouse brain morphological tools. Evaluation is per-  
<sup>204</sup> formed on an independent open dataset<sup>70</sup> comprising longitudinal acquisitions of multiple  
<sup>205</sup> specimens.

206 **2 Results**

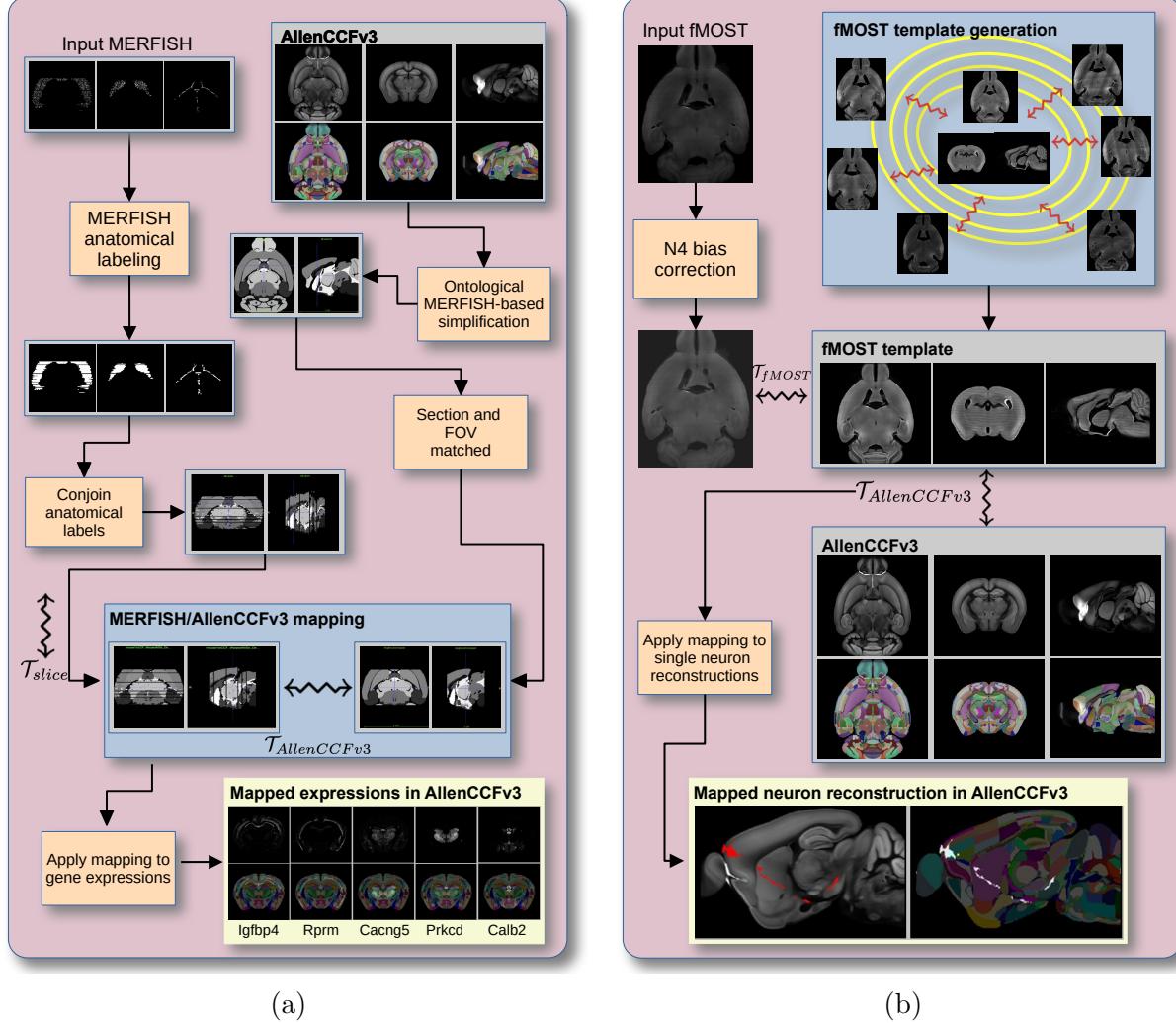


Figure 1: Diagram of the two ANTsX-based pipelines for mapping (a) MERFISH and (b)fMOST data into the space of AllenCCFv3. Each generates the requisite transforms,  $\mathcal{T}$ , to map individual images to the CCF.

207 **2.1 AllenCCFv3 brain image mapping**

208 **2.1.1 Mapping multiplexed error-robust fluorescence *in situ* hybridization  
209 (MERFISH) data**

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

212 data onto the AllenCCFv3 (see Figure 1(a)). This pipeline, used recently in creating a  
213 high-resolution transcriptomic atlas of the mouse brain<sup>50</sup>, performs mappings by first gen-  
214 erating anatomical labels from tissue related gene expressions in the MERFISH data, and  
215 then spatially matching these labels to corresponding anatomical tissue parcellations in the  
216 AllenCCFv3. The pipeline consists of MERFISH data specific preprocessing which includes  
217 section reconstruction, mapping corresponding anatomical labels between AllenCCFv3 and  
218 the spatial transcriptomic maps of the MERFISH data, and matching MERFISH sections  
219 to the atlas space. Following preprocessing, two main alignment steps were performed: 1)  
220 3-D global affine mapping and section matching of the AllenCCFv3 into the MERFISH data  
221 and 2) 2-D global and deformable mapping between each MERFISH section and matched  
222 AllenCCFv3 section. Mappings learned via each step in the pipeline are preserved and con-  
223 catenated to provide point-to-point correspondence between the original MERFISH data  
224 and AllenCCFv3, thus allowing individual gene expressions to be transferred into the Al-  
225 lenCCFv3.

226 **Data.** MERFISH mouse brain data was acquired using a previously detailed procedure<sup>50</sup>.  
227 Briefly, a brain of C57BL/6 mouse was dissected according to standard procedures and  
228 placed into an optimal cutting temperature (OCT) compound (Sakura FineTek 4583) in  
229 which it was stored at -80°C. The fresh frozen brain was sectioned at 10 $\mu m$  on Leica 3050  
230 S cryostats at intervals of 200 $\mu m$  to evenly cover the brain. A set of 500 genes were imaged  
231 that had been carefully chosen to distinguish the ~5200 clusters of our existing RNAseq  
232 taxonomy. For staining the tissue with MERFISH probes, a modified version of instructions  
233 provided by the manufacturer was used<sup>50</sup>. Raw MERSCOPE data were decoded using  
234 Vizgen software (v231). Cells were segmented based on DAPI and PolyT staining using  
235 Cellpose<sup>71,72</sup>. Segmentation was performed on a median z-plane (fourth out of seven) and  
236 cell borders were propagated to z-planes above and below. To assign cluster identity to each  
237 cell in the MERFISH dataset, we mapped the MERFISH cells to the scRNA-seq reference  
238 taxonomy.

239 **Evaluation.** Alignment of the MERFISH data into the AllenCCFv3 was qualitatively as-  
240 sessed by an expert anatomist at each iteration of the registration using known correspon-

<sup>241</sup> dence of gene markers and their associations with the AllenCCFv3. As previously reported<sup>50</sup>,  
<sup>242</sup> further assessment of the alignment showed that, of the 554 terminal regions (gray matter  
<sup>243</sup> only) in the AllenCCFv3, only seven small subregions were missed from the MERFISH  
<sup>244</sup> dataset: frontal pole, layer 1 (FRP1), FRP2/3, FRP5; accessory olfactory bulb, glomerular  
<sup>245</sup> layer (AOBgl); accessory olfactory bulb, granular layer (AOBgr); accessory olfactory bulb,  
<sup>246</sup> mitral layer (AOBmi); and accessory supraoptic group (ASO).

### <sup>247</sup> 2.1.2 Mapping fluorescence micro-optical sectioning tomography (fMOST) data

<sup>248</sup> **Overview.** We developed a pipeline for mapping fluorescence micro-optical sectioning to-  
<sup>249</sup> mography (fMOST) mouse brain images into the AllenCCFv3 (see Figure 1(b)). The pipeline  
<sup>250</sup> is adapted from previously developed frameworks for human brain mapping<sup>59</sup>, and uses a  
<sup>251</sup> modality specific (fMOST) average atlas to assist in the image registration and mapping.  
<sup>252</sup> This approach has been well validated in human studies<sup>73–75</sup>, and successfully used in other  
<sup>253</sup> mouse data<sup>12,15,34</sup>. Briefly, we construct an intensity- and shape-based average fMOST atlas  
<sup>254</sup> using 30 fMOST images to serve as an intermediate registration target for mapping fMOST  
<sup>255</sup> images from individual specimens into the AllenCCFv3. Preprocessing steps include down-  
<sup>256</sup> sampling to match the  $25\mu m$  isotropic AllenCCFv3, acquisition-based stripe artifact removal,  
<sup>257</sup> and inhomogeneity correction<sup>52</sup>. Preprocessing also includes a single annotation-driven reg-  
<sup>258</sup> istration to establish a canonical mapping between the fMOST atlas and the AllenCCFv3.  
<sup>259</sup> This step allows us to align expert determined landmarks to accurately map structures  
<sup>260</sup> with large morphological differences between the modalities, which are difficult to address  
<sup>261</sup> using standard approaches. Once this canonical mapping is established, standard intensity-  
<sup>262</sup> based registration is used to align each new fMOST image to the fMOST specific atlas.  
<sup>263</sup> This mapping is concatenated with the canonical fMOST atlas-to-AllenCCFv3 mapping to  
<sup>264</sup> further map each individual brain into the latter without the need to generate additional  
<sup>265</sup> landmarks. Transformations learned through this mapping can be applied to single neuron  
<sup>266</sup> reconstructions from the fMOST images to evaluate neuronal distributions across different  
<sup>267</sup> specimens into the AllenCCFv3 for the purpose of cell census analyses.

<sup>268</sup> **Data.** The high-throughput and high-resolution fluorescence micro-optical sectioning to-

269 mography (fMOST)<sup>76,77</sup> platform was used to image 55 mouse brains containing gene-defined  
270 neuron populations, with sparse transgenic expression<sup>78,79</sup>. In short, the fMOST imaging  
271 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  
272 imaging system where the green channel displays the green fluorescent protein (GFP) labeled  
273 neuron morphology and the red channel is used to visualize the counterstained propidium  
274 iodide cytoarchitecture. The spatial normalizations described in this work were performed  
275 using the red channel, which offered higher tissue contrast for alignment, although other  
276 approaches are possible including multi-channel registration.

277 **Evaluation.** Evaluation of the canonical fMOST atlas to Allen CCFv3 mapping was per-  
278 formed via quantitative comparison at each step of the registration and qualitative assess-  
279 ment of structural correspondence after alignment by an expert anatomist. Dice values were  
280 generated for the following structures: whole brain, 0.99; fimbria, 0.91; habenular com-  
281 missure, 0.63; posterior choroid plexus, 0.93; anterior choroid plexus, 0.96; optic chiasm,  
282 0.77; caudate putamen, 0.97. Similar qualitative assessment was performed for each fMOST  
283 specimen including the corresponding neuron reconstruction data.

## 284 2.2 Continuously mapping the DevCCF developmental trajectory 285 with a velocity flow model

286 The DevCCF is an openly accessible resource for the mouse brain research community<sup>16</sup>. It  
287 consists of multi-modal MRI and LSFM symmetric ANTsX-generated templates<sup>59</sup> sampling  
288 the mouse brain developmental trajectory, specifically the embryonic (E) and postnatal (P)  
289 days E11.5, E13.5, E15.5, E18.5 P4, P14, and P56. Each template space includes structural  
290 labels defined by a developmental ontology. Its utility is also enhanced by a coordinated  
291 construction with AllenCCFv3. Although this work represents a significant contribution,  
292 the gaps between time points potentially limit its applicability which could be addressed  
293 through the development of the ability to map not only between time points but also within  
294 and across time points.

295 To continuously generate transformations between the different stages of the DevCCF atlases,

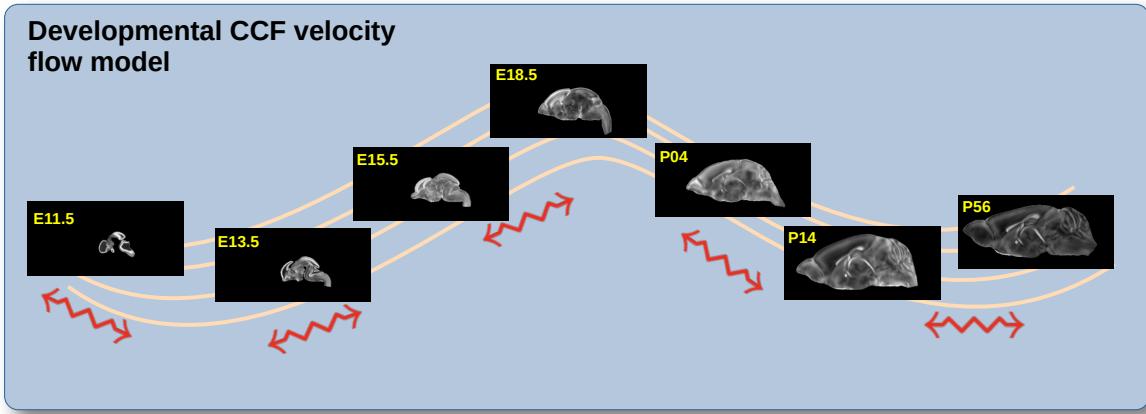


Figure 2: The spatial transformation between any two time points within the continuous DevCCF longitudinal developmental trajectory is available through the use of ANTsX functionality for generating a velocity flow model.

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

### 2.2.1 Data

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

Prior to velocity field optimization, all data were rigidly transformed to DevCCF P56 using the centroids of the common label sets. In order to determine the landmark correspondence

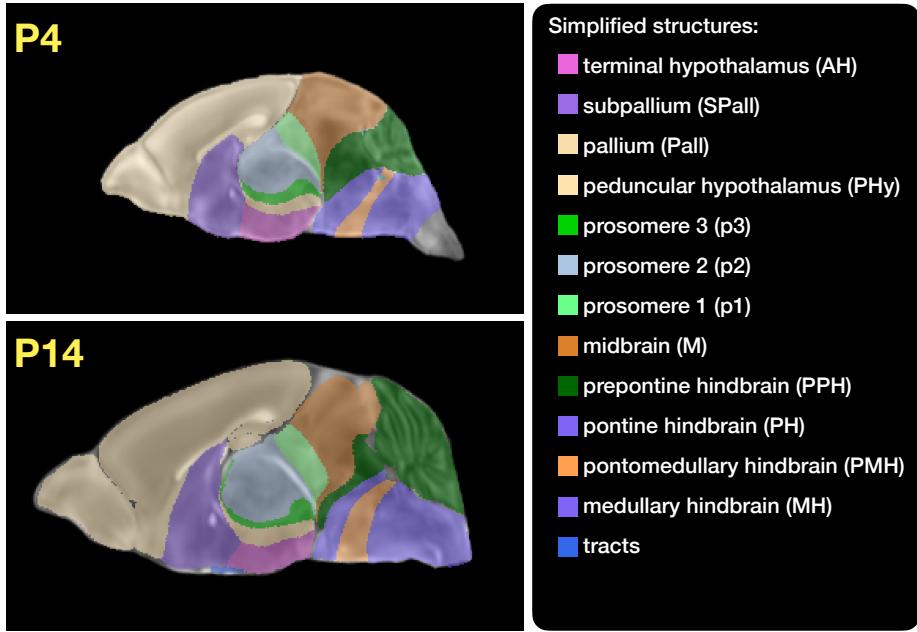


Figure 3: Annotated regions representing common labels across developmental stages which are illustrated for both P4 and P14.

312 across DevCCF stages, the multi-metric capabilities of `ants.registration(...)` were used.  
 313 Instead of performing intensity-based pairwise registration directly on these multi-label im-  
 314 ages, each label was used to construct a separate fixed and moving image pair resulting in a  
 315 multi-metric registration optimization scenario involving 24 binary image pairs (each label  
 316 weighted equally) for optimizing diffeomorphic correspondence between neighboring time  
 317 point atlases using the mean squares metric and the symmetric normalization transform<sup>51</sup>.  
 318 To generate the set of common point sets across all seven developmental atlases, the label  
 319 boundaries and whole regions were sampled in the P56 atlas and then propagated to each  
 320 atlas using the transformations derived from the pairwise registrations. We selected a sam-  
 321 pling rate of 10% for the contour points and 1% for the regional points for a total number  
 322 of points being per atlas being 173303 ( $N_{contour} = 98151$  and  $N_{region} = 75152$ ). Regional  
 323 boundary points were weighted twice as those of non-boundary points during optimization.

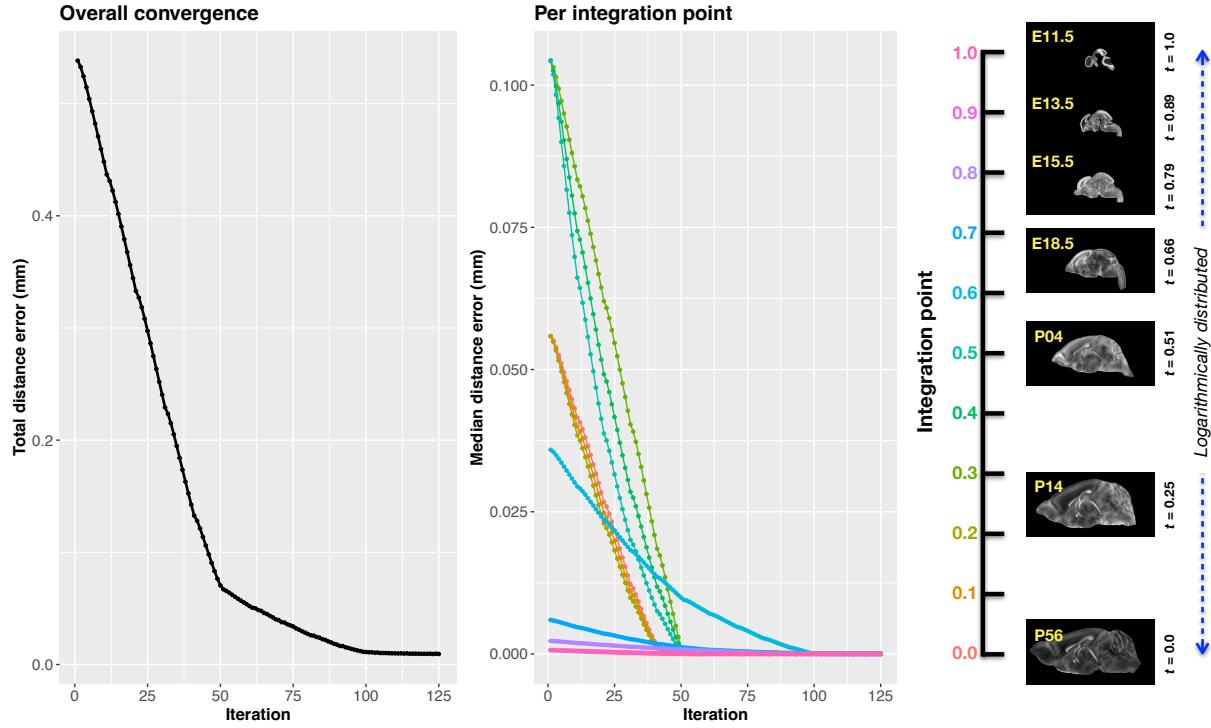


Figure 4: Convergence of the optimization of the velocity field for describing the transformation through the developmental stages from E11.5 through P56. Integration points in diagram on the right are color-coordinated with the center plot and placed in relation to the logarithmically situated temporal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

### <sup>324</sup> 2.2.2 Velocity field optimization

<sup>325</sup> The velocity field was optimized using the input composed of the seven corresponding point  
<sup>326</sup> sets and their associated weight values, the selected number of integration points for the  
<sup>327</sup> velocity field ( $N = 11$ ), and the parameters defining the geometry of the spatial dimensions  
<sup>328</sup> of the velocity field. Thus, the optimized velocity field described here is of size [256, 182, 360]  
<sup>329</sup> ( $50\mu\text{m}$  isotropic)  $\times 11$  integration points for a total compressed size of a little over 2 GB.

<sup>330</sup> This choice represented weighing the trade-off between tractability, portability, and accuracy.  
<sup>331</sup> However, all data and code to reproduce the results described are available in the dedicated  
<sup>332</sup> GitHub repository.

<sup>333</sup> The normalized time point scalar value for each atlas/point-set in the temporal domains  $[0, 1]$   
<sup>334</sup> was also defined. Given the increasingly larger gaps in the postnatal time point sampling,  
<sup>335</sup> we made two adjustments. Based on known mouse brain development, we used 28 days for

336 the P56 data. We then computed the log transform of the adjusted set of time points prior  
337 to normalization between 0 and 1 (see the right side of Figure 4). This log transform, as  
338 part of the temporal normalization, significantly improves the temporal spacing of data.

339 The maximum number of iterations was set to 200 with each iteration taking approximately  
340 six minutes on a 2020 iMac (processor, 3.6 GHz 10-Core Intel Core i9; memory, 64 GB 2667  
341 MHz DDR4) At each iteration we looped over the 11 integration points. At each integration  
342 point, the velocity field estimate was updated by warping the two immediately adjacent  
343 point sets to the integration time point and determining the regularized displacement field  
344 between the two warped point sets. As with any gradient-based descent algorithm, this field  
345 was multiplied by a small step size ( $\delta = 0.2$ ) before adding to the current velocity field.  
346 Convergence is determined by the average displacement error over each of the integration  
347 points. As can be seen in the left panel of Figure 4, convergence occurred around 125  
348 iterations when the average displacement error over all integration points is minimized. The  
349 median displacement error at each of the integration points also trends towards zero but at  
350 different rates.

### 351 2.2.3 The velocity flow transformation model

352 Once optimized, the resulting velocity field can be used to generate the deformable transform  
353 between any two continuous points within the time interval bounded by E11.5 and P56. As  
354 a demonstration, in Figure 5, we transform each atlas to the space of every other atlas  
355 using the DevCCF transform model. Additionally, one can use this transformation model  
356 to construct virtual templates in the temporal gaps of the DevCCF. Given an arbitrarily  
357 chosen time point within the normalized time point interval, the existing adjacent DevCCF  
358 atlases on either chronological side can be warped to the desired time point. A subsequent  
359 call to one of the ANTsX template building functions then permits the construction of the  
360 template at that time point. In Figure 6, we illustrate the use of the DevCCF velocity flow  
361 model for generating two such virtual templates for two arbitrary time points. Note that  
362 both of these usage examples can be found in the GitHub repository previously given.

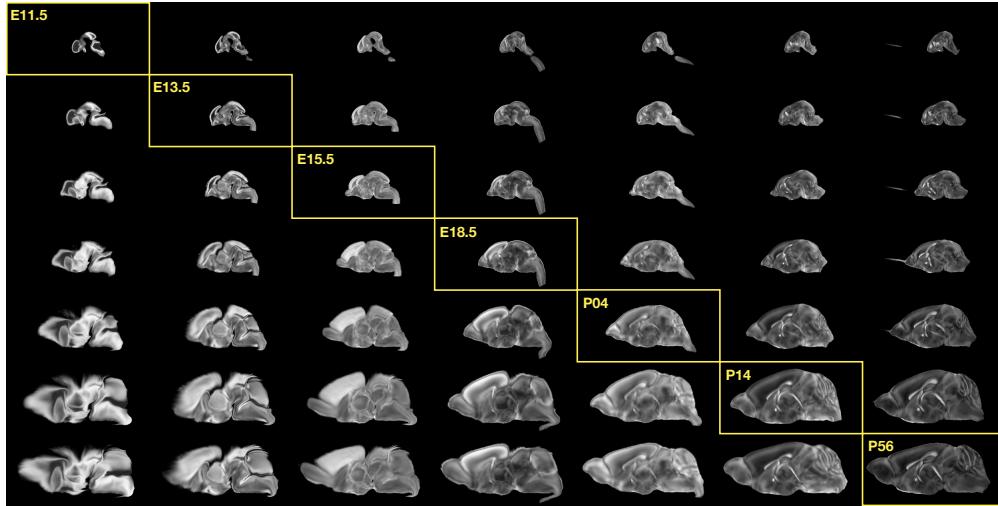


Figure 5: Mid-sagittal visualization of the effects of the transformation model in warping every developmental stage to the time point of every other developmental stage. The original images are located along the diagonal. Columns correspond to the warped original image whereas the rows represent the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

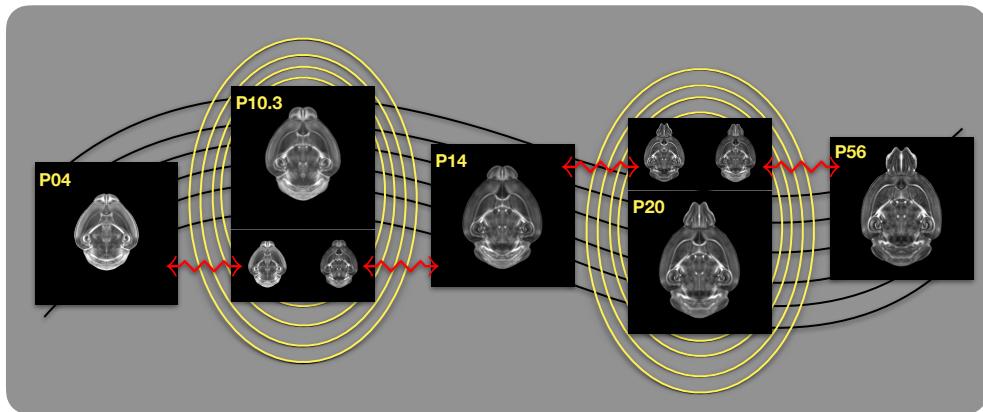


Figure 6: Illustration of the use of the velocity flow model for creating virtual templates at continuous time points not represented in one of the existing DevCCF time points. For example, FA templates at time point P10.3 and P20 can be generated by warping the existing temporally adjacent developmental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

### 363 2.3 Automated structural parcellations of the mouse brain

364 Brain parcellation strategies for the mouse brain are pivotal for understanding the complex  
365 organization and function of murine nervous system<sup>80</sup>. By dividing the brain into distinct  
366 regions based on anatomical, physiological, or functional characteristics, researchers can  
367 investigate specific areas in isolation and identify their roles in various behaviors and pro-

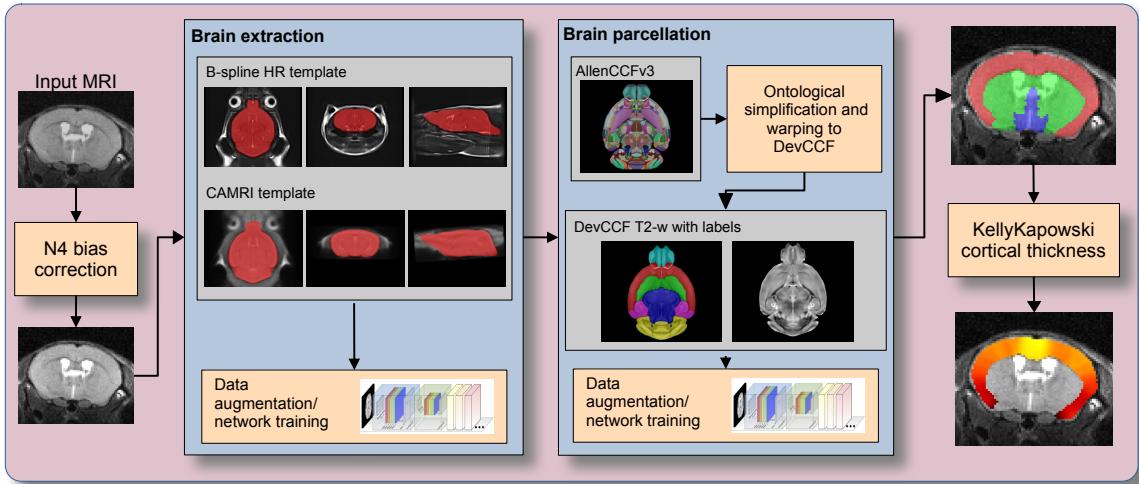


Figure 7: The mouse brain cortical parcellation pipeline integrating two deep learning components for brain extraction and brain parcellation prior to estimating cortical labels. Both deep learning networks rely heavily on aggressive data augmentation on templates built from open data and provide an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

cesses. For example, such parcellation schemes can help elucidate the spatial distribution of gene expression patterns<sup>81</sup> as well as identify functional regions involved in specific cognitive tasks<sup>82</sup>.

Although deep learning techniques have been used to develop useful parcellation tools for human brain research (e.g., SynthSeg<sup>83</sup>, ANTsXNet<sup>46</sup>), analogous development for the mouse brain is limited. In addition, mouse data is often characterized by unique imaging issues 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) dataset<sup>68</sup> from the University of North Carolina at Chapel Hill consists of 16 T2-w MRI volumes of

<sup>382</sup> voxel resolution  $0.16 \times 0.16 \times 0.16\text{mm}^3$ . The second high-resolution dataset<sup>69</sup> comprises  
<sup>383</sup> 88 specimens each with three spatially aligned canonical views with in-plane resolution of  
<sup>384</sup>  $0.08 \times 0.08\text{mm}^2$  with a slice thickness of  $0.5\text{mm}$ . These three orthogonal views were used  
<sup>385</sup> to reconstruct a single high-resolution volume per subject using a B-spline fitting algorithm  
<sup>386</sup> available in ANTsX<sup>84</sup>.

<sup>387</sup> From these two datasets, two ANTsX templates<sup>59</sup> were generated. Bias field simulation,  
<sup>388</sup> intensity histogram warping, noise simulation, random translation and warping, and random  
<sup>389</sup> anisotropic resampling in the three canonical directions were used for data augmentation  
<sup>390</sup> in training an initial T2-w brain extraction network. This network was posted and the  
<sup>391</sup> corresponding functionality was immediately made available within ANTsXNet, similar to  
<sup>392</sup> our previous contributions to the community.

<sup>393</sup> 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  
<sup>394</sup> user-made third template and extracted brain mask generated from T2-w ex-vivo data with  
<sup>395</sup> isotropic spacing of  $0.08\text{ mm}$  in each voxel dimension. This third template, in conjunction  
<sup>396</sup> with the other two, were used with the same aggressive data augmentation to refine the  
<sup>397</sup> network weights which were subsequently posted and made available through ANTsPyNet  
<sup>398</sup> using the function `antspynet.mouse_brain_extraction(...)`.

#### <sup>400</sup> 2.3.2 Single-shot mouse brain parcellation network

<sup>401</sup> AllenCCFv3 and its hierarchical ontological labeling, along with the DevCCF, provides the  
<sup>402</sup> necessary data for developing a tailored structural parcellation network for multi-modal  
<sup>403</sup> imaging. The `allensdk` Python library permits the creation of any gross parcellation based  
<sup>404</sup> on the AllenCCFv3 ontology. Specifically, using `allensdk` we coalesced the labels to the  
<sup>405</sup> following six major structures: cerebral cortex, cerebral nuclei, brain stem, cerebellum, main  
<sup>406</sup> olfactory bulb, and hippocampal formation. This labeling was mapped to the P56 component  
<sup>407</sup> of the DevCCF for use with the T2-w template component.

<sup>408</sup> The T2-w P56 DevCCF and labelings, in conjunction with the data augmentation  
<sup>409</sup> described previously for brain extraction, were used to train the proposed brain

410 parcellation network. This is available in ANTsXNet (e.g. in ANTsPyNet using  
 411 `antspynet.mouse_brain_parcellation(...)`). Note that other brain parcellation  
 412 networks have also been trained using alternative regions and parcellation schemes and are  
 413 available in the same ANTsXNet functionality. One usage note is that the data augmenta-  
 414 tion used to train the network permits a learned interpolation in 0.08 mm isotropic space.  
 415 Since the training data is isotropic and data augmentation includes downsampling in the  
 416 canonical directions, each of the two networks learns mouse brain-specific interpolation such  
 417 that one can perform prediction on thick-sliced images, as, for example, in these evaluation  
 418 data, and return isotropic probability and thickness maps (a choice available to the user).  
 419 This permits robust cortical thickness estimation even in the case of anisotropic data (see  
 420 `antspynet.mouse_cortical_thickness(...)`).

#### 421 2.3.3 Evaluation

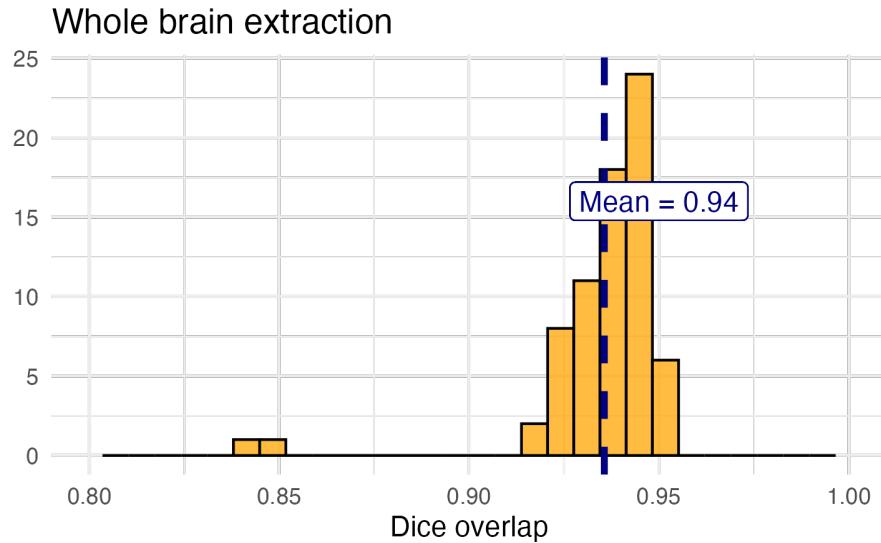
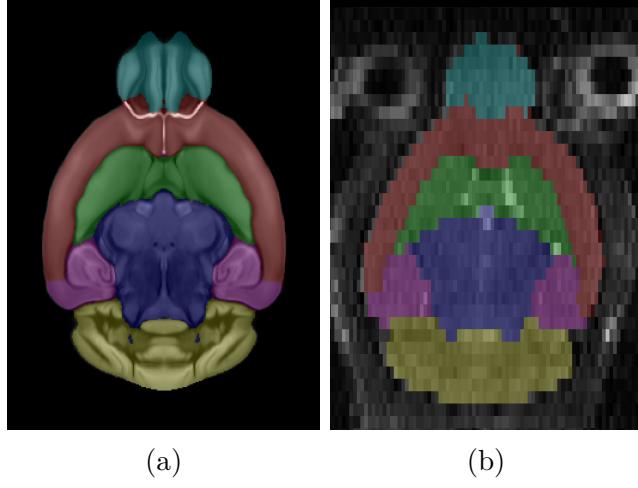


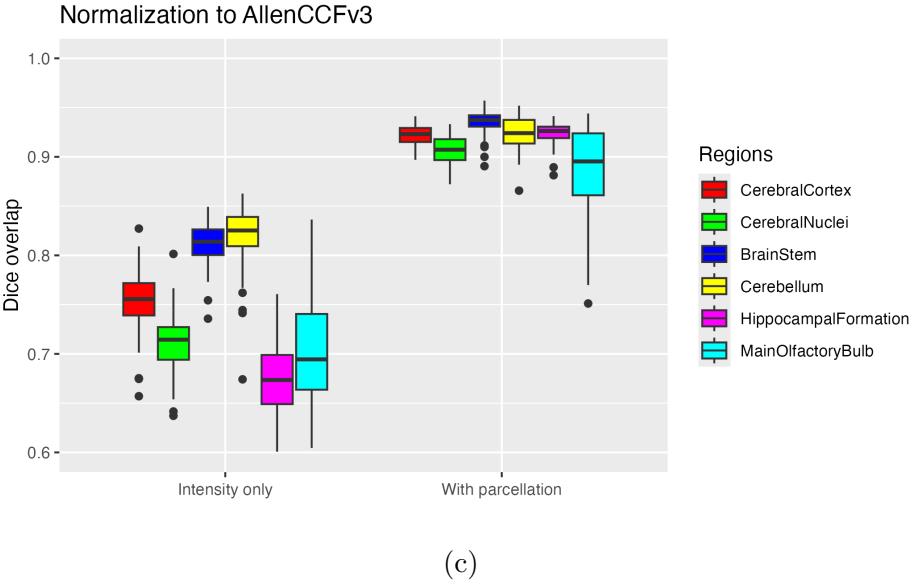
Figure 8: Evaluation of the ANTsX mouse brain extraction on an independent, publicly available dataset consisting of 12 specimens  $\times$  7 time points = 84 total images. Dice overlap comparisons with the user-generated brain masks provide good agreement with the automated results from the brain extraction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

422 For evaluation, we used an additional publicly available dataset<sup>70</sup> that is completely inde-  
 423 pendent from the data used in training the brain extraction and parcellation networks. Data  
 424 includes 12 specimens each imaged at seven time points (Day 0, Day 3, Week 1, Week 4,



(a)

(b)



(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 Week 8, Week 20) with in-house-generated brain masks for a total of 84 images. Spacing is  
426 anisotropic with an in-plane resolution of  $0.1 \times 0.1 mm^2$  and a slice thickness of  $0.5 mm$ .

427 Figure 8 summarizes the whole brain overlap between the provided segmentations for all  
428 84 images and the results of applying the proposed network. Also, since mapping to the  
429 AllenCCFv3 atlas is crucial for many mouse studies, we demonstrate the utility of the second  
430 network by leveraging the labeled regions to perform anatomically-explicit alignment using  
431 ANTsX multi-component registration instead of intensity-only registration. For these data,  
432 the whole brain extraction demonstrates excellent performance across the large age range.  
433 And although the intensity-only image registration provides adequate alignment, intensity  
434 with the regional parcellations significantly improves those measures.

435 **3 Discussion**

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

448 Our MERFISH pipeline provides an example of how to map high-resolution spatial transcripto-  
449 mic data into the AllenCCFv3. While the techniques employed for mapping the sectioned  
450 data can be generally applicable to spatially transform other serial histology images, much  
451 of the pipeline was designed to specifically address known alignment challenges in the MER-  
452 FISH data. Thus pipeline shows how general ANTsX tools can be adapted to target highly  
453 specialized problems in mouse cell type data.

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

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

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

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

<sup>492</sup> without requiring the integration of multiple packages or lengthy software development.  
<sup>493</sup> These open-source components not only perform well but are available across multiple plat-  
<sup>494</sup> forms which facilitates the construction of tailored pipelines for individual study solutions.  
<sup>495</sup> These components are also supported by years of development not only by the ANTsX  
<sup>496</sup> development team but by the larger ITK community.

497 **4 Methods**

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

503 **4.1 General ANTsX utilities**

504 Although they focus on distinct data types, the three pipelines presented share common  
505 components that are generally applicable when mapping mouse cell type data. These include,  
506 addressing intensity biases and noise in the data, image registration to solve the mapping,  
507 creating custom templates and atlases from the data, and visualization of the results. Table  
508 1 provides a brief summary of key general functionalities in ANTsX for addressing these  
509 challenges.

510 **4.1.1 Preprocessing: bias field correction and denoising**

511 Bias field correction and image denoising are standard preprocessing steps in improving over-  
512 all image quality in mouse brain images. The bias field, a gradual spatial intensity variation  
513 in images, can arise from various sources such as magnetic field inhomogeneity or acquisition  
514 artifacts, leading to distortions that can compromise the quality of brain images. Correct-  
515 ing for bias fields ensures a more uniform and consistent representation of brain structures,  
516 enabling more accurate quantitative analysis. Additionally, brain images are often suscep-  
517 tible to various forms of noise, which can obscure subtle features and affect the precision  
518 of measurements. Denoising techniques help mitigate the impact of noise, enhancing the  
519 signal-to-noise ratio and improving the overall image quality. The well-known N4 bias field  
520 correction algorithm<sup>52</sup> has its origins in the ANTs toolkit which was implemented and intro-  
521 duced into the ITK toolkit, i.e. `ants.n4_bias_field_correction(...)`. Similarly, ANTsX

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>).

522 contains an implementation of a well-performing patch-based denoising technique<sup>61</sup> and is  
523 also available as an image filter to the ITK community, `ants.denoise_image(...)`.

524 **4.1.2 Image registration**

525 The ANTs registration toolkit is a complex framework permitting highly tailored solutions  
526 to pairwise image registration scenarios<sup>85</sup>. It includes innovative transformation models  
527 for biological modeling<sup>51,67</sup> and has proven capable of excellent performance<sup>56,86</sup>. Vari-  
528 ous parameter sets targeting specific applications have been packaged with the different  
529 ANTsX packages, specifically ANTs, ANTsPy, and ANTsR<sup>46</sup>. In ANTsPy, the function  
530 `ants.registration(...)` is used to register a pair of images or a pair of image sets where  
531 `type_of_transform` is a user-specified option that invokes a specific parameter set. For ex-  
532 ample `type_of_transform='antsRegistrationSyNQuick[s]'` encapsulates an oft-used pa-  
533 rameter set for quick registration whereas `type_of_transform='antsRegistrationSyN[s]'`  
534 is a more aggressive alternative. Transforming images using the derived transforms is per-  
535 formed via the `ants.apply_transforms(...)` function.

536 Initially, linear optimization is initialized with center of (intensity) mass alignment typically  
537 followed by optimization of both rigid and affine transforms using the mutual information  
538 similarity metric. This is followed by diffeomorphic deformable alignment using symmetric  
539 normalization (SyN) with Gaussian<sup>51</sup> or B-spline regularization<sup>67</sup> where the forward trans-  
540 form is invertible and differentiable. The similarity metric employed at this latter stage is  
541 typically either neighborhood cross-correlation or mutual information. Note that these pa-  
542 rameter sets are robust to input image type (e.g., light sheet fluorescence microscopy, Nissl  
543 staining, and the various MRI modalities) and are adaptable to mouse image geometry and  
544 scaling. Further details can be found in the various documentation sources for these ANTsX  
545 packages.

546 **4.1.3 Template generation**

547 ANTsX provides functionality for constructing templates from a set (or multi-modal sets) of  
548 input images as originally described<sup>59</sup> and recently used to create the DevCCF templates<sup>16</sup>.

549 An initial template estimate is constructed from an existing subject image or a voxelwise  
550 average derived from a rigid pre-alignment of the image population. Pairwise registration  
551 between each subject and the current template estimate is performed using the Symmetric  
552 Normalization (SyN) algorithm<sup>51</sup>. The template estimate is updated by warping all subjects  
553 to the space of the template, performing a voxelwise average, and then performing a “shape  
554 update” of this latter image by warping it by the average inverse deformation, thus yielding  
555 a mean image of the population in terms of both intensity and shape. The corresponding  
556 ANTsPy function is `ants.build_template(...)`.

#### 557 4.1.4 Visualization

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

## 562 4.2 Mapping fMOST data to AllenCCFv3

### 563 4.2.1 Preprocessing

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

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

- 580 • *Stripe artifact removal.* Repetitive pattern artifacts are a common challenge in fMOST  
581 imaging where inhomogeneity during the cutting and imaging of different sections can  
582 leave stripes of hyper- and hypo-intensity across the image. These stripe artifacts  
583 can be latched onto by the registration algorithm as unintended features that are  
584 then misregistered to non-analogous structures in the AllenCCFv3. We address these  
585 artifacts by fitting a 3-D bandstop (notch) filter to target the frequency of the stripe  
586 patterns and removing them prior to the image registration.
- 587 • *Inhomogeneity correction.* Regional intensity inhomogeneity can also occur within  
588 and between sections in fMOST imaging due to staining or lighting irregularity during  
589 acquisition. Similar to stripe artifacts, intensity gradients due to inhomogeneity  
590 can be misconstrued as features during the mapping and result in matching of non-  
591 corresponding structures. Our pipeline addresses these intensity inhomogeneities using  
592 N4 bias field correction<sup>52</sup>, `ants.n4_bias_field_correction(...)`.

#### 593 4.2.2 Steps for spatial normalization to AllenCCFv3

- 594 1. *Average fMOST atlas as an intermediate target.* Due to the preparation of the mouse  
595 brain for fMOST imaging, the resulting structure in the mouse brain has several large  
596 morphological deviations from the AllenCCFv3 atlas. Most notable of these is an en-  
597 largement of the ventricles, and compression of cortical structures. In addition, there is  
598 poor intensity correspondence for the same anatomic features due to intensity dissim-  
599 ilarity between imaging modalities. We have found that standard intensity-base reg-  
600 istration is insufficient to capture the significant deformations required to map these  
601 structures correctly into the AllenCCFv3. We address this challenge in ANTsX by  
602 using explicitly corresponding parcellations of the brain, ventricles and surrounding

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

611 2. *Average fMOST atlas construction.* An intensity and shape-based contralaterally sym-  
612 metric average of the fMOST image data is constructed from 30 images and their  
613 contralateral flipped versions. We ran three iterations of the atlas construction using  
614 the default settings. Additional iterations (up to six) were evaluated and showed mini-  
615 mal changes to the final atlas construction, suggesting a convergence of the algorithm.

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

- 632     4. *Alignment of individual fMOST mouse brains.* The canonical transformation between  
633       the fMOST atlas and AllenCCFv3 greatly simplifies the registration of new individ-  
634       ual fMOST mouse brains into the AllenCCFv3. Each new image is first registered  
635       into the fMOST average atlas, which shares intensity, modality, and morphologi-  
636       cal characteristics. This allows us to leverage standard, intensity-based registration  
637       functionality<sup>85</sup> available in ANTsX to perform this alignment. Transformations are  
638       then concatenated to the original fMOST image to move it into the AllenCCFv3 space  
639       using `ants.apply_transforms(...)`.
- 640     5. *Transformation of single cell neurons.* A key feature of fMOST imaging is the ability  
641       to reconstruct and examine whole-brain single neuron projections<sup>79</sup>. Spatial mapping  
642       of these neurons from individual brains into the AllenCCFv3 allows investigators to  
643       study different neuron types within the same space and characterize their morphology  
644       with respect to their transcriptomics. Mappings found between the fMOST image  
645       and the AllenCCFv3 using our pipeline can be applied in this way to fMOST neuron  
646       reconstruction point set data using `ants.apply_transforms_to_points(..)`.

## 647     4.3 Mapping MERFISH data to AllenCCFv3

### 648     4.3.1 Preprocessing

- 649       • *Initial volume reconstruction.* Alignment of MERFISH data into a 3-D atlas space  
650       requires an estimation of anatomical structure within the data. For each section,  
651       this anatomic reference image was created by aggregating the number of detected  
652       genetic markers (across all probes) within each pixel of a  $10 \times 10 \mu\text{m}^2$  grid to match  
653       the resolution of the  $10 \mu\text{m}$  AllenCCFv3 atlas. These reference image sections are then  
654       coarsely reoriented and aligned across sections using manual annotations of the most  
655       dorsal and ventral points of the midline. The procedure produces an anatomic image  
656       stack that serves as an initialization for further global mappings into the AllenCCFv3.
- 657       • *Anatomical correspondence labeling.* Mapping the MERFISH data into the AllenCCFv3  
658       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.

685 **4.3.2 2.5D deformable, landmark-driven alignment to AllenCCFv3**

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

699 **4.4 DevCCF velocity flow transformation model**

700 Given multiple, linearly or non-linearly ordered point sets where individual points across the  
701 sets are in one-to-one correspondence, we developed an approach for generating a velocity  
702 flow transformation model to describe a time-varying diffeomorphic mapping as a variant of  
703 the landmark matching solution. Integration of the resulting velocity field can then be used  
704 to describe the displacement between any two time points within this time-parameterized  
705 domain. Regularization of the sparse correspondence between point sets is performed using  
706 a generalized B-spline scattered data approximation technique<sup>84</sup>, also created by the ANTsX  
707 developers and contributed to ITK.

708 **4.4.1 Velocity field optimization**

709 To apply this methodology to the developmental templates<sup>16</sup>, we coalesced the manual an-  
710 notations of the developmental templates into 26 common anatomical regions (see Figure 3).

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

## 723 4.5 ANTsXNet mouse brain applications

### 724 4.5.1 General notes regarding deep learning training

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

735 Data augmentation is crucial for generalizability and accuracy of the trained networks.  
736 Intensity-based data augmentation consisted of randomly added noise (i.e., Gaussian, shot,  
737 salt-and-pepper), simulated bias fields based on N4 bias field modeling, and histogram warp-

738 ing for mimicking well-known MRI intensity nonlinearities<sup>46,87</sup>. These augmentation tech-  
739 niques are available in ANTsXNet (only ANTsPyNet versions are listed with ANTsRNet  
740 versions available) and include:

- 741 • image noise: `ants.add_noise_to_image(...)`,
- 742 • simulated bias field: `antspynet.simulate_bias_field(...)`, and
- 743 • nonlinear intensity warping: `antspynet.histogram_warp_image_intensities(...)`.

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

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

#### 749 4.5.2 Brain extraction

750 Similar to human neuroimage processing, brain extraction is a crucial preprocessing step  
751 for accurate brain mapping. We developed similar functionality for T2-weighted mouse  
752 brains. This network uses a conventional U-net architecture<sup>88</sup> and, in ANTsPyNet, this  
753 functionality is available in the program `antspynet.mouse_brain_extraction(...)`.  
754 For the two-shot T2-weighted brain extraction network, two brain templates were gen-  
755 erated along with their masks. One of the templates was generated from orthogonal  
756 multi-plane, high resolution data<sup>69</sup> which were combined to synthesize isotropic volu-  
757 metric data using the B-spline fitting algorithm<sup>84</sup>. This algorithm is encapsulated in  
758 `ants.fit_bspline_object_to_scattered_data(...)` where the input is the set of voxel  
759 intensity values and each associated physical location. Since each point can be assigned  
760 a confidence weight, we use the normalized gradient value to more heavily weight edge  
761 regions. Although both template/mask pairs are available in the GitHub repository  
762 associated with this work, the synthesized volumetric B-spline T2-weighted pair is available  
763 within ANTsXNet through the calls:

- 764     • template: `antspynet.get_antsxnet_data("bsplineT2MouseTemplate")` and  
765     • mask: `antspynet.get_antsxnet_data("bsplineT2MouseTemplateBrainMask")`.

766     **4.5.3 Brain parcellation**

767     The T2-weighted brain parcellation network is also based on a 3-D U-net architecture and the  
768     T2-w DevCCF P56 template component with extensive data augmentation, as described pre-  
769     viously. Intensity differences between the template and any brain extracted input image are  
770     minimized through the use of the rank intensity transform (`ants.rank_intensity(...)`).  
771     Shape differences are reduced by the additional preprocessing step of warping the brain ex-  
772     tracted input image to the template. Additional input channels include the prior probability  
773     images created from the template parcellation. These images are also available through the  
774     ANTsXNet `get_antsxnet_data(...)` interface.

775 **Data availability**

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

<sup>785</sup> **Acknowledgments**

<sup>786</sup> Support for the research reported in this work includes funding from the National Institute  
<sup>787</sup> of Biomedical Imaging and Bioengineering (R01-EB031722) and National Institute of Mental  
<sup>788</sup> Health (RF1-MH124605 and U24-MH114827).

<sup>789</sup> **Author contributions**

<sup>790</sup> N.T., M.C., and J.G. wrote the main manuscript text and figures. M.C., M.K., R.D., S.S.,  
<sup>791</sup> Q.W., L.G., J.D., C.G., and J.G. developed the Allen registration pipelines. N.T. and F.K.  
<sup>792</sup> developed the time-varying velocity transformation model for the DevCCF. N.T. and M.T.  
<sup>793</sup> developed the brain parcellation and cortical thickness methodology. All authors reviewed  
<sup>794</sup> the manuscript.

795 **References**

- 796 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).
- 797 2. La Manno, G. *et al.* Molecular architecture of the developing mouse brain. *Nature*  
**596**, 92–96 (2021).
- 798 3. Wen, L. *et al.* Single-cell technologies: From research to application. *Innovation  
(Camb)* **3**, 100342 (2022).
- 799 4. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–14  
(2014).
- 800 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).
- 801 6. Li, A. *et al.* Micro-optical sectioning tomography to obtain a high-resolution atlas of  
the mouse brain. *Science* **330**, 1404–8 (2010).
- 802 7. Ueda, H. R. *et al.* Tissue clearing and its applications in neuroscience. *Nat Rev  
Neurosci* **21**, 61–79 (2020).
- 803 8. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by  
spatial transcriptomics. *Science* **353**, 78–82 (2016).
- 804 9. Burgess, D. J. Spatial transcriptomics coming of age. *Nat Rev Genet* **20**, 317 (2019).
- 805 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).
- 806 11. Hawrylycz, M. *et al.* A guide to the BRAIN initiative cell census network data  
ecosystem. *PLoS biology* **21**, e3002133 (2023).
- 807 12. Wang, Q. *et al.* The allen mouse brain common coordinate framework: A 3D reference  
atlas. *Cell* **181**, 936–953.e20 (2020).
- 808 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).
- 809 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).

- 810 15. Qu, L. *et al.* Cross-modal coherent registration of whole mouse brains. *Nature Methods* **19**, 111–118 (2022).
- 811 16. 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).
- 812 17. Chuang, N. *et al.* An MRI-based atlas and database of the developing mouse brain. *Neuroimage* **54**, 80–89 (2011).
- 813 18. Dries, R. *et al.* Advances in spatial transcriptomic data analysis. *Genome research* **31**, 1706–1718 (2021).
- 814 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).
- 815 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).
- 816 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).
- 817 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).
- 818 23. Cahill, L. S. *et al.* Preparation of fixed mouse brains for MRI. *Neuroimage* **60**, 933–939 (2012).
- 819 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).
- 820 25. Kim, Y. *et al.* Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell* **171**, 456–469 (2017).
- 821 26. Fürth, D. *et al.* An interactive framework for whole-brain maps at cellular resolution. *Nat Neurosci* **21**, 139–149 (2018).

- 822 27. Li, Y. *et al.* mBrainAligner-web: A web server for cross-modal coherent registration  
of whole mouse brains. *Bioinformatics* **38**, 4654–4655 (2022).
- 823 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).
- 824 29. Eastwood, B. S. *et al.* Whole mouse brain reconstruction and registration to a ref-  
erence atlas with standard histochemical processing of coronal sections. *Journal of  
Comparative Neurology* **527**, 2170–2178 (2019).
- 825 30. Ni, H. *et al.* A robust image registration interface for large volume brain atlas. *Sci  
Rep* **10**, 2139 (2020).
- 826 31. Pallast, N. *et al.* Processing pipeline for atlas-based imaging data analysis of struc-  
tural and functional mouse brain MRI (AIDAmri). *Front Neuroinform* **13**, 42 (2019).
- 827 32. Celestine, M., Nadkarni, N. A., Garin, C. M., Bougacha, S. & Dhenain, M. **Sammbar-  
MRI: A library for processing SmAll-MaMmal BrAin MRI data in python.** *Front  
Neuroinform* **14**, 24 (2020).
- 828 33. Ioanas, H.-I., Marks, M., Zerbi, V., Yanik, M. F. & Rudin, M. **An optimized regis-  
tration workflow and standard geometric space for small animal brain imaging.** *Neu-  
roimage* **241**, 118386 (2021).
- 829 34. Perens, J. *et al.* Multimodal 3D mouse brain atlas framework with the skull-derived  
coordinate system. *Neuroinformatics* **21**, 269–286 (2023).
- 830 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).
- 831 36. Goubran, M. *et al.* **Multimodal image registration and connectivity analysis for inte-  
gration of connectomic data from microscopy to MRI.** *Nat Commun* **10**, 5504 (2019).
- 832 37. Chandrashekhar, V. *et al.* CloudReg: Automatic terabyte-scale cross-modal brain  
volume registration. *Nature methods* **18**, 845–846 (2021).
- 833 38. Jin, M. *et al.* **SMART: An open-source extension of WholeBrain for intact mouse  
brain registration and segmentation.** *eNeuro* **9**, (2022).

- 834 39. Negwer, M. *et al.* FriendlyClearMap: An optimized toolkit for mouse brain mapping  
and analysis. *Gigascience* **12**, (2022).
- 835 40. Lin, W. *et al.* Whole-brain mapping of histaminergic projections in mouse brain.  
*Proceedings of the National Academy of Sciences* **120**, e2216231120 (2023).
- 836 41. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by  
MERFISH. *Nature* **598**, 137–143 (2021).
- 837 42. Shi, H. *et al.* Spatial atlas of the mouse central nervous system at molecular resolution.  
*Nature* **622**, 552–561 (2023).
- 838 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).
- 839 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).
- 840 45. Fedorov, A. *et al.* 3D slicer as an image computing platform for the quantitative  
imaging network. *Magnetic resonance imaging* **30**, 1323–1341 (2012).
- 841 46. Tustison, N. J. *et al.* The ANTsX ecosystem for quantitative biological and medical  
imaging. *Sci Rep* **11**, 9068 (2021).
- 842 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).
- 843 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).
- 844 49. Allan Johnson, G. *et al.* Whole mouse brain connectomics. *Journal of Comparative  
Neurology* **527**, 2146–2157 (2019).
- 845 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).

- 846 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).
- 847 52. Tustison, N. J. *et al.* [N4ITK: Improved N3 bias correction](#). *IEEE Trans Med Imaging*  
**29**, 1310–20 (2010).
- 848 53. Bajcsy, R. & Broit, C. Matching of deformed images. in *Sixth International Conference on Pattern Recognition (ICPR'82)* 351–353 (1982).
- 849 54. Bajcsy, R. & Kovacic, S. [Multiresolution elastic matching](#). *Computer Vision, Graphics, and Image Processing* **46**, 1–21 (1989).
- 850 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).
- 851 56. Klein, A. *et al.* [Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration](#). *Neuroimage* **46**, 786–802 (2009).
- 852 57. Murphy, K. *et al.* [Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge](#). *IEEE Trans Med Imaging* **30**, 1901–20 (2011).
- 853 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).
- 854 59. Avants, B. B. *et al.* [The optimal template effect in hippocampus studies of diseased populations](#). *Neuroimage* **49**, 2457–66 (2010).
- 855 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).
- 856 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).
- 857 62. Wang, H. *et al.* [Multi-atlas segmentation with joint label fusion](#). *IEEE Trans Pattern Anal Mach Intell* **35**, 611–23 (2013).

- 858 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](https://doi.org/10.1007/s12021-014-9245-2).
- 859 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).
- 860 65. McCormick, M., Liu, X., Jomier, J., Marion, C. & Ibanez, L. **ITK: Enabling reproducible research and open science**. *Front Neuroinform* **8**, 13 (2014).
- 861 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).
- 862 67. Tustison, N. J. & Avants, B. B. **Explicit B-spline regularization in diffeomorphic image registration**. *Front Neuroinform* **7**, 39 (2013).
- 863 68. Hsu, L.-M. *et al.* CAMRI mouse brain MRI data.
- 864 69. Reshetnikov, V. *et al.* High-resolution MRI data of brain C57BL/6 and BTBR mice in three different anatomical views.
- 865 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).
- 866 71. Liu, J. *et al.* **Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing**. *Life Sci Alliance* **6**, (2023).
- 867 72. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. **Cellpose: A generalist algorithm for cellular segmentation**. *Nat Methods* **18**, 100–106 (2021).
- 868 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).
- 869 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).

- 870 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).
- 871 76. Gong, H. *et al.* High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. *Nat Commun* **7**, 12142 (2016).
- 872 77. Wang, J. *et al.* Divergent projection patterns revealed by reconstruction of individual neurons in orbitofrontal cortex. *Neurosci Bull* **37**, 461–477 (2021).
- 873 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).
- 874 79. Peng, H. *et al.* Morphological diversity of single neurons in molecularly defined cell types. *Nature* **598**, 174–181 (2021).
- 875 80. 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).
- 876 81. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci* **19**, 335–46 (2016).
- 877 82. Bergmann, E., Gofman, X., Kavushansky, A. & Kahn, I. Individual variability in functional connectivity architecture of the mouse brain. *Commun Biol* **3**, 738 (2020).
- 878 83. Billot, B. *et al.* SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. *Med Image Anal* **86**, 102789 (2023).
- 879 84. 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).
- 880 85. Avants, B. B. *et al.* The Insight ToolKit image registration framework. *Front Neuroinform* **8**, 44 (2014).
- 881 86. Avants, B. B. *et al.* A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **54**, 2033–44 (2011).

- 882 87. 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).
- 883 88. Falk, T. *et al.* U-net: Deep learning for cell counting, detection, and morphometry.  
*Nat Methods* **16**, 67–70 (2019).