

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

<sup>4</sup> 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  
<sup>5</sup> 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>,  
<sup>6</sup> Yongsoo Kim<sup>3</sup>, and James C. Gee<sup>2</sup>

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

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

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

<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

J. C. G. B. D.

19 James C. Gee, PhD  
20 Department of Radiology

20 Department of Radiol  
21 University of Pennsylvania

21 University of Pe  
22 geo@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 have recently begun aggregating the most compre-  
<sup>26</sup> hensive reference database to date for diverse cell type profiling of the mouse brain, which  
<sup>27</sup> encompasses over 40 different multi-modal profiling techniques from more than 30 research  
<sup>28</sup> groups. One central challenge for this integrative effort across different investigators and  
<sup>29</sup> laboratories has been the need to map these unique datasets into common reference spaces  
<sup>30</sup> such that the spatial, structural, and functional information from different cell types can be  
<sup>31</sup> jointly analyzed across modalities. However, significant variations in the acquisition, tissue  
<sup>32</sup> processing, and imaging techniques across data types makes mapping such diverse data a  
<sup>33</sup> multifarious problem. Different data types exhibit unique tissue distortion and signal char-  
<sup>34</sup> acteristics that precludes a single mapping strategy from being generally applicable across  
<sup>35</sup> all cell type data. Diverse, and often specialized, mapping approaches are needed to address  
<sup>36</sup> the unique barriers present in each modality. This work highlights modular atlas mapping  
<sup>37</sup> strategies developed across three separate BICCN studies using the Advanced Normalization  
<sup>38</sup> Tools Ecosystem (ANTsX) to map spatial transcriptomic (MERFISH) and high-resolution  
<sup>39</sup> morphology (fMOST) mouse brain data into the Allen Common Coordinate Framework  
<sup>40</sup> (AllenCCFv3), and developmental (MRI and LSFM) data into the Developmental Common  
<sup>41</sup> Coordinate Framework (DevCCF). We discuss both common mapping strategies that can be  
<sup>42</sup> shared across modalities, and targeted strategies driven by specific challenges from each data  
<sup>43</sup> type. Novel open-source contributions, made publicly available through ANTSX, include a  
<sup>44</sup> generic velocity flow-based approach for continuously mapping developmental trajectories  
<sup>45</sup> such as that characterizing the DevCCF as well as an automated framework for determining  
<sup>46</sup> structural morphology made possible through the leveraging of public resources such as the  
<sup>47</sup> AllenCCFv3 and the DevCCF. Finally, we provide general guidance to aid investigators in  
<sup>48</sup> their efforts to tailor these strategies to address unique challenges in their data without the  
<sup>49</sup> need to develop additional specialized software.

## 50 1 Introduction

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

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

76 **1.1 Mouse brain mapping**

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

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

<sup>105</sup> difficult to apply the same mapping to their own data without direct collaboration with the  
<sup>106</sup> platform owners.

<sup>107</sup> The second category are specialized approaches specifically designed for mapping one or  
<sup>108</sup> more modalities into a CCF. These approaches use combinations of specialized manual and  
<sup>109</sup> automated processes that address specific challenges in each modality. Examples include ap-  
<sup>110</sup> proaches for mapping histology<sup>28–30</sup>, magnetic resonance imaging (MRI)<sup>37</sup>, micro-computed  
<sup>111</sup> tomography (microCT)<sup>35,37</sup>, light-sheet fluorescence microscopy (LSFM)<sup>34,36–39</sup>, fluorescence  
<sup>112</sup> micro-optical sectioning tomography (fMOST)<sup>15,40</sup> and transcriptomic data<sup>41–43</sup>. As special-  
<sup>113</sup> ized approaches, these techniques tend to boast higher mapping accuracy, robustness, and  
<sup>114</sup> ease of use when ran with applicable modalities. Conversely, their specialized designs often  
<sup>115</sup> rely on base assumptions regarding the data type that can make them rigid and difficult  
<sup>116</sup> to adapt for new modalities or unexpected artifacts and distortions in the data. Retooling  
<sup>117</sup> these specialize software to use with new data can require significant development, validation  
<sup>118</sup> time, and engineering expertise that may not be readily available for all investigators.

<sup>119</sup> The last category are modular mapping approaches constructed using general image analy-  
<sup>120</sup> sis toolkits, which are software packages that include varied collections of image processing,  
<sup>121</sup> segmentation and registration tools that have been previously developed, and validated for  
<sup>122</sup> multiple application areas. Examples of such toolkits include elastix<sup>44</sup>, Slicer3D<sup>45</sup>, ANTsX<sup>46</sup>,  
<sup>123</sup> and several others which have all been applied towards mouse brain spatial mapping. The  
<sup>124</sup> main challenge, in these mouse-specific study scenarios, is that tailored pipelines often need  
<sup>125</sup> be constructed from available software components. Investigators must therefore be familiar  
<sup>126</sup> with the these tools for formulating new or adapting existing pipelines. However, in com-  
<sup>127</sup> parison to previously described specialized mapping approaches, these approaches are often  
<sup>128</sup> easier to create and prone to robustness, being typically constructed from pipelin compo-  
<sup>129</sup> nents which have been previously vetted in other contexts. In this work, we highlight such  
<sup>130</sup> mapping strategies designed using the ANTsX framework to map three distinct mouse cell  
<sup>131</sup> type data with different characteristics into existing CCFs.

<sup>132</sup> **1.2 Advanced Normalization Tools (ANTsX)**

<sup>133</sup> The Advanced Normalization Tools (ANTsX) framework has been used in a number of  
<sup>134</sup> applications for mapping mouse brain data as part of core processing steps in various  
<sup>135</sup> workflows<sup>30,47–50</sup>, particularly its pairwise, intensity-based image registration capabilities and  
<sup>136</sup> bias field correction. Historically, ANTsX development is originally based on fundamental  
<sup>137</sup> approaches to image mapping<sup>51–53</sup>, particularly in the human brain, which has resulted  
<sup>138</sup> in core contributions to the field such as the widely-used Symmetric Normalization (SyN)  
<sup>139</sup> algorithm<sup>54</sup>. Since its development, various independent platforms have been used to eval-  
<sup>140</sup> uate ANTsX image registration capabilities in the context of different application foci which  
<sup>141</sup> include multi-site brain MRI data<sup>55</sup>, pulmonary CT data<sup>56</sup>, and most recently, multi-modal  
<sup>142</sup> brain registration in the presence of tumors<sup>57</sup>.

<sup>143</sup> Apart from its registration capabilities, ANTsX comprises additional functionality such  
<sup>144</sup> as template generation<sup>58</sup>, intensity-based segmentation<sup>59</sup>, preprocessing<sup>60,61</sup>, deep learning  
<sup>145</sup> networks<sup>46</sup>, and other utilities relevant to brain mapping (see Methods - Table 1). The  
<sup>146</sup> use of the toolkit has demonstrated high performance in multiple application areas (e.g.,  
<sup>147</sup> consensus labeling<sup>62</sup>, brain tumor segmentation<sup>63</sup>, and cardiac motion estimation<sup>64</sup> ). Im-  
<sup>148</sup> portantly, ANTs is built on the Insight Toolkit (ITK)<sup>65</sup> deriving benefit from the open-source  
<sup>149</sup> community of scientists and programmers as well as providing an important resource for al-  
<sup>150</sup> gorithmic development, evaluation, and improvement. In this paper we demonstrate how  
<sup>151</sup> ANTs' comprehensive toolset provides the basis to develop modular frameworks for map-  
<sup>152</sup> ping diverse mouse cell type data into common coordinate frameworks (CCFs). Specifically,  
<sup>153</sup> we highlight its application for mapping data from three separate BICCN projects focused  
<sup>154</sup> on distinct data types: morphology data using fluorescence micro-optical sectioning tomog-  
<sup>155</sup> raphy (fMOST), spatial transcriptomics from multiplexed error-robust fluorescence in situ  
<sup>156</sup> hybridization (MERFISH) data, and time-series developmental data using light sheet fluores-  
<sup>157</sup> cence microscopy (LSFM) and magnetic resonance imaging (MRI). We describe both shared  
<sup>158</sup> and targeted strategies developed to address the specific challenges of these modalities.

<sup>159</sup> **1.3 Novel ANTsX-based open-source contributions**

<sup>160</sup> We introduce two novel inclusions to the ANTsX toolset that were developed as part of  
<sup>161</sup> the MRI mapping and analysis pipeline for the Developmental Common Coordinate Frame-  
<sup>162</sup> work (DevCCF). Consistent with previous ANTsX development, newly introduced capa-  
<sup>163</sup> bilities introduced below are available through ANTsX (specifically, via R and Python  
<sup>164</sup> ANTsX packages), and illustrated through self-contained examples in the ANTsX tuto-  
<sup>165</sup> rial (<https://tinyurl.com/antsxtutorial>) with a dedicated GitHub repository specific to this  
<sup>166</sup> work (<https://github.com/ntustison/ANTsXMouseBrainMapping>). To complement stan-  
<sup>167</sup> dard preprocessing steps (e.g., bias correction, brain masking), additional mouse brain  
<sup>168</sup> specific tools have also been introduced to the ANTsX ecosystem, such as section recon-  
<sup>169</sup> struction and landmark-based alignment with corresponding processing scripts available at  
<sup>170</sup> <https://github.com/dontminchenit/CCFAlignmentToolkit>.

<sup>171</sup> **1.3.1 The DevCCF velocity flow model**

<sup>172</sup> Recently, the Developmental Common Coordinate Framework (DevCCF) was introduced to  
<sup>173</sup> the mouse brain research community as a public resource<sup>16</sup> comprising symmetric atlases  
<sup>174</sup> of multimodal image data and anatomical segmentations defined by developmental ontol-  
<sup>175</sup> ogy. These templates sample the mouse embryonic days (E) 11.5, E13.5, E15.5, E18.5 and  
<sup>176</sup> postnatal day (P) 4, P14, and P56. Modalities include light sheet fluorescence microscopy  
<sup>177</sup> (LSFM) and at least four MRI contrasts per developmental stage. Anatomical parcellations  
<sup>178</sup> are also available for each time point and were generated from ANTsX-based mappings of  
<sup>179</sup> gene expression and other cell type data. Additionally, the P56 template was integrated  
<sup>180</sup> with the AllenCCFv3 to further enhance the practical utility of the DevCCF. These pro-  
<sup>181</sup> cesses, specifically template generation and multi-modal image mapping, were performed  
<sup>182</sup> using ANTsX functionality in the presence of image mapping difficulties such as missing  
<sup>183</sup> data and tissue distortion.<sup>16</sup>

<sup>184</sup> Given the temporal gaps in the discrete set of developmental atlases, we also provide an  
<sup>185</sup> open-source framework for inferring correspondence within the temporally continuous do-

<sup>186</sup> main sampled by the existing set of embryonic and postnatal atlases of the DevCCF. This  
<sup>187</sup> recently developed functionality permits the generation of a diffeomorphic velocity flow trans-  
<sup>188</sup> formation model<sup>66</sup>, influenced by previous work<sup>67</sup>. The resulting time-parameterized veloc-  
<sup>189</sup> ity field spans the stages of the DevCCF where mappings between any two continuous time  
<sup>190</sup> points within the span bounded by the E11.5 and P56 atlases is determined by integration  
<sup>191</sup> of the optimized velocity field.

### <sup>192</sup> 1.3.2 Structural morphology and parcellations of the mouse brain

<sup>193</sup> In contrast to the pipeline development in human data<sup>46</sup>, limited tools exist yet to cre-  
<sup>194</sup> ate adequate training data for automated parcellations of the mouse brain. In addition,  
<sup>195</sup> mouse brain data acquisition often has unique issues, such as lower data quality or sampling  
<sup>196</sup> anisotropy which limits its applicability to high resolution resources (e.g., AllenCCFv3, De-  
<sup>197</sup> vCCF), specifically with respect to the corresponding granular brain parcellations derived  
<sup>198</sup> from numerous hours of expert annotation leveraging multimodal imaging resources.

<sup>199</sup> Herein, we introduce a mouse brain parcellation pipeline for T2-weighted (T2-w) MRI com-  
<sup>200</sup> prising two novel deep learning components: two-shot learning brain extraction from data  
<sup>201</sup> augmentation of two ANTsX templates generated from two open datasets<sup>68,69</sup> and single-  
<sup>202</sup> shot brain parcellation derived from the AllenCCFv3 labelings mapped to the corresponding  
<sup>203</sup> DevCCF P56 T2-w component. Although we anticipate that this pipeline will be beneficial  
<sup>204</sup> to the research community, this work demonstrates more generally how one can leverage  
<sup>205</sup> ANTsX tools for developing tailored brain parcellation schemes using publicly available re-  
<sup>206</sup> sources. Evaluation is performed on an independent open dataset<sup>70</sup> comprising longitudinal  
<sup>207</sup> acquisitions of multiple specimens.

208 **2 Results**

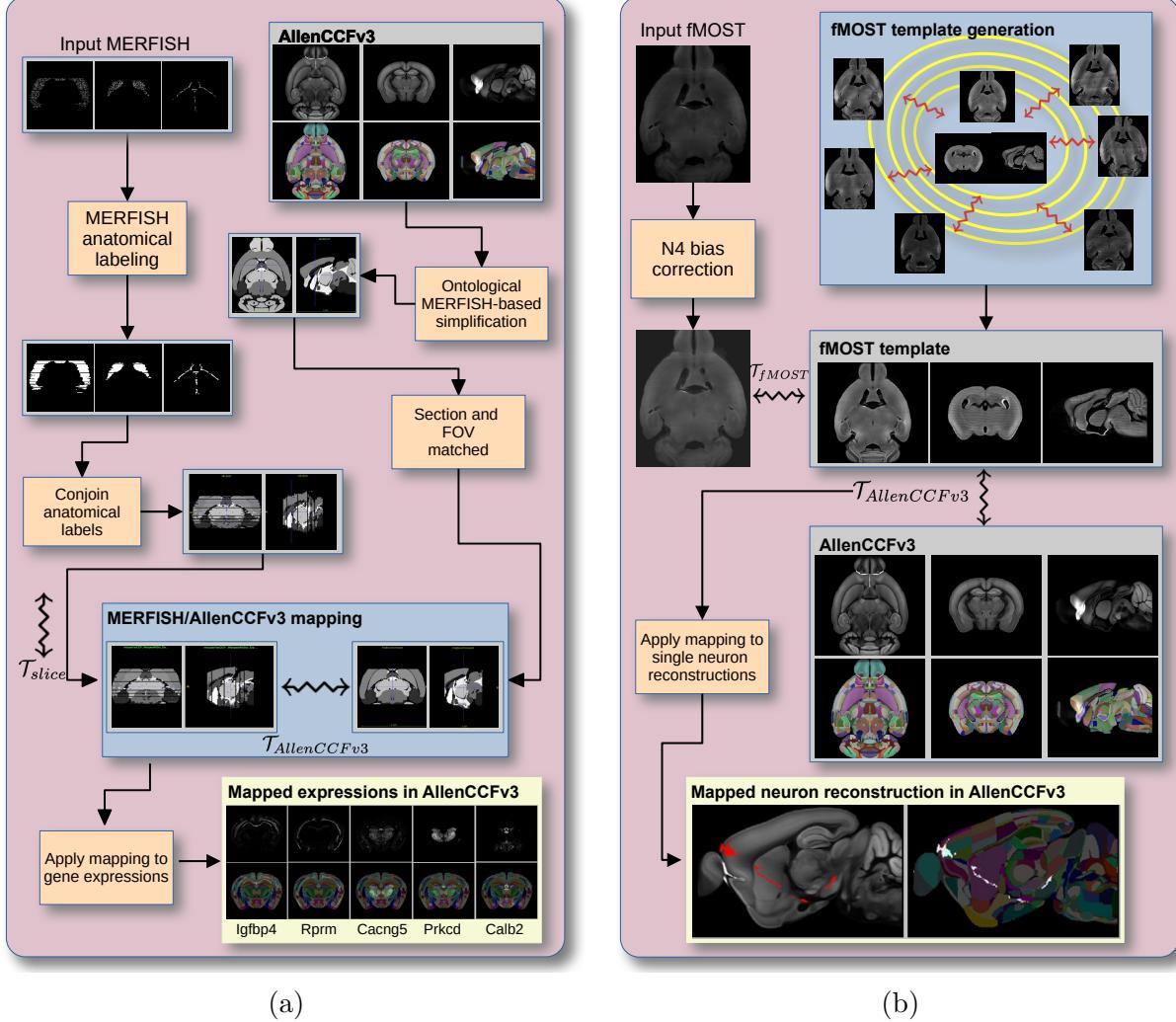


Figure 1: Diagram of the two ANTsX-based pipelines for mapping (a) MERFISH and (b)fMOST data into the space of AllenCCFv3. Each generates the requisite transforms,  $\mathcal{T}$ , to map individual images.

209 **2.1 AllenCCFv3 brain image mapping**

210 **2.1.1 Mapping multiplexed error-robust fluorescence *in situ* hybridization  
211 (MERFISH) data**

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

214 data onto the AllenCCFv3 (see Figure 1(a)). This pipeline, used recently in creating a  
215 high-resolution transcriptomic atlas of the mouse brain<sup>50</sup>, performs mappings by first gen-  
216 erating anatomical labels from tissue related gene expressions in the MERFISH data, and  
217 then spatially matching these labels to corresponding anatomical tissue parcellations in the  
218 AllenCCFv3. The pipeline consists of MERFISH data specific preprocessing which includes  
219 section reconstruction, mapping corresponding anatomical labels between AllenCCFv3 and  
220 the spatial transcriptomic maps of the MERFISH data, and matching MERFISH sections to  
221 the atlas space. Following pre-processing, two main alignment steps were performed: 1) 3-D  
222 global affine mapping and section matching of the AllenCCFv3 into the MERFISH data and  
223 2) 2D global and deformable mapping between each MERFISH section and matched AllenC-  
224 CFv3 section. Mappings learned via each step in the pipeline are preserved and concatenated  
225 to provide point-to-point correspondence between the original MERFISH data and AllenC-  
226 CFv3, thus allowing individual gene expressions to be transferred into the AllenCCFv3.

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

240 **Evaluation.** Alignment of the MERFISH data into the AllenCCFv3 was qualitatively as-  
241 sessed by an expert anatomist at each iteration of the registration using known correspon-  
242 dence of gene markers and their associations with the AllenCCFv3. As previously reported<sup>50</sup>,

<sup>243</sup> further assessment of the alignment showed that, of the 554 terminal regions (gray matter  
<sup>244</sup> only) in the AllenCCFv3, only seven small subregions were missed from the MERFISH  
<sup>245</sup> dataset: frontal pole, layer 1 (FRP1), FRP2/3, FRP5; accessory olfactory bulb, glomerular  
<sup>246</sup> layer (AOBgl); accessory olfactory bulb, granular layer (AOBgr); accessory olfactory bulb,  
<sup>247</sup> mitral layer (AOBmi); and accessory supraoptic group (ASO).

### <sup>248</sup> 2.1.2 Mapping fluorescence micro-optical sectioning tomography (fMOST) data

<sup>249</sup> **Overview.** We developed a pipeline for mapping fluorescence micro-optical sectioning to-  
<sup>250</sup> mography (fMOST) mouse brain images into the AllenCCFv3 (see Figure 1(b)). The pipeline  
<sup>251</sup> is adapted from previously developed frameworks for human brain mapping<sup>58</sup>, and uses a  
<sup>252</sup> modality specific (fMOST) average atlas to assist in the image registration and mapping.  
<sup>253</sup> This approach has been well validated in human studies<sup>73–75</sup>, and successfully used in other  
<sup>254</sup> mouse data<sup>12,15,34</sup>. Briefly, we construct an intensity- and shape-based average fMOST atlas  
<sup>255</sup> using 30 fMOST images to serve as an intermediate registration target for mapping fMOST  
<sup>256</sup> images from individual specimens into the AllenCCFv3. Preprocessing steps include down-  
<sup>257</sup> sampling to match the 25 $\mu$ m isotropic AllenCCFv3, acquisition-based stripe artifact removal,  
<sup>258</sup> and inhomogeneity correction<sup>61</sup>. Preprocessing also includes a single annotation-driven reg-  
<sup>259</sup> istration to establish a canonical mapping between the fMOST atlas and the AllenCCFv3.  
<sup>260</sup> This step allows us to align expert determined landmarks to accurately map structures  
<sup>261</sup> with large morphological differences between the modalities, which are difficult to address  
<sup>262</sup> using standard approaches. Once this canonical mapping is established, standard intensity-  
<sup>263</sup> based registration is used to align each new fMOST image to the fMOST specific atlas.  
<sup>264</sup> This mapping is concatenated with the canonical fMOST atlas-to-AllenCCFv3 mapping to  
<sup>265</sup> further map each individual brain into the latter without the need to generate additional  
<sup>266</sup> landmarks. Transformations learned through this mapping can be applied to single neuron  
<sup>267</sup> reconstructions from the fMOST images to evaluate neuronal distributions across different  
<sup>268</sup> specimens into the AllenCCFv3 for the purpose of cell census analyses.

<sup>269</sup> **Data.** The high-throughput and high-resolution fluorescence micro-optical sectioning to-  
<sup>270</sup> mography (fMOST)<sup>76,77</sup> platform was used to image 55 mouse brains containing gene-defined

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

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

285 **2.2 The DevCCF velocity flow model**

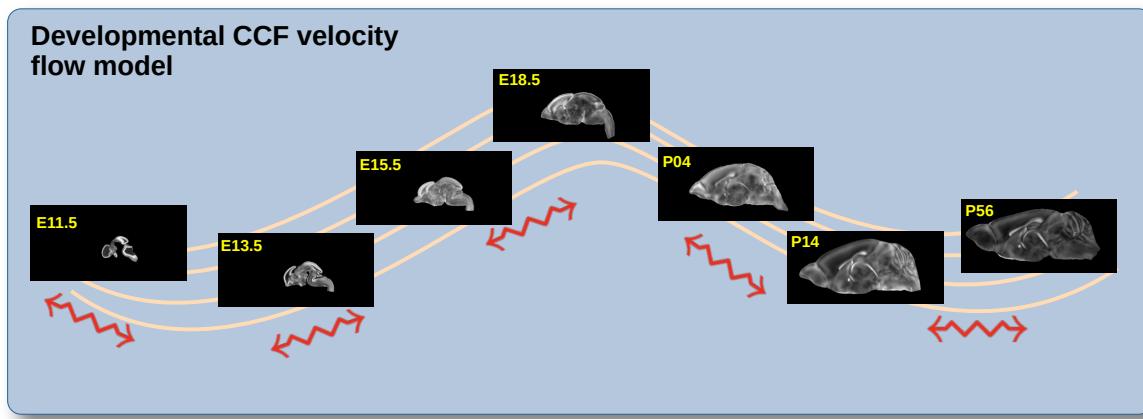


Figure 2: The spatial transformation between any two time points within the DevCCF longitudinal developmental trajectory is available through the use of ANTsX functionality for generating a velocity flow model.

286 To continuously interpolate transformations between the different stages of the DevCCF

atlases, a velocity flow model was constructed using DevCCF derived data and functionality recently introduced into both the ANTsR and ANTsPy packages. Both platforms include a complete suite of functions for determining dense correspondence from sparse landmarks based on a variety of transformation models ranging from standard linear models (i.e., rigid, affine) to deformable diffeomorphic models (e.g, symmetric normalization<sup>54</sup>). The latter set includes transformation models for both the pairwise scenario and for multiple sets, as in the case of the DevCCF. 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

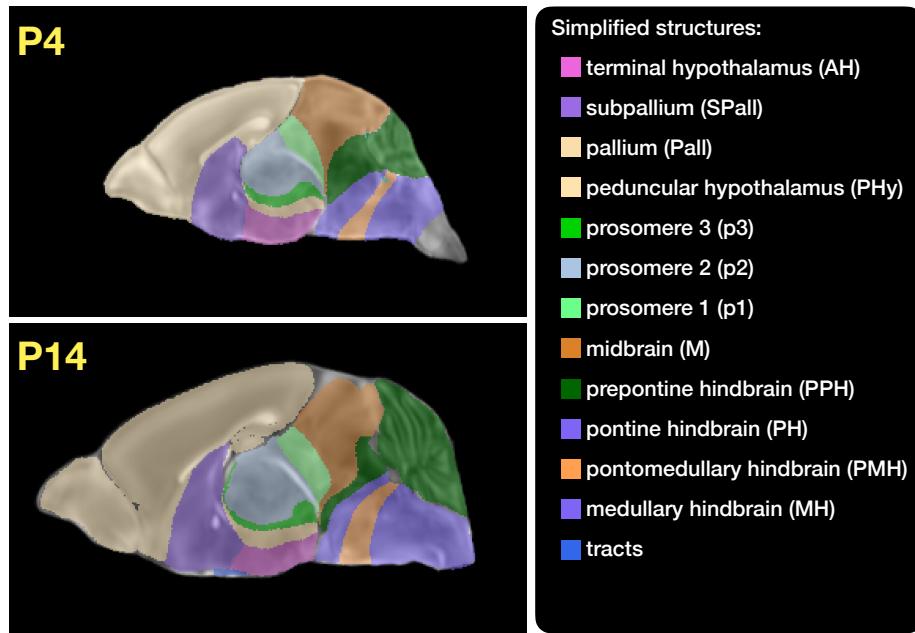


Figure 3: Annotated regions representing common labels across developmental stages which are illustrated for both P4 and P14.

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

302 hypothalamus, prosomere, prosomere, prosomere, midbrain, prepontine hindbrain, pontine  
303 hindbrain, pontomedullary hindbrain, medullary hindbrain, and tracts (see Figure 3).

304 Prior to velocity field optimization, all data were rigidly transformed to DevCCF P56 using  
305 the centroids of the common label sets. In order to determine the landmark correspondence  
306 across DevCCF stages, the multi-metric capabilities of `ants.registration(...)` were used.  
307 Instead of performing intensity-based pairwise registration directly on these multi-label im-  
308 ages, each label was used to construct a separate fixed and moving image pair resulting in a  
309 multi-metric registration optimization scenario involving 24 binary image pairs (each label  
310 weighted equally) for optimizing diffeomorphic correspondence between neighboring time  
311 point atlases using the mean squares metric and the symmetric normalization transform<sup>54</sup>.  
312 To generate the set of common point sets across all seven developmental atlases, the label  
313 boundaries and whole regions were sampled in the P56 atlas and then propagated to each  
314 atlas using the transformations derived from the pairwise registrations. We selected a sam-  
315 pling rate of 10% for the contour points and 1% for the regional points for a total number  
316 of points being per atlas being 173303 ( $N_{contour} = 98151$  and  $N_{region} = 75152$ ). Regional  
317 boundary points were weighted twice as those of non-boundary points during optimization.

### 318 2.2.2 Optimization

319 The velocity field was optimized using the input composed of the seven corresponding point  
320 sets and their associated weight values, the selected number of integration points for the  
321 velocity field ( $N = 11$ ), and the parameters defining the geometry of the spatial dimensions  
322 of the velocity field. Thus, the optimized velocity field described here is of size [256, 182, 360]  
323 (50 $\mu\text{m}$  isotropic)  $\times$  11 integration points for a total compressed size of a little over 2 GB.  
324 This choice represented weighing the trade-off between tractability, portability, and accuracy.  
325 However, all data and code to reproduce the results described (with possible variation in the  
326 input parameters) are available in the dedicated GitHub repository.  
327 The normalized time point scalar value for each atlas/point-set in the temporal domains [0, 1]  
328 was also defined. Given the increasingly larger gaps in the postnatal timepoint sampling, we  
329 made two adjustments. Based on known mouse brain development, we used 28 days for the

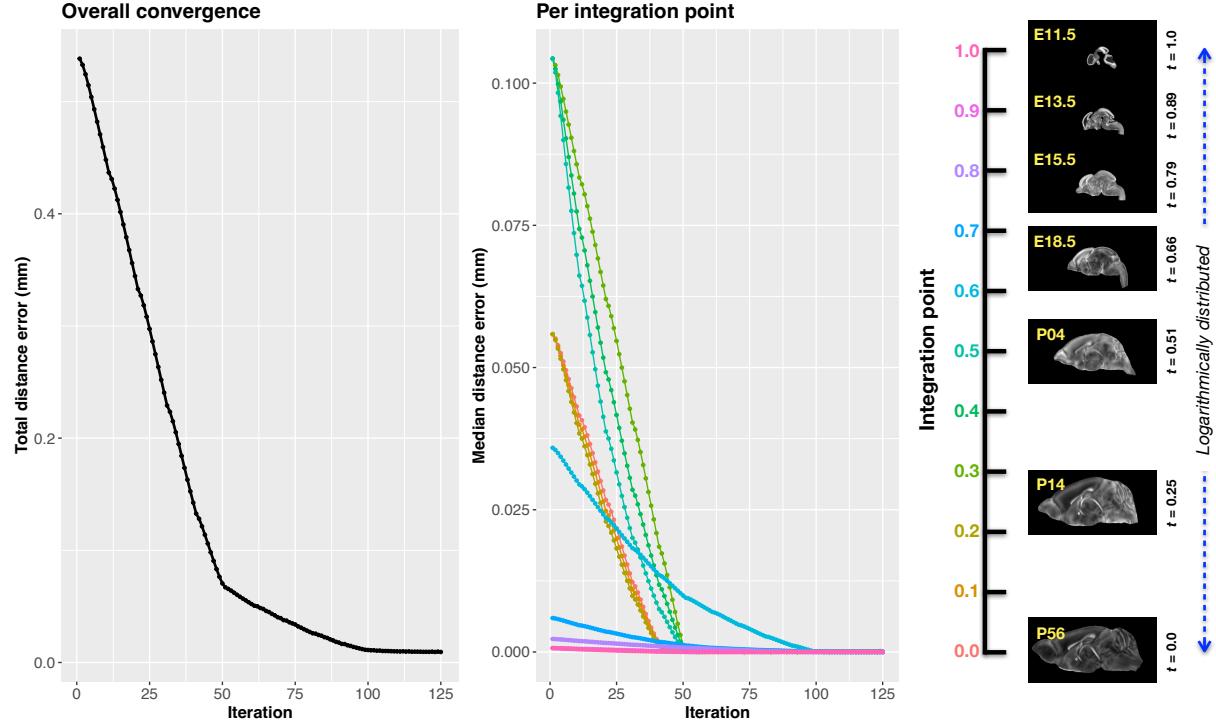


Figure 4: Convergence of the optimization of the velocity field for describing the transformation through the developmental stages from E11.5 through P56. Integration points in diagram on the right are color-coordinated with the center plot and placed in relation to the logarithmically situated temporal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

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

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

343 median displacement error at each of the integration points also trends towards zero but at  
 344 different rates.

345 **2.2.3 The transformation model**

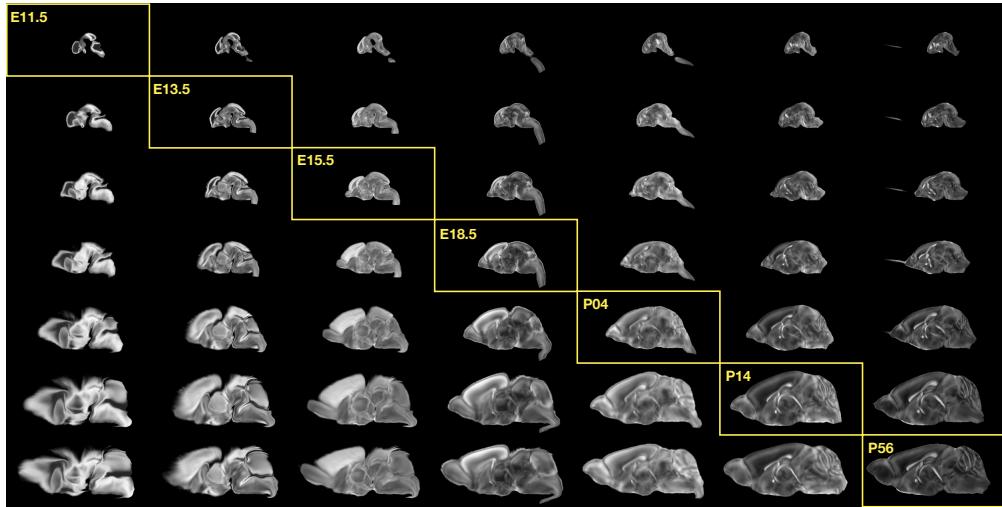


Figure 5: Mid-sagittal visualization of the effects of the transformation model in warping every developmental stage to the time point of every other developmental stage. The original images are located along the diagonal. Columns correspond to the warped original image whereas the rows represent the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

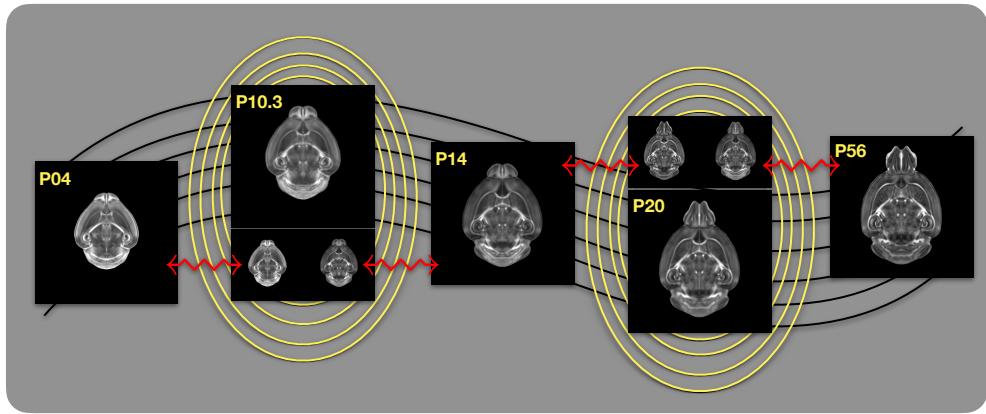


Figure 6: Illustration of the use of the velocity flow model for creating virtual templates at continuous time points not represented in one of the existing DevCCF time points. For example, FA templates at time point P10.3 and P20 can be generated by warping the existing temporally adjacent developmental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

346 Once optimized, the resulting velocity field can be used to generate the deformable transform

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

### 357 2.3 The Mouse Brain Parcellation Pipeline

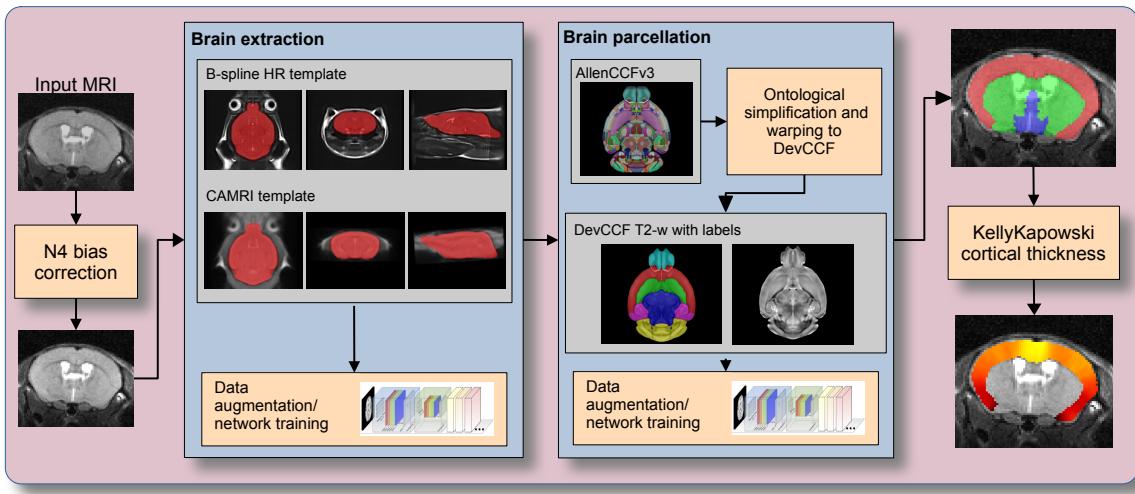


Figure 7: The mouse brain cortical parcellation pipeline integrating two deep learning components for brain extraction and brain parcellation prior to estimating cortical labels. Both deep learning networks rely heavily on data augmentation on templates built from open data and provide an outline for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

358 One of the most well-utilized pipelines in the ANTsX toolkit is the generation of corti-  
 359 cal thickness maps in the human brain from T1-weighted MRI. Starting with the novel  
 360 Diffeomorphic Registration-based Cortical Thickness (DiReCT) algorithm<sup>80</sup>, a complete al-  
 361 gorithmic workflow was developed for both cross-sectional<sup>81</sup> and longitudinal<sup>82</sup> T1-weighted

<sup>362</sup> MR image data. This contribution was later refactored using deep learning<sup>46</sup> leveraging the  
<sup>363</sup> earlier results<sup>81</sup> for training data.

<sup>364</sup> In the case of the mouse brain, the lack of training data and/or tools to generate training  
<sup>365</sup> data making analogous algorithmic development difficult. In addition, mouse data is often  
<sup>366</sup> characterized by unique issues such as frequent anisotropic sampling which are often in sharp  
<sup>367</sup> contrast to the high resolution resources available within the community, e.g., AllenCCFv3  
<sup>368</sup> and DevCCF. Using ANTsX and other publicly available data resources, we developed a  
<sup>369</sup> complete mouse brain structural morphology pipeline as illustrated in Figure 7 and detailed  
<sup>370</sup> below.

### <sup>371</sup> 2.3.1 Two-shot mouse brain extraction network

<sup>372</sup> In order to create a generalized mouse brain extraction network, we built whole-head tem-  
<sup>373</sup> plates from two publicly available datasets. The Center for Animal MRI (CAMRI) dataset<sup>68</sup>  
<sup>374</sup> from the University of North Carolina at Chapel Hill consists of 16 T2-weighted MRI volumes  
<sup>375</sup> of voxel resolution  $0.16 \times 0.16 \times 0.16 mm^3$ . The second high-resolution dataset<sup>69</sup> comprises  
<sup>376</sup> 88 specimens each with three spatially aligned canonical views with in-plane resolution of  
<sup>377</sup>  $0.08 \times 0.08 mm^2$  with a slice thickness of  $0.5 mm$ . These three orthogonal views were used to  
<sup>378</sup> reconstruct a single high-resolution volume per subject using a B-spline fitting algorithm de-  
<sup>379</sup> veloped in ANTsX<sup>83</sup>. From these two datasets, two symmetric isotropic ANTsX templates<sup>58</sup>  
<sup>380</sup> were generated analogous to the publicly available ANTsX human brain templates used in  
<sup>381</sup> previous research<sup>81</sup>. Bias field simulation, intensity histogram warping, noise simulation,  
<sup>382</sup> random translation and warping, and random anisotropic resampling in the three canon-  
<sup>383</sup> ical directions were used for data augmentation in training a T2-weighted brain extraction  
<sup>384</sup> network.

### <sup>385</sup> 2.3.2 Single-shot mouse brain parcellation network

<sup>386</sup> To create the network for generating a brain parcellation consistent with cortical thickness  
<sup>387</sup> estimation, we used the AllenCCFv3 and the associated `allensdk` Python library. Using

388 allensdk, a gross parcellation labeling was generated from the fine Allen CCFv3 labeling  
 389 which includes the cerebral cortex, cerebral nuclei, brain stem, cerebellum, main olfactory  
 390 bulb, and hippocampal formation. This labeling was mapped to the P56 component of  
 391 the DevCCF. Both the T2-w P56 DevCCF and labelings, in conjunction with the data  
 392 augmentation described previously for brain extraction, was used to train a brain parcellation  
 393 network.

### 394 2.3.3 Evaluation

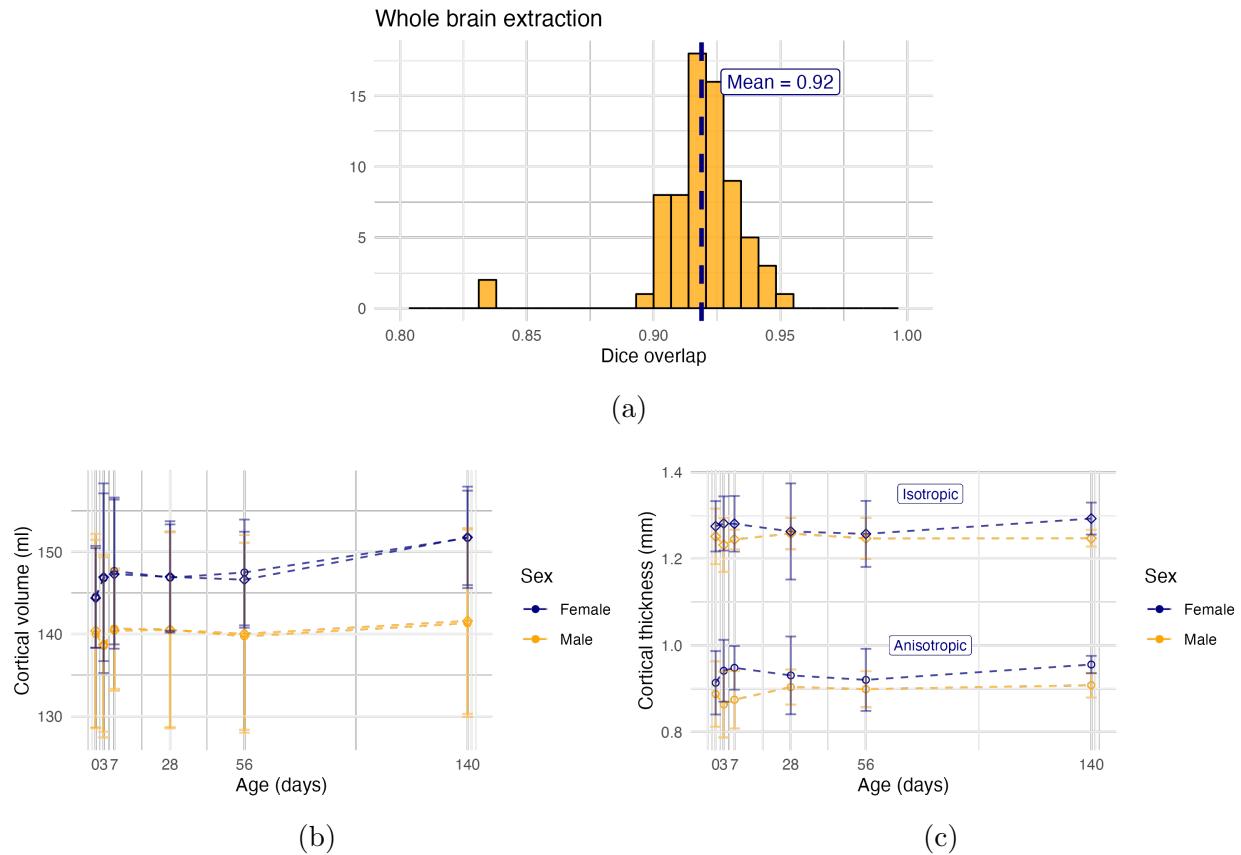


Figure 8: Evaluation of the ANTsX mouse brain extraction, parcellation, and cortical thickness pipeline on an independent dataset consisting of 12 specimens  $\times$  7 time points = 84 total images. (a) Dice overlap comparisons with the provided brain masks provide generally good agreement with the brain extraction network. (b) Cortical volume measurements show similar average quantities over growth and development between the original anisotropic data and interpolated isotropic data. (c) These results contrast with the cortical thickness measurements which show that cortical thickness estimation in anisotropic space severely underestimates the actual values.

395 For evaluation, we used an additional publicly available dataset<sup>70</sup> which is completely in-  
396 dependent from the data used in training the brain extraction and parcellation networks.  
397 Data includes 12 specimens each imaged at seven time points (Day 0, Day 3, Week 1, Week  
398 4, Week 8, Week 20) with available brain masks. In-plane resolution is  $0.1 \times 0.1 mm^2$  with  
399 a slice thickness of  $0.5 mm$ . Since the training data is isotropic and data augmentation in-  
400 cludes downsampling in the canonical directions, each of the two networks learns mouse  
401 brain-specific interpolation such that one can perform prediction on thick-sliced images, as,  
402 for example, in these evaluation data, and return isotropic probability and thickness maps (a  
403 choice available to the user). Figure 8 summarizes the results of the evaluation and compar-  
404 ison between isotropic and anisotropic cortical measurements in male and female specimens.

405 **3 Discussion**

406 The diverse mouse brain cell type profiles gathered through BICCN and associated efforts  
407 provides a rich multi-modal resource to the community. However, despite significant progress,  
408 full integration of these valuable resources is not yet complete. Central to the data integra-  
409 tion is a continued need to accurately map each unique dataset into common coordinate  
410 frameworks (CCFs) so that they can be accessed in connection with each other. Addition-  
411 ally, the ability to map novel cell type data in the future to these existing BICCN resources  
412 is vital for effective utilization of this endeavor and the continuation of its goals. To meet  
413 these needs, tools for mapping mouse cell type data must be both generally accessible to  
414 a wide audience of investigators, and still capable of handling distinct challenges unique to  
415 each data type.

416 In this work, we describe modular ANTsX-based pipelines developed to address the needs  
417 of three BICCN projects that cover distinct cell type data, including spatial transcriptomic,  
418 morphology, and developmental data. We highlight how a modular toolbox like ANTsX can  
419 be tailored to address problems unique to each modality while still leveraging a variety of  
420 ready-to-use powerful tools that have been externally validated.

421 Our MERFISH pipeline provides an example of how to map high-resolution spatial tran-  
422 scriptomic data into the AllenCCFv3. Since full brain large-scale transcriptomics is still  
423 rare and difficult to collect, the pipeline focuses on achieving the best possible anatomical  
424 alignment and fully utilizing the available data. While the techniques employed for mapping  
425 the sectioned data can be generally applicable to map other serial histology images, many  
426 parts of the pipeline were designed to address very specific known alignment challenges in  
427 the MERFISH data using a series of iterative registration steps. The pipeline shows how  
428 general tools available in ANTsX can be adapted to target highly specialized problems in  
429 mouse cell type data.

430 In contrast to the MERFISH pipeline, our fMOST pipeline was designed to be a more  
431 general solution that can be employed in other modalities. The pipeline primarily uses  
432 previously developed ANTsX preprocessing and atlasing tools to map fMOST data into the

433 AllenCCFv3. The key component of the pipeline is the use of a fMOST specific average  
434 atlas to greatly simplify the image registration problem. This average atlas, also constructed  
435 using pre-existing ANTsX tools, allows for a one-time canonical alignment from the fMOST  
436 atlas to the AllenCCFv3 to be transferred and used for mapping new fMOST images. Lastly,  
437 ANTsX provides point set transformation tools to allow the mappings found through the  
438 pipeline to be directly applied to associated single-cell reconstructions from the fMOST data  
439 to study neuronal morphology.

440 Our DevCCF pipeline shows the application of the toolkit for temporospatial developmental  
441 data. ANTsX was crucial in providing necessary functionality for yielding high quality  
442 output. For the generation of the individual developmental stage multi-modal, symmetric  
443 templates, ANTsX is unique amongst image analysis software packages in providing existing  
444 solutions for template generation which have been thoroughly vetted, including being used  
445 in several studies over the years, and which continue to be under active refinement. At its  
446 core, computationally efficient and quality template generation requires the use of precision  
447 pairwise image mapping functionality which, historically, is at the origins of the ANTsX  
448 ecosystem. Moreover, these mapping capabilities extend beyond template generation to the  
449 mapping of other image data (e.g., gene expression maps) to a selected template for providing  
450 further insight into the mouse brain.

451 With respect to the DevCCF, despite the significant expansion of available developmental age  
452 templates beyond what existed previously, there are still temporal gaps in the DevCCF which  
453 can be potentially sampled by future research efforts. However, pioneering work involving  
454 time-varying diffeomorphic transformations allow us to continuously situate the existing  
455 templates within a velocity flow model. This allows one to determine the diffeomorphic  
456 transformation from any one temporal location to any other temporal location within the  
457 time span defined by the temporal limits of the DevCCF. This functionality is built on  
458 multiple ITK components including the B-spline scattered data approximation technique for  
459 field regularization and velocity field integration. This velocity field model permits intra-  
460 template comparison and the construction of virtual templates where a template can be  
461 estimated at any continuous time point within the temporal domain. This novel application

462 can potentially enhance our understanding of intermediate developmental stages.

463 We also presented a mouse brain pipeline for brain extraction, parcellation, and cortical  
464 thickness using single-shot and two-shot learning with data augmentation. This approach  
465 attempts to circumvent (or at least minimize) the typical requirement of large training  
466 datasets as with the human ANTsX pipeline analog. However, even given our initial success  
467 on independent data, we fully anticipate that refinements will be necessary. Given that the  
468 ANTsX toolkit is a dynamic effort undergoing continual improvement, we manually correct  
469 cases that fail and use them for future training and refinement of network weights as we have  
470 done for our human-based networks. Generally, these approaches provide a way to bootstrap  
471 training data for manual refinement and future generation of more accurate deep learning  
472 networks in the absence of other applicable tools.

473 The ANTsX ecosystem is a powerful framework that has demonstrated applicability to di-  
474 verse cell type data in the mouse brain. This is further evidenced by the many software  
475 packages that use various ANTsX components in their own mouse-specific workflows. In  
476 and of itself, the extensive functionality of ANTsX makes it possible to create complete pro-  
477 cessing pipelines without requiring the integration of multiple packages or lengthy software  
478 development. These open-source components not only perform well but are available across  
479 multiple platforms which facilitates the construction of tailored pipelines for individual study  
480 solutions. These components are also supported by years of development not only by the  
481 ANTsX development team but by the larger ITK community.

482 **4 Methods**

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

488 **4.1 General ANTsX utilities**

489 Although they focus on distinct data types, the three pipelines presented share common  
490 components that are generally applicable when mapping mouse cell type data. These include,  
491 addressing intensity biases and noise in the data, image registration to solve the mapping,  
492 creating custom templates and atlases from the data, and visualization of the results. Table  
493 1 provides a brief summary of key general functionalities in ANTsX for addressing these  
494 challenges.

495 **4.1.1 Preprocessing: bias field correction and denoising**

496 Bias field correction and image denoising are standard preprocessing steps in improving over-  
497 all image quality in mouse brain images. The bias field, a gradual spatial intensity variation  
498 in images, can arise from various sources such as magnetic field inhomogeneity or acquisition  
499 artifacts, leading to distortions that can compromise the quality of brain images. Correct-  
500 ing for bias fields ensures a more uniform and consistent representation of brain structures,  
501 enabling more accurate quantitative analysis. Additionally, brain images are often suscep-  
502 tible to various forms of noise, which can obscure subtle features and affect the precision  
503 of measurements. Denoising techniques help mitigate the impact of noise, enhancing the  
504 signal-to-noise ratio and improving the overall image quality. The well-known N4 bias field  
505 correction algorithm<sup>61</sup> has its origins in the ANTs toolkit which was implemented and intro-  
506 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> <code>mouse_brain_extraction(...modality="ex5"...)</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>).

507 contains an implementation of a well-performing patch-based denoising technique<sup>60</sup> and is  
508 also available as an image filter to the ITK community, `ants.denoise_image(...)`.

509 **4.1.2 Image registration**

510 The ANTs registration toolkit is a complex framework permitting highly tailored solutions  
511 to pairwise image registration scenarios<sup>84</sup>. It includes innovative transformation models  
512 for biological modeling<sup>54,67</sup> and has proven capable of excellent performance<sup>55,85</sup>. Vari-  
513 ous parameter sets targeting specific applications have been packaged with the different  
514 ANTsX packages, specifically ANTs, ANTsPy, and ANTsR<sup>46</sup>. In ANTsPy, the function  
515 `ants.registration(...)` is used to register a pair of images or a pair of image sets where  
516 `type_of_transform` is a user-specified option that invokes a specific parameter set. For ex-  
517 ample `type_of_transform='antsRegistrationSyNQuick[s]'` encapsulates an oft-used pa-  
518 rameter set for quick registration whereas `type_of_transform='antsRegistrationSyN[s]'`  
519 is a more detailed alternative. Transforming images using the derived transforms is performed  
520 via the `ants.apply_transforms(...)` function.

521 Initially, linear optimization is initialized with center of (intensity) mass alignment typically  
522 followed by optimization of both rigid and affine transforms using the mutual information  
523 similarity metric. This is followed by diffeomorphic deformable alignment using symmetric  
524 normalization (SyN) with Gaussian<sup>54</sup> or B-spline regularization<sup>67</sup> where the forward trans-  
525 form is invertible and differentiable. The similarity metric employed at this latter stage is  
526 typically either neighborhood cross-correlation or mutual information. Note that these pa-  
527 rameter sets are robust to input image type (e.g., light sheet fluorescence microscopy, Nissl  
528 staining, and the various MRI modalities) and are adaptable to mouse image geometry and  
529 scaling. Further details can be found in the various documentation sources for these ANTsX  
530 packages.

531 **4.1.3 Template generation**

532 ANTsX provides functionality for constructing templates from a set (or multi-modal sets) of  
533 input images as originally described<sup>58</sup> and recently used to create the DevCCF templates<sup>16</sup>.

534 An initial template estimate is constructed from an existing subject image or a voxelwise  
535 average derived from a rigid pre-alignment of the image population. Pairwise registration  
536 between each subject and the current template estimate is performed using the Symmetric  
537 Normalization (SyN) algorithm<sup>54</sup>. The template estimate is updated by warping all subjects  
538 to the space of the template, performing a voxelwise average, and then performing a “shape  
539 update” of this latter image by warping it by the average inverse deformation, thus yielding  
540 a mean image of the population in terms of both intensity and shape. The corresponding  
541 ANTsPy function is `ants.build_template(...)`.

#### 542 4.1.4 Visualization

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

## 547 4.2 Mapping fMOST data to AllenCCFv3

### 548 4.2.1 Preprocessing

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

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

- 565 • *Stripe artifact removal.* Repetitive pattern artifacts are a common challenge in fMOST  
566 imaging where inhomogeneity during the cutting and imaging of different sections can  
567 leave stripes of hyper- and hypo-intensity across the image. These stripe artifacts  
568 can be latched onto by the registration algorithm as unintended features that are  
569 then misregistered to non-analogous structures in the AllenCCFv3. We address these  
570 artifacts by fitting a 3-D bandstop (notch) filter to target the frequency of the stripe  
571 patterns and removing them prior to the image registration.
- 572 • *Inhomogeneity correction.* Regional intensity inhomogeneity can also occur within  
573 and between sections in fMOST imaging due to staining or lighting irregularity during  
574 acquisition. Similar to stripe artifacts, intensity gradients due to inhomogeneity  
575 can be misconstrued as features during the mapping and result in matching of non-  
576 corresponding structures. Our pipeline addresses these intensity inhomogeneities using  
577 N4 bias field correction<sup>61</sup>, `ants.n4_bias_field_correction(...)`.

#### 578 4.2.2 Steps for spatial normalization to AllenCCFv3

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

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

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

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

- 617     4. *Alignment of individual fMOST mouse brains.* The canonical transformation between  
618       the fMOST atlas and AllenCCFv3 greatly simplifies the registration of new individ-  
619       ual fMOST mouse brains into the AllenCCFv3. Each new image is first registered  
620       into the fMOST average atlas, which shares intensity, modality, and morphologi-  
621       cal characteristics. This allows us to leverage standard, intensity-based registration  
622       functionality<sup>84</sup> available in ANTsX to perform this alignment. Transformations are  
623       then concatenated to the original fMOST image to move it into the AllenCCFv3 space  
624       using `ants.apply_transforms(...)`.
- 625     5. *Transformation of single cell neurons.* A key feature of fMOST imaging is the ability  
626       to reconstruct and examine whole-brain single neuron projections<sup>79</sup>. Spatial mapping  
627       of these neurons from individual brains into the AllenCCFv3 allows investigators to  
628       study different neuron types within the same space and characterize their morphology  
629       with respect to their transcriptomics. Mappings found between the fMOST image  
630       and the AllenCCFv3 using our pipeline can be applied in this way to fMOST neuron  
631       reconstruction point set data using `ants.apply_transforms_to_points(..)`.

632     **4.3 Mapping MERFISH data to AllenCCFv3**

633     **4.3.1 Preprocessing**

- 634     • *Initial volume reconstruction.* Alignment of MERFISH data into a 3-D atlas space  
635       requires an estimation of anatomical structure within the data. For each section,  
636       this anatomic reference image was created by aggregating the number of detected  
637       genetic markers (across all probes) within each pixel of a  $10 \times 10 \mu\text{m}^2$  grid to match  
638       the resolution of the  $10 \mu\text{m}$  AllenCCFv3 atlas. These reference image sections are then  
639       coarsely reoriented and aligned across sections using manual annotations of the most  
640       dorsal and ventral points of the midline. The procedure produces an anatomic image  
641       stack that serves as an initialization for further global mappings into the AllenCCFv3.
- 642     • *Anatomical correspondence labeling.* Mapping the MERFISH data into the AllenCCFv3  
643       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.

670 **4.3.2 2.5D deformable, landmark-driven alignment to AllenCCFv3**

671 After global alignment of the AllenCCFv3 into the MERFISH dataset, 2D per-section de-  
672 formable refinements are used to address local differences between the MERFISH sections  
673 and the resampled AllenCCFv3 sections. Nine registrations were performed in sequence us-  
674 ing a single label at each iteration in the following order: 1) brain mask, 2) isocortex (layer  
675 2+3), 3) isocortex (layer 5), 4) isocortex (layer 6), 5) striatum, 6) medial habenula, 7) lateral  
676 habenula, 8) thalamus, and 9) hippocampus. This ordering was determined empirically by  
677 an expert anatomist who prioritized which structure to use in each iteration by evaluat-  
678 ing the anatomical alignment from the previous iteration. Global and local mappings are  
679 then all concatenated (with appropriate inversions) to create the final mapping between the  
680 MERFISH data and AllenCCFv3. This mapping is then used to provide a point-to-point  
681 correspondence between the original MERFISH coordinate space and the AllenCCFv3 space,  
682 thus allowing mapping of individual genes and cell types located in the MERFISH data to  
683 be directly mapped into the AllenCCFv3.

684 **4.4 DevCCF velocity flow transformation model**

685 Given multiple, linearly or non-linearly ordered point sets where individual points across the  
686 sets are in one-to-one correspondence, we developed an approach for generating a velocity  
687 flow transformation model to describe a time-varying diffeomorphic mapping as a variant of  
688 the landmark matching solution. Integration of the resulting velocity field can then be used  
689 to describe the displacement between any two time points within this time-parameterized  
690 domain. Regularization of the sparse correspondence between point sets is performed using  
691 a generalized B-spline scattered data approximation technique<sup>83</sup>, also created by the ANTsX  
692 developers and contributed to ITK.

693 **4.4.1 Velocity field optimization**

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

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

## 708 4.5 ANTsXNet mouse brain applications

### 709 4.5.1 General notes regarding deep learning training

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

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

723 ing for mimicking well-known MRI intensity nonlinearities<sup>46,86</sup>. These augmentation tech-  
724 niques are available in ANTsXNet (only ANTsPyNet versions are listed with ANTsRNet  
725 versions available) and include:

- 726 • image noise: `ants.add_noise_to_image(...)`,
- 727 • simulated bias field: `antspynet.simulate_bias_field(...)`, and
- 728 • nonlinear intensity warping: `antspynet.histogram_warp_image_intensities(...)`.

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

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

#### 734 4.5.2 Brain extraction

735 Similar to human neuroimage processing, brain extraction is a crucial preprocessing step  
736 for accurate brain mapping. We developed similar functionality for T2-weighted mouse  
737 brains. This network uses a conventional U-net architecture<sup>87</sup> and, in ANTsPyNet, this  
738 functionality is available in the program `antspynet.mouse_brain_extraction(...)`.  
739 For the two-shot T2-weighted brain extraction network, two brain templates were gen-  
740 erated along with their masks. One of the templates was generated from orthogonal  
741 multi-plane, high resolution data<sup>69</sup> which were combined to synthesize isotropic volu-  
742 metric data using the B-spline fitting algorithm<sup>83</sup>. This algorithm is encapsulated in  
743 `ants.fit_bspline_object_to_scattered_data(...)` where the input is the set of voxel  
744 intensity values and each associated physical location. Since each point can be assigned  
745 a confidence weight, we use the normalized gradient value to more heavily weight edge  
746 regions. Although both template/mask pairs are available in the GitHub repository  
747 associated with this work, the synthesized volumetric B-spline T2-weighted pair is available  
748 within ANTsXNet through the calls:

- 749     • template: `antspynet.get_antsxnet_data("bsplineT2MouseTemplate")` and  
750     • mask: `antspynet.get_antsxnet_data("bsplineT2MouseTemplateBrainMask")`.

751   **4.5.3 Brain parcellation**

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

760 **Data availability**

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

<sup>770</sup> **Acknowledgments**

<sup>771</sup> Support for the research reported in this work includes funding from the National Institute  
<sup>772</sup> of Biomedical Imaging and Bioengineering (R01-EB031722) and National Institute of Mental  
<sup>773</sup> Health (RF1-MH124605 and U24-MH114827).

## <sup>774</sup> Author contributions

<sup>775</sup> N.T., M.C., and J.G. wrote the main manuscript text and figures. M.C., M.K., R.D., S.S.,  
<sup>776</sup> Q.W., L.G., J.D., C.G., and J.G. developed the Allen registration pipelines. N.T. and F.K.  
<sup>777</sup> developed the time-varying velocity transformation model for the DevCCF. N.T. and M.T.  
<sup>778</sup> developed the brain parcellation and cortical thickness methodology. All authors reviewed  
<sup>779</sup> the manuscript.

780 **References**

- 781 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).
- 782 2. La Manno, G. *et al.* Molecular architecture of the developing mouse brain. *Nature*  
**596**, 92–96 (2021).
- 783 3. Wen, L. *et al.* Single-cell technologies: From research to application. *Innovation  
(Camb)* **3**, 100342 (2022).
- 784 4. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–14  
(2014).
- 785 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).
- 786 6. Li, A. *et al.* Micro-optical sectioning tomography to obtain a high-resolution atlas of  
the mouse brain. *Science* **330**, 1404–8 (2010).
- 787 7. Ueda, H. R. *et al.* Tissue clearing and its applications in neuroscience. *Nat Rev  
Neurosci* **21**, 61–79 (2020).
- 788 8. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by  
spatial transcriptomics. *Science* **353**, 78–82 (2016).
- 789 9. Burgess, D. J. Spatial transcriptomics coming of age. *Nat Rev Genet* **20**, 317 (2019).
- 790 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).
- 791 11. Hawrylycz, M. *et al.* A guide to the BRAIN initiative cell census network data  
ecosystem. *PLoS biology* **21**, e3002133 (2023).
- 792 12. Wang, Q. *et al.* The allen mouse brain common coordinate framework: A 3D reference  
atlas. *Cell* **181**, 936–953.e20 (2020).
- 793 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).
- 794 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).

- 795 15. Qu, L. *et al.* Cross-modal coherent registration of whole mouse brains. *Nature Methods* **19**, 111–118 (2022).
- 796 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).
- 797 17. Chuang, N. *et al.* An MRI-based atlas and database of the developing mouse brain. *Neuroimage* **54**, 80–89 (2011).
- 798 18. Dries, R. *et al.* Advances in spatial transcriptomic data analysis. *Genome research* **31**, 1706–1718 (2021).
- 799 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).
- 800 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).
- 801 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).
- 802 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).
- 803 23. Cahill, L. S. *et al.* Preparation of fixed mouse brains for MRI. *Neuroimage* **60**, 933–939 (2012).
- 804 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).
- 805 25. Kim, Y. *et al.* Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell* **171**, 456–469 (2017).
- 806 26. Fürth, D. *et al.* An interactive framework for whole-brain maps at cellular resolution. *Nat Neurosci* **21**, 139–149 (2018).

- 807 27. Li, Y. *et al.* mBrainAligner-web: A web server for cross-modal coherent registration  
of whole mouse brains. *Bioinformatics* **38**, 4654–4655 (2022).
- 808 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).
- 809 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).
- 810 30. Ni, H. *et al.* A robust image registration interface for large volume brain atlas. *Sci  
Rep* **10**, 2139 (2020).
- 811 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).
- 812 32. Celestine, M., Nadkarni, N. A., Garin, C. M., Bougacha, S. & Dhenain, M. **Sammba-  
MRI: A library for processing SmAll-MaMmal BrAin MRI data in python.** *Front  
Neuroinform* **14**, 24 (2020).
- 813 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).
- 814 34. Perens, J. *et al.* Multimodal 3D mouse brain atlas framework with the skull-derived  
coordinate system. *Neuroinformatics* **21**, 269–286 (2023).
- 815 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).
- 816 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).
- 817 37. Chandrashekhar, V. *et al.* CloudReg: Automatic terabyte-scale cross-modal brain  
volume registration. *Nature methods* **18**, 845–846 (2021).
- 818 38. Jin, M. *et al.* **SMART: An open-source extension of WholeBrain for intact mouse  
brain registration and segmentation.** *eNeuro* **9**, (2022).

- 819 39. Negwer, M. *et al.* FriendlyClearMap: An optimized toolkit for mouse brain mapping  
and analysis. *Gigascience* **12**, (2022).
- 820 40. Lin, W. *et al.* Whole-brain mapping of histaminergic projections in mouse brain.  
*Proceedings of the National Academy of Sciences* **120**, e2216231120 (2023).
- 821 41. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by  
MERFISH. *Nature* **598**, 137–143 (2021).
- 822 42. Shi, H. *et al.* Spatial atlas of the mouse central nervous system at molecular resolution.  
*Nature* **622**, 552–561 (2023).
- 823 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).
- 824 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).
- 825 45. Fedorov, A. *et al.* 3D slicer as an image computing platform for the quantitative  
imaging network. *Magnetic resonance imaging* **30**, 1323–1341 (2012).
- 826 46. Tustison, N. J. *et al.* The ANTsX ecosystem for quantitative biological and medical  
imaging. *Sci Rep* **11**, 9068 (2021).
- 827 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).
- 828 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).
- 829 49. Allan Johnson, G. *et al.* Whole mouse brain connectomics. *Journal of Comparative  
Neurology* **527**, 2146–2157 (2019).
- 830 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).

- 831 51. Bajcsy, R. & Broit, C. Matching of deformed images. in *Sixth International Conference on Pattern Recognition (ICPR'82)* 351–353 (1982).
- 832 52. Bajcsy, R. & Kovacic, S. [Multiresolution elastic matching](#). *Computer Vision, Graphics, and Image Processing* **46**, 1–21 (1989).
- 833 53. Gee, J. C., Reivich, M. & Bajcsy, R. [Elastically deforming 3D atlas to match anatomical brain images](#). *J Comput Assist Tomogr* **17**, 225–36 (1993).
- 834 54. 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).
- 835 55. Klein, A. *et al.* [Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration](#). *Neuroimage* **46**, 786–802 (2009).
- 836 56. Murphy, K. *et al.* [Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge](#). *IEEE Trans Med Imaging* **30**, 1901–20 (2011).
- 837 57. 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).
- 838 58. Avants, B. B. *et al.* [The optimal template effect in hippocampus studies of diseased populations](#). *Neuroimage* **49**, 2457–66 (2010).
- 839 59. 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).
- 840 60. 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).
- 841 61. Tustison, N. J. *et al.* [N4ITK: Improved N3 bias correction](#). *IEEE Trans Med Imaging* **29**, 1310–20 (2010).
- 842 62. Wang, H. *et al.* [Multi-atlas segmentation with joint label fusion](#). *IEEE Trans Pattern Anal Mach Intell* **35**, 611–23 (2013).

- 843 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).
- 844 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).
- 845 65. McCormick, M., Liu, X., Jomier, J., Marion, C. & Ibanez, L. **ITK: Enabling reproducible research and open science**. *Front Neuroinform* **8**, 13 (2014).
- 846 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).
- 847 67. Tustison, N. J. & Avants, B. B. **Explicit B-spline regularization in diffeomorphic image registration**. *Front Neuroinform* **7**, 39 (2013).
- 848 68. Hsu, L.-M. *et al.* CAMRI mouse brain MRI data.
- 849 69. Reshetnikov, V. *et al.* High-resolution MRI data of brain C57BL/6 and BTBR mice in three different anatomical views.
- 850 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).
- 851 71. Liu, J. *et al.* **Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing**. *Life Sci Alliance* **6**, (2023).
- 852 72. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. **Cellpose: A generalist algorithm for cellular segmentation**. *Nat Methods* **18**, 100–106 (2021).
- 853 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).
- 854 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).

- 855 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).
- 856 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).
- 857 77. Wang, J. *et al.* Divergent projection patterns revealed by reconstruction of individual neurons in orbitofrontal cortex. *Neurosci Bull* **37**, 461–477 (2021).
- 858 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).
- 859 79. Peng, H. *et al.* Morphological diversity of single neurons in molecularly defined cell types. *Nature* **598**, 174–181 (2021).
- 860 80. Das, S. R., Avants, B. B., Grossman, M. & Gee, J. C. Registration based cortical thickness measurement. *Neuroimage* **45**, 867–79 (2009).
- 861 81. Tustison, N. J. *et al.* Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *Neuroimage* **99**, 166–79 (2014).
- 862 82. Tustison, N. J. *et al.* Longitudinal mapping of cortical thickness measurements: An Alzheimer’s Disease Neuroimaging Initiative-based evaluation study. *J Alzheimers Dis* (2019) doi:10.3233/JAD-190283.
- 863 83. 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).
- 864 84. Avants, B. B. *et al.* The Insight ToolKit image registration framework. *Front Neuroinform* **8**, 44 (2014).
- 865 85. Avants, B. B. *et al.* A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **54**, 2033–44 (2011).
- 866 86. 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).

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