# D-LMBmapX: Generalized Deep-Learning Pipeline for Automated Whole-brain Neural Circuitry Profiling Across Any Developmental Stages

Zhongyu Li[1,2,4†*], Peiqi Li[2†], Hangzhen Pan[2†], Annabel M.J. Adams[1†], Zuqi Huang[2], He Liu[2], Jinjiang Li[2], Yaoyu Su[2], Baoming Zhong[2], Liam Bray[1], Robyn N. Sturgess[1], Jihua Zhu[2], Dingwen Zhang[3], Shaoting Zhang[4,5], Jing Ren[1*]

1. Neurobiology Division, MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, CB2 0QH, United Kingdom

2. School of Software Engineering, Xi'an Jiaotong University, Xi'an, China

3. Hefei Comprehensive National Science Center, Hefei, China

4. Shanghai Jiao Tong University, Shanghai, China

5. Shanghai Artificial Intelligence Laboratory, Shanghai, China

†These authors contributed equally to this work

*Correspondence: Jing Ren (jren@mrc-lmb.cam.ac.uk), Zhongyu Li (zhongyuli@xjtu.edu.cn)

## SUMMARY

Understanding whole-brain circuitry development in mice has been hindered by the lack of tools for accurate and comprehensive analysis. No existing 3D brain atlas provides daily postnatal resolution, as constructing such atlases is highly labour-intensive. Axonal morphologies change dynamically, making reliable segmentation challenging, and many 2D datasets lack sufficient Z-resolution for cross-modality 3D analysis. Here, we present D-LMBmapX, a deep-learning pipeline for automated whole-brain circuitry profiling across postnatal development. D-LMBmapX constructs high-resolution 3D mouse brain atlases spanning seven postnatal stages and employs an adaptive registration strategy for whole-brain alignment at any postnatal day. It also integrates a foundation model for axon and soma segmentation, enabling the quantitative circuitry assessment across development. We implemented a diffusion model-based style transfer for cross-modality and cross-dimensional registration, validated by aligning genetically defined neuronal types from 2D ISH datasets to our 3D atlas. Using D-LMBmapX, we profiled whole-brain dopaminergic projections throughout postnatal maturation.

## INTRODUCTION

In-depth knowledge of neuronal connections in the adult brain is fundamental to understanding how the brain is functionally organized. Moreover, knowledge of neuronal connectivity in the developing brain not only offers insights into how adult brain networks are established but also provides crucial anatomical

information for understanding the pathology of neurodevelopmental and mental disorders that arise from deviant brain development[1]. Recent whole-brain connectomic analysis of adult and larval Drosophila brains revealed that around 0.5% of neurons have developmental wiring defects, leading to the hypothesis that such miswiring might contribute to individuality and susceptibility to mental health disorders[2].

In the past decades, great success has been achieved in developing powerful new techniques and platforms for mapping mammalian neural circuits, particularly in mesoscale connectomics[3,4]. Single-cell level, cell-type-specific connections across various mouse brain regions have been mapped using a combination of viral-genetic labelling techniques and advanced imaging methods, such as serial two-photon tomography (STPT)[5,6] and fluorescence micro-optical sectioning tomography (fMOST)[7,8]. Meanwhile, the integration of tissue-clearing techniques with light-sheet fluorescence microscopy (LSFM) has opened up remarkable opportunities for high-throughput, 3D mesoscale connectivity mapping of the entire brain[9–11]. However, resources and analytical tools for mapping connectivity in developing brains remain limited. For example, D-LMBmap employs deep-learning algorithms for whole-brain circuitry profiling, addressing automated axon segmentation, brain region segmentation and whole-brain 3D registration in a single workflow[4]. However, it is not equipped to process data collected from brains at various developmental stages due to the critical challenges listed below. First, resources for mouse brain atlases across developmental stages are limited. The Allen Developing Mouse Brain Atlas (ADMBA) was created using a limited number of 2D histological sections, and the annotation is not sufficiently detailed[12,13]. Recently, the 3D Developmental Common Coordinate Framework (DevCCF) for mouse brains was released, offering an atlas with a 10 μm isotropic resolution[14]. However, for postnatal developmental stages, both the ADMBA and DevCCF only include two time points: P4 and P14. Both platforms are restricted in scope because it is extremely labour-intensive to create 3D reference atlases for every postnatal day. Second, axonal fibers exhibit diverse morphologies across different developmental stages, often differing significantly from those in adult brains. To our knowledge, there are currently no axon segmentation tools available that can effectively profile projectomes throughout various stages of development. Third, no existing platform can integrate gene expression or tracing data from the developing brains across multiple modalities while establishing whole-brain registration within a standardized framework.

Here, we developed D-LMBmapX, a deep-learning pipeline designed for automatic whole-brain circuitry profiling at any postnatal stage. We validated

our methods by profiling whole-brain dopaminergic projections throughout postnatal development. We also achieved cross-modality registration and validated this by registering genetically defined neuronal types from the Allen Institute's developmental ISH datasets[15] to our framework.

## RESULTS

### A Comprehensive Pipeline for Quantifying Developmental Neuronal Circuits

D-LMBmapX provides comprehensive tools and newly generated high-resolution 3D mouse brain atlas for profiling of whole-brain neuronal circuits at any postnatal developmental stage. We first used a novel cross-stage registration method to construct 10 μm voxel resolution average template brains across seven postnatal stages, i.e., P0, P4, P7, P10, P14, P21, and P28, where each stage includes 30 to 40 brains collected by light-sheet fluorescence microscopy (LSFM) imaging after tissue clearing. Subsequently, hundreds of hierarchical brain structures are successively delineated based on registration from Allen Mouse Brain Common Coordinate Framework (CCFv3)[16] to P28 average template and so forth (Figure 1A). Here, we also considered the case where users' interesting stages are not included in the above seven atlases. Given an experimental brain imaged in any developmental stage, the brain structures can be adaptatively quantified using on a designed cross-stage registration strategy based on two atlases from the flanking stages (Figure 1B). We further established a generalized model for axon and soma segmentation, enabling the quantification of axon densities throughout development. We developed a neuron foundation model using advanced deep learning techniques, trained on more than 100,000 3D cubes, to achieve automated segmentation of whole-brain axon and soma across different developmental stages. By inputting sample brains from any postnatal stage, D-LMBmapX outputs the quantification and visualization of region-specific axon and soma distributions at the 3D whole-brain level, registered to the corresponding developmental stage (Figure 1C). In addition to the light-sheet modality, we extended the whole-brain analysis to various imaging modalities (e.g., STPT, MRI, confocal imaging) through a newly designed diffusion model for brain style transfer. Leveraging this model, D-LMBmapX also integrates data analysis from 2D histological slides, employing a similarity localization pipeline to map and quantitatively project 2D slides onto a 3D brain atlas (Figure 1D). Finally, D-LMBmapX is a user-friendly software that integrates multiple models and offers functions such as developmental atlas visualization, cross-stage brain registration, whole-brain axon and soma segmentation and quantification,

cross-modality brain style transfer, and 2D-to-3D whole-brain projection (Figure 1E).

**3D mouse brain atlas for any postnatal stage analysis**

Considering the lack of high-resolution 3D brain atlases in postnatal stages, we collected light-sheet imaged brains and constructed brain atlas in 7 stages from P0 to P28. Firstly, average brain template is computed based on sequential registration to the initial reference brain template. To alleviate the brain image diversity during sample collection and microscopy imaging, we employ a series of image pre-processing strategies, including N4 bias field correction[17], rescaling intensity, and 3D adaptive histogram equalization[18]. In each stage, an initial reference brain template is constructed based on the voxel average of all brains. Then the reference brain template is iteratively registered and updated based on the average deformation space of all processed brains (Figure 2A). Different from commonly used brain atlas construction methods[14,16] only using rigid, affine, and B-spline registration, we develop an extra 2.5D non-rigid registration method to further refine the reference brain template (Figure S1A), which can well address brain diversity in development. Moreover, unlike most widely used 3D deep learning-based non-rigid methods, the proposed 2.5D method can achieve high-resolution 3D registration without huge computational expense (Figure S1B). Finally, high-resolution average brain template (in 10 μm) is constructed in 7 stages from P0 to P28. To delineate hundreds of brain structures in each brain template, we carry out sequential registration from the adult brain template Allen CCFv3 to P28 template in mapping and constructing P28 atlas (Figure 2B). Then, the delineated P28 template is registered to P21 template in mapping and constructing P21 atlas, and so forth (Figure S1C). To evaluate the effectiveness of the constructed 7 postnatal atlas, we first compare the atlas with recently developed KimDevCCF atlas[14] in P4 and P14 stages (Figure 2C) by computing the Dice score of major brain regions from experimental brains in registering the two atlases respectively. For other stages KimDevCCF not included, we also record the Dice score of major brain regions from experimental brains in registering our constructed atlases (Figure S1D).

Despite the constructed atlases including 7 postnatal stages, they cannot cover every postnatal day for diverse whole-brain projection analysis. Here, given an experiment brain at any postnatal day, we develop a cross-stage strategy for adaptative brain structure quantification. The experiment brain registers to its two flanking stages respectively to compute two deformation fields afterwards (Figure 2D, Figure S1E). We quantify region-wise deformation fields and Dice

score in reflecting the similarity of experimental brain with its two flanking stages in structure level (Figure 2E, Figure S1F). Users can decide which flanking stage to register based on the deformation score of interested brain regions. Additionally, to explore the dynamic brain structure changes during development, we sequentially register each brain atlas to its flanking atlas in computing the deformation score in hundreds of brain regions (Figure 2F). The deformation changes per day are first recorded in five major brain regions (CB, CTX, CP, HPF, BS). We find that deformation changes in most brain regions are consistent. The deformation changes of inconsistent brain regions are also provided (Figure S1G).

**Generalized whole-brain signal quantification by implementing neuron foundation model**

Automated recognition of neuronal structure (e.g., axon and soma) at whole-brain level generally require specific deep learning models using abundant well-annotated training samples. Considering the labor-intensive annotation of complex and dense axons and artefacts, automated annotation and data augmentation strategy has been explored in whole-brain projection. However, the existing axon segmentation performance relies heavily on the quality and diversity of annotated training samples. Besides, annotating, training, and tuning the deep model every time from scratch for different stained types and developmental stages, are still not feasible and applicable for the connetomic analysis in the neuroscience community. Here, we develop a large-scale neuron foundation model which can achieve generalized axon and soma segmentation for whole-brain projection mapping.

First, D-LMBmapX trains a self-supervised foundation model based on more than 100,000 cubes in 3D, which are randomly cropped from diverse light-sheet fluorescence microscopy (LSFM) imaged whole brains (Fig. 3A). During the training process, the neuron foundation model divides each 3D cube into regular non-overlapping patches and masking part of them. The Vision Transformer (ViT)[19] encoder is employed for the embedding of unmasked patches, while the decoder tries to reconstruct the original 3D cube in pixel level. Unlike the general masked autoencoder which only reconstruct the original 3D cubes in pixel target[20], D-LMBmapX also reconstructs the original 3D cubes after computing the Histogram of Oriented Gradients (HOG)[21], which can better represent edges in 3D images for the recognition of tree-topological axons. After the foundation model training, automated annotation is applied for task-specific whole brain data, where the axons are automatically annotated based on the previously developed pipeline. The 3D somata are semi-automatically

annotated based on the segment anything model (SAM)[22] and post-processing, where users only need to point few somata in 2D slices (Fig.3B). For the profiling of specific neuron circuits, D-LMBmapX presents a fine-tuning network by paralleling ViT and CNN for the whole-brain predicting of axons and somata respectively, which can quickly convergent and achieve excellent performance by only using few annotated samples.

In comparison with traditional deep models every time training from scratch (e.g., TrailMap[23], D-LMBmap[4]), D-LMBmapX presents the neuron foundation model which demonstrates strong generalization for whole-brain axon segmentation across different brain regions and developmental stages (Fig. 3D, Fig.S2A). Moreover, D-LMBmapX can effectively demonstrate superior quantitative results in both axon and soma segmentation even under a few annotated samples (Fig. 3D, Fig. S2B). The effectiveness of developed foundation model pre-training, as well as the dual target strategy are validated in axon and soma segmentation (Fig. S2C). Based on the neuron foundation model, D-LMBmapX successfully reconstructs dopamine-stained 3D whole-brain axon and soma in 7 postnatal stages from P0 to P28 (Fig. 2E). By registering each whole brain in autofluorescence channel (488nm) to corresponding developmental atlas, D-LMBmapX also computes the average 3D heatmap related to dopamine-stained region-wise axon and soma density in each postnatal stage (Fig. S2D, E). Moreover, benefiting from the developed accurate 2.5D registration method, D-LMBmapX also provides the region-wise dopamine-stained axon and soma density from P0 to P28 (Fig.3F, Extended Fig. S2F). In addition to the dopamine neuron, users can achieve developmental analytics for any type of neuron circuits based on the D-LMBmapX pipeline.

**Cross-Modality 2D-to-3D Whole-Brain Projection Mapping and Validation at Any Postnatal Stage**

D-LMBmapX can achieve 3D whole-brain circuits profiling at any postnatal stage by the constructed developmental atlas and the neuron foundation model. However, in addition to LSFM imaging, brain data can be imaged by other methods, e.g., serial two-photon tomography (STPT)[5,6], fluorescence micro-optical sectioning tomography (fMOST)[8,9], volumetric imaging with synchronized on-the-fly-scan and readout (VISoR)[24]. Besides, many neuroscience labs can only generate 2D histological slices instead of 3D whole brain for neuron circuits profiling. Therefore, to enable the broader neuroscience community to benefit from D-LMBmapX, we further develop cross-modality and cross-dimension methods for the neuron projection mapping and validation in any postnatal stages.

Firstly, D-LMBmapX develops a novel pipeline for fine-grained brain image style transfer based on the diffusion model, which can transfer source brain image modality to any given target style (Fig. 4A). In the training process, given the target style brain image $x_0$, the diffusion process incrementally makes the image a bit noisier until pura noisy $x_T$. Subsequently, the pura noisy image $x_T$ tries to reconstruct the target style image $x_0$ through the denoising process. Different from original denoising diffusion probabilistic model, D-LMBmapX incorporates two modules in the denoising process for more accurate and stable target style reconstruction, i.e., brain edge constrains, and denoising U-Net. Especially, to ensure the consistency of brain structures after style transfer, edges in brain images are automatically extracted which serve as constrains in the denoising processing. Moreover, we introduce a segmentation network (U-Net) in each denoising processing for the automated segmentation of major brain regions (BS, CB, CTX, HPF, CP). This strategy can well preserve the denoising optimization from whole-brain to each major brain region. In comparison with other brain style transfer methods, D-LMBmapX can achieve excellent performance for the brain style transfer in three viewpoints under different evaluation metrics (Fig. 4B). In addition to the brain style transfer from LSFM to STPT (Allen CCFv3), D-LMBmapX can even achieve excellent brain style transfer from Allen histological 2D slices to our developmental LSFM brain atlas (Fig. S3A, Fig. S3B).

Secondly, D-LMBmapX also develops a novel pipeline for the automated 2D histological slices localization and registration to 3D brain atlas (Fig. 4C). The 2D histological slices (e.g., in sagittal view) are processed and style transferred to the 3D brain atlas style. D-LMBmapX can automated localize similar 2D slices in an orthogonal view from atlas. Then similar 2D slices and their neighboring slices within two angles of freedom (in horizontal and coronal views) are further ranked by computing the slices similarity with the preprocessed image and the style transferred image. The most similar slice is considered as the matched 2D slice in 3D brain atlas. Finally, the histological 2D slice is registered to the matched 2D slice in 3D atlas to compute the region-wise soma density, where the soma in histological 2D slices are automatically segmented. To the best of our knowledge, this is the first pipeline to achieve the 2D brain slice localization to 3D brain. Given the 2D histological slices, the pipeline can localize reasonable and corresponding slices in 3D atlas, as well as achieve reliable registration for whole-brain soma profiling.

## DISCUSSION

Here, we presents D-LMBmapX, a machine-learning powered tool for whole-brain profiling of developmental neuronal circuitry in mice.. D-LMBmapX contructed a high-resolution 3D light-sheet based brain atlas in 7 postnatal

stages (P0, P4, P7, P10, P14, P21, P28), and implemented a newly designed cross-stage registration module for 3D alignment at any stages,. It can support volumn quantification of hundreds of brain regions during development. D-LMBmapX also provides a neuron foundation model for accurate and robust whole-brain axon and soma segmentation. Moreover, D-LMBmapX successfully quantifies region-wise signal mapping of 2D histological slices to developmental 3D atlas through novel brain style transfer and 2D to 3D localization methods. All developmental atlas, tools and methods have packaged as user-friendly software, which is publicly available for the neuroscience community.

Accurate 3D brain atlas with fine-grained structure delineation is considered a golden standard for neuroscience research. Allen Institute provided three versions of the adult mouse brain common coordinate framework (CCF) in the last 15 years. Allen CCFv3 is the widely accepted 3D atlas for adult whole mouse brain analytics. However, for the developmental atlas, Allen only provides tens of 2D histological slices with annotations in sagittal view at two postnatal stages (P4 and P14), which cannot meet the wider range of analytical tasks for related developmental circuits. Accordingly, D-LMBmapX constructs postnatal atlases in 7 stages, and provides cross-stage registration for any developmental analysis. By introducing major and small brain region constrains, the postnatal atlas is more accurate and reliable with high-resolution than other developmental atlases.

Automated whole-brain axon and soma segmentation in mesoscale is generally the most time-consuming and labor-intensive step for circuitry profiling. Due to the diverse brain image backgrounds, staining methods, neuron types, and morphological changes during development, deep models for axon and soma segmentation generally require massive annotated data with every time training from scratch, which is not convenient for the accessibility in neuroscience field. D-LMBmapX presents a powerful neuron foundation model enabling massive brain images for self-supervised pre-training. By only using few training samples for target-specific fine-tuning, D-LMBmapX achieves excellent and robust performance for axon and soma segmentation in a generalized framework.

Additionally, to further extend D-LMBmapX with wider applicability for developmental whole-brain neural circuits profiling, we consider two practical scenarios for neuroscience labs. First, despite D-LMbmapX constructs 7 postnatal atlas based on LSFM with the tissue clearing of iDISCO, many labs uses other brain imaging microscopes and tissue preparation methods. D-LMbmapX provides a novel diffusion model-based brain style transfer

technique to merge cross-modality gaps for circuitry profiling. Second, many labs can only afford the brain imaging of 2D histological slices. In this case, D-LMBmapX provides a 2D slices to 3D atlas localization and registration pipeline, finding the most similar section in 3D atlas regardless of imaging position, rotation, and deformation. We believe D-LMbmapX is a powerful tool for automated developmental profiling of whole-brain neural circuits with generalized applicability to neuroscience researchers.

## ACKNOWLEDGEMENTS

We thank the MRC Laboratory of Molecular Biology (LMB) Biological Service Group and the Ares staff for their support with animal husbandry. This work was supported by the Medical Research Council, as part of United Kingdom Research and Innovation (UK Research and Innovation) (MC_UP_1201/22). For the purpose of open access, the Medical Research Council Laboratory of Molecular Biology has applied a CC BY public copyright license to any Author Accepted Manuscript version arising. A.M.A was supported by the MRC LMB PhD program with an MRC studentship. This work was also partially funded by NARSAD Young Investigator Award (2020, BBRF) to J.R..

## AUTHOR CONTRIBUTIONS

J.R. conceptualized the project. J.R. and Z.L. supervised the project. Z.L. led the development of D-LMBmapX. J.R. and Z.L. designed all the experiments. Z.L., P. L., Z.H., and S.Z. designed and implemented the neuron foundation model. Z. L., H.P. and J. L. designed and constructed the developmental brain atlas. Z. L., B.Z., and J.Z. designed and implemented the 2.5D registration methods. Z.L., H.P., H. L., P.L. and D.Z. designed and implemented the brain style transfer and the 2D to 3D localization pipeline. Z.L., Y.S., and P.L. developed the D-LMBmapX software. Z.L., P.L., H.P., H.L., Z.H., Y.S., A.MJ.A, and J.R. co-annotated and checked the testing data. A.MJ.A. performed brain sample collections, tissue-clearing, immunotaining and light-sheet imaging. L.B. performed brain sample collections. R.N.S. assisted with the tissue-clearing. Z.L. and J.R. wrote the manuscript with input from all co-authors.

## REFERENCE

1. Ren, J. *et al.* Anatomically Defined and Functionally Distinct Dorsal Raphe Serotonin Sub-systems. *Cell* **175**, 472-487.e20 (2018).
2. Winding, M. *et al.* The connectome of an insect brain. *Science* **379**, eadd9330 (2023).
3. Ngai, J. BRAIN @ 10: A decade of innovation. *Neuron* **112**, 3003–3006 (2024).

4. Li, Z. *et al.* D-LMBmap: a fully automated deep-learning pipeline for whole-brain profiling of neural circuitry. *Nature Methods* **20**, 1593–1604 (2023).
5. Ragan, T. *et al.* Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nat Methods* **9**, 255–258 (2012).
6. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
7. 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).
8. Harris, J. A. *et al.* Hierarchical organization of cortical and thalamic connectivity. *Nature* **575**, 195–202 (2019).
9. Xu, F. *et al.* High-throughput mapping of a whole rhesus monkey brain at micrometer resolution. *Nat Biotechnol* **39**, 1521–1528 (2021).
10. Dodt, H.-U. *et al.* Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nat Methods* **4**, 331–336 (2007).
11. Hobson, C. M. *et al.* Practical considerations for quantitative light sheet fluorescence microscopy. *Nat Methods* **19**, 1538–1549 (2022).
12. Thompson, C. L. A High-Resolution Spatiotemporal Atlas of Gene Expression of the Developing Mouse Brain. *Neuron* **83**, 309–323 (2014).
13. Young, D. M. *et al.* Constructing and optimizing 3D atlases from 2D data with application to the developing mouse brain. *eLife* **10**, 1–40 (2021).
14. Kronman, F. N., Liwang, J. K., Betty, R., Vanselow, D. J. & Wu, Y.-T. Developmental mouse brain common coordinate framework. *Nature Communications* **15**, 1–17 (2024).
15. 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).
16. Wang, Q. *et al.* The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell* **181**, 936-953.e20 (2020).
17. Tustison, N. J. *et al.* N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* **29**, 1310–1320 (2010).
18. Amorim, P., Moraes, T., Silva, J. & Pedrini, H. 3D Adaptive Histogram Equalization Method for Medical Volumes. *VISIGRAPP* (2018).
19. Han, K. *et al.* A Survey on Vision Transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 87–110 (2023).
20. He, K. *et al.* Masked Autoencoders Are Scalable Vision Learners. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2022).
21. Dalal, N. & Triggs, B. Histograms of Oriented Gradients for Human Detection. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* vol. 1 886–893 (IEEE, San Diego, CA, USA, 2005).
22. Kirillov, A. *et al.* Segment Anything. *Proceedings of the IEEE/CVF*

*International Conference on Computer Vision* (2023).
23. Friedmann, D. *et al.* Mapping mesoscale axonal projections in the mouse brain using a 3D convolutional network. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11068–11075 (2020).
24. Wang, H. *et al.* Scalable volumetric imaging for ultrahigh-speed brain mapping at synaptic resolution. *National Science Review* **6**, 982–992 (2019).

## FIGURE LEGENDS

**Figure 1.** Overview of D-LMBmapX.

A. The constructed 3D brain atlases in postnatal developmental stages P0, P4, P7, P10, P14, P21, and P28. P4 and P10 brain samples also serve as intermediate stages and are utilized for validating cross-stage registration.

B. The pipeline for cross-stage registration of any postnatal stage. Given an intermediate stage (using the P4 brain as an example), brain structures can be adaptively quantified through cross-stage registration with its two flanking atlases.

C. The pipeline for quantitative analysis of axon and soma from the same brain samples. Utilizing a neuron foundation model, extensive 3D cubes are extracted from brains at postnatal stages and then self-supervised training is applied for whole-brain segmentation of axon and soma.

D. Cross-modality and cross-dimension whole-brain profiling. Example of registering and quantitative analyzing a 2D histological slice from Allen. The slice is style-transferred to match the LSFM format, and the corresponding slice from the digitally sectioned 3D atlas is automatically localized and registered. Soma segmentation is then performed, followed by mapping for region-wise quantification of soma density.

E. The software interface of the D-LMBmapX for developmental whole-brain projection and visualization.

**Figure 2.** Construction of 3D mouse brain atlas for any postnatal stage analysis.

A. The pipeline for developmental 3D brain atlas construction. Original brain images are first pre-processed and voxel-wise averaged as initial reference brain. Then, each brain registers and updates the reference average brain via affine, B-spline, and a proposed 2.5D non-rigid registration. Finally, the average brain upsampled from 25 μm to 10 μm and refined based on high-resolution original brains. (will update)

B. Brain structures delineation in each developmental stage by sequential registration from Allen CCFv3 to P28, then P28 to P21, and so forth. (will update)

C. Quantitative comparison between KimDevCCF atlas and D-LMBmapX atlas in P4 and P14 stages, under the evaluation of Dice score.

D. Cross-stage registration and region-wise quantification for a P10 brain, A P10 brain is registered to atlases in its flanking stages (P7 and P14) respectively, then deformation fields of 5 major brain regions (BS, CB, CP, CTX, HPF) are presented.
E. Quantitative evaluation of cross-stage registration on 5 P10 brains in registering to P7 and P14 atlas. Median Dice score and Deformation score of five individual major brain regions (BS, CB, CP, CTX, HPF) are recorded.

**Figure 3.** Generalized whole-brain projection quantification by implementing neuron foundation model.
A. The pre-training of neuron foundation model. Massive 3D cubes (more than 100,000) are randomly cropped from diverse whole-brain data. The masked autoencoder is self-supervised trained by reconstructing pixel and HOG target simultaneously.
B. Automated annotation of axon and soma in selected 3D cubes. Axon annotation can be achieved by binarization and skeletonization following the developed pipeline in D-LMBmap. Soma annotation can be achieved by 2D point prompting, SAM predicting and post-processing.
C. The fine-tuning network of neuron foundation model. Given few annotated cubes, ViT and CNN are parallelly connected and fine-tuned for fast prediction of whole-brain axon and soma.
D. Quantitative comparison of different axon segmentation methods under the evaluation metric of ClDice in 4 major brain regions from 21 dopamine-stained brains, where each stage includes 3 dopamine-stained brains (P0, P4, P7, P10, P14, P21, P28).
E. Quantitative of different soma segmentation methods under the evaluation metrics of Dice and Precision in 21 dopamine-stained brains, where each stage includes 3 dopamine-stained brains (P0, P4, P7, P10, P14, P21, P28).

**Figure 4.** Cross-modality and cross-dimension whole-brain mapping and validation across various developmental stages.
A. Brain image style transfer of 4 histological 2D slices to light-sheet based developmental atlas style. From left to right, original histological 2D slices, histological 2D slices after image preprocessing, extracted edges in histological 2D slices, after brain style transfer using the proposed diffusion based deep model, corresponding referred slices in developmental 3D atlas.
B. Quantitative evaluation and comparison of the proposed diffusion-based style transfer with CyclaGAN and DDPM, where three evaluation metrics (SSIM, PSNR, FID) are used for evaluation. Nine 3D brain images are used for evaluation in three viewpoints (Sagittal, coronal, and horizontal).
C. Schematic drawings showing the Pipeline developed for automated whole-brain mapping and quantification of histological 2D slices. Given a histological

2D slice, the preprocessed and style transferred images are used to find the similar 2D slices in corresponding developmental 3D atlas. A similarity measuring algorithm is designed to find the matched 2D slices in atlas. Then the histological 2D slice and segmented soma are registered to the matched 2D slices for the region-wise soma density projection mapping.
D. Three examples of the pre-processed 2D histological slices, the corresponding matched 2D slices in developmental 3D atlas, and the registration results of the match 2D slices to histological slices.
E. Cross-validation of the region-wise heatmaps generated by D-LMBmap based on Allen histological in situ 2D slices and light-sheet imaged 3D whole-brain of TH+ somata at P28, P14 and P4.

**SUPPLEMENTAL INFORMATION**

**Figure S1.** Pipeline of 2.5D resignation for 3D brain atlas construction and brain region segmentation.
A. Schematic drawings showing the pipeline developed for 2.5D registration. The 3D moving brain is registered to the 3D fixed brain in coronal and horizontal views respectively based on 2D image slices. Then the registered 2D slices are combined in two views in computing the 3D consistency loss.
B. Quantitative comparison of the proposed 2.5D registration method with other 3D brain registration methods. (will update)
C. The constructed 3D brain atlas in 7 postnatal stage stages, including P0, P4, P7, P10, P14, P21, and P28.

**Figure S2.** Quantitative and qualitative analysis of TH+ whole-brain projectome.
A. Quantitative comparison of different axon segmentation methods under the evaluation metrics of ClDice, ClPrecision, ClRecall, and Dice score each developmental stage, where each stage includes 3 dopamine-stained brains.
B. Annotation efficiency evaluation and comparison of the neuron foundation model for axon and soma segmentation. Given a set of annotated training samples, ClDice are recorded when using 5%, 20%, 50%, and 100% of them respectively.
C. Ablation study for the effectiveness evaluation of the proposed neuron foundation model pre-training and the HOG target reconstruction.
D. Average density heatmap of dopamine-stained whole-brain axon across each developmental stage from P0 to P28, where n=3 in each stage.
E. Average density heatmap of dopamine-stained whole-brain soma across each developmental stage from P0 to P28, where n=3 in each stage.

**Figure S3.** Cross-modality and cross-dimension processing for whole-brain projection mapping.
A. The framework of the proposed diffusion-based brain image style transfer.

Different with commonly used diffusion model, a denoising U-Net is introduced in the diffusion phase for preserving the brain region after style transfer.
B. Illustrations of brain image style transfer from LSFM to Allen CCFv3 in three viewpoints.

## STAR★METHODS

### Animal care and use

All experiments related to the use of mice in the Medical Research Council Laboratory of Molecular Biology were carried out in accordance with the UK Animals (Scientific Procedures) Act of 1986, with local ethical approval provided by the Medical Research Council Laboratory of Molecular Biology Animal Welfare Ethical Review Board and overseen institutionally by designated animal welfare officers (Animal Project License PP6471806).

### Data preparation

All 3D brains were cleared based on the iDISCO+ protocol, and then imaged in horizontal view through a light-sheet microscope (Ultramicroscope II, LaVision BioTec). 3D brains were scanned on 647nm (axon and soma stained specific) and 488nm (autofluorescence) channels with dynamic focus and the step size of 3 μm. The original brain images in 488nm channel were pre-processed by rotation, cropping, mirroring, contrasting, then resized to 1320*800*1140 for developmental atlas construction. In the developmental atlas construction, each 3D brain image in autofluorescence channel was further pre-processed and normalized through N4 bias correction, intensity rescale, and 3D contrast limited adaptive histogram equalization, which can enhance original image quality and normalize 3D brains in each stage.

In the evaluation of neuron foundation model for the whole-brain axon segmentation, we manual annotated 105 cubes from P0 to P28 in the size of 150×150×150 with voxel-level ground truth of dense axons. In each stage, 3 cubes were randomly selected from 5 brain regions respectively (CB, BS, CTX, HPF, CP). For the evaluation of soma segmentation, 44 cubes were manual annotated with voxel-level ground truth of somata. All manual annotated cubes were only used for evaluation. In the pre-training of neuron foundation model, 100,000 cubes were randomly selected from dopamine-stained brains in each developmental stage from P0 to P28. The pre-training of neuron foundation model was performed using 8 NVIDIA A100 Tensor Core GPU. Other training and testing tasks were performed using 4 NVIDIA GeForce RTX 3090 GPU.

### Developmental atlas construction

We designed a 2.5D registration method for the developmental atlas

construction and cross-stage brain alignment (Extended Data Fig. 1A). First, the 3D moving brain was initially aligned to the 3D fixed brain through rigid and affine transformation. Afterwards, 2D slices were extracted in both coronal and horizontal views. For corresponding slices in moving and fixed brains, the U-Net deep neural network was used for the 2D registration in coronal and horizontal slices respectively, under the deformation loss $L_{smo}$ and the registration loss $L_{sim}$ in two views. Besides, we introduced a consistency loss $L_{consis}$ by computing the similarity of two register 3D brains assembled individually in coronal and horizontal views. In comparison with general 3D registration method, the proposed 2.5D registration method do not have high computational overhead, since the main registration and optimization computation is in 2D. This advantage allows our algorithm to handle higher-resolution brain images.

D-LMBmapX collected at least 30 half brains in each stage and mirroring them for the average brain construction. In each stage, given $N$ 3D brains after pre-processing, the initial reference brain is first obtained based on all 3D brains by computing the center alignment of their voxel average. Subsequently, each brain is continuously registered to the initial reference brain and compute the voxel average based on the rigid and affine transformation. To achieve more accurate and fine-grained average brain atlas, each brain is iteratively registered to the average brain using the designed non-rigid 2.5D registration method (Fig. 2A).

After the averaged brain construction in each developmental stage from P0 to P28, the key problem is the delineation of hundreds of brain regions for each average brain. It is extremely labor-intensive to manually delineate all brain regions by exhaustive staining of diverse genetic specific 3D brains for all 7 postnatal stages. Here, we used the most widely accepted adult mouse brain atlas Allen CCFv3 as the reference for brain region delineation of developmental atlas by cross-stage and cross-modality registration with multiple constrains. Considering the shape similarity between the adult atlas Allen CCFv3 and our constructed atlas P28 averaged brain, we first register Allen CCFv3 to the P28 averaged brain based on the 2.5D registration method, embedding brain style transfer and multiple brain regions as registration constrains. Especially, the proposed diffusion-based brain style transfer is used to transform brain image modality from Allen CCFv3 (STPT modality) to P28 average brain (LSFM modality). Brain regions in each developmental averaged brain are manual annotated, including 6 major brain regions (CTX, CB, BS, HPF, CP, OLF), and 10 small brain regions, i.e., Red nucleus (RN), Dorsal part of the lateral geniculate complex (LGd), Medial habenula (MH), Lateral habenula (LH), fasciculus retroflexus (fr), anterior commissure olfactory limb (aco), anterior

commissure temporal limb (act), Basolateral amygdalar nucleus (BLA), Ventral part of the lateral geniculate complex (LGv), Anterior pretectal nucleus (APN). These small brain regions mainly located in brain stem, which are also relatively obvious in LSFM based averaged brain. After the cross-modality and cross-stage registration from Allen CCFv3 to P28 average brain, hundreds of brain regions can be automatically delineated. Afterwards, brain regions in P21 average brain can be also automatically delineated by the registration from P28 to P21, using 6 major and 10 small brain regions constrains, and so forth in delineating brain regions for each developmental atlas from P14 to P0. Due to the structural and regional shape diverse between Allen CCFv3 and each developmental atlas, directly registration may be unable to match hundreds of brain regions and cause the alignment shifting and error accumulation problem during the delineation in earlier stages, especially the small and numerous brain regions in BS. In our solution, the introduction constrains of ten small brain regions, the alignment shifting and registration error accumulation problem can be well alleviated.

**Neuron foundation model**

D-LMBmapX developed a neuron foundation model for generalized axon and soma segmentation, which can effectively adapt to diverse brain images and neuron types. The neuron foundation model can be divided into three steps. The first step is the self-supervised pre-training of neuron foundation model. In this step, we randomly selected 100,000 cubes with the size of 150×150×150 from the dopamine-stained whole brains in 7stages. All these cubes were set as input in the 3D masked autoencoder (MAE) network, which trained parameters of the Vision Transformer (ViT) encoder and decoder, by try to reconstruct all voxels of the original 3D cubes. Different with the common decoder module in MAE network which only reconstruct the original image in pixels, we further introduce another decoder branch for the reconstruction of the corresponding 3D Histogram of Oriented Gradients (HOG) images. The 3D HOG cubes were extracted by computing the edge of the original 3D cubes. The dual decoder reconstruction setting can better represent the long-range tree-topological structure of axons.

The second step is the annotation of axon and soma for target-specific fine-tuning (Fig. 3B). For the annotation of axons, we follow the automated annotation method in D-LMBmap. Users first selected tens of predominantly “axon” and “artificial” cubes. Then artificial cubes were set as background. The “axon” cubes were binarized and skeletonized through the pipeline in D-LMBmap. For the annotation of somata, we designed a semi-automated pipeline for 3D soma annotation. Especially, we employed the most widely used segmentation foundation model in natural images, i.e., segment anything model

(SAM). Given a 3D cube for annotation, users first provided point prompt for each soma in 2D slices, which is the only step from manual input. Afterwards, we employed the widely used segmentation foundation model in natural images, i.e., segment anything model (SAM), which can automatically segment somata based on the point prompts. Subsequently, a watershed algorithm based on distance maps is introduced to further separate adhered somata. These lightweight annotation strategies can be quickly extended to diverse axon and soma segmentation tasks.

The third step is the fine-tuning of the pre-trained neuron foundation model using the annotated cubes in specific tasks (Fig. 3C). In the network, the pre-trained ViT encoder is directly used for target-specific fine-tuning. However, due to the large number of parameters in the ViT encoder, directly fine-tuning is generally inefficient, which still requires a considerable amount of training data. Here, we introduced a parallel convolutional neural network, which not only enhances the fine-tuning and model's convergence speed, but also improve the model's accuracy. The output of ViT encoder in each level is combined with the CNN output in each corresponding level. Based on this network architecture, only few training samples can finish the fine-tuning of neuron foundation model. Experimental results also demonstrated that in the neuron foundation model and the fine-tuning framework, few annotated samples can also achieve excellent and comparable performance. This advantage can be easily adapted to axon and soma segmentation tasks, as well as the whole-brain segmentation in different stages and stained signals.

**Diffusion-based brain style transfer**

In the previous work, D-LMBmap presented the first brain style transfer method based on CycleGAN, which learns an unsupervised network for the brain image transfer from modality A to B. However, the network performance relied on the two brain modalities, which requires train from scratch if one or two brain modalities are different in a new task. Besides, we find that the CyclaGAN based method cannot well preserve detailed brain structure consistency after style transfer. Here, we presented a novel brain style transfer solution based on the diffusion model (Fig. 4A). In the training step, the model only using the target brain image (e.g., developmental atlas brain). The original target brain image is processed with a set of noise adding step (forward diffusion process) until pure Gaussian noise. In the denoising process, we introduce a denoising U-Net to segment 6 major brain regions (BS, CB, CTX, HPF, CP, OLF) as the constrains for the generation control. Besides, we also introduced edges in brain images as constrain, where the generated image has consistent edges with the original input image $X_t$. In the predicting step, source brains in other modalities can be processed based on the denoising process, where the source brain edges were

extracted in controlling the generation of the target images. The presented diffusion-based brain style transfer has two advantages: 1) brain edges and details are consistent between the source image and the target image; 2) the diffusion deep model only need to train one time for specific target modality, with no requirements of the source modality.

**2D slice localization and registration to 3D atlas**

Based on the constructed 3D atlas, D-LMBmapX further proposed a novel pipeline for the cross-dimensional whole-brain neuron circuitry projection of 2D slices, especially for the widely used 2D histological images. In the first step, we designed a 2D to 3D localization method for the finding of the most similar section in 3D brain atlas with regard to cross-modality 2D slices (Fig. 4C). 2D slices were first pre-processed and style transferred to atlas style if brain modalities are different. Here, we introduce two metrics in measuring the similarity between 2D slices and sections in 3D atlas, i.e., image-level structural similarity index measure (SSIM) and landmark-level corner feature similarity (CFS). The SSIM is the generally used metric in measuring the similarity of two images. The landmark-level is computed based on the corner detection and landmark feature representation. Afterwards, initial *top-K* similar sections can be located by computing and ranking the two metrics between the 2D slice and all section in 3D atlas orthogonal view (e.g., $N$ sections in sagittal view). Considering the 2D slices may be cut in diverse angles during the tissue preparation, we added two freedoms of degrees with the range from $-M^{o}$ to $M^{o}$. Subsequently, the matched sections can be automatically located by the second round of computing and ranking based on the two metrics. In our 2D to 3D localization solution, the image-level and landmark-level metrics need to be calculated in total $2 \times N \times K \times 2M \times 2M$ times in finding the matched sections in 3D atlas.

After finding the most similar sections in 3D atlas, next steps are the signals (axon and soma) extraction and segmentation in 2D slices, and the registration to corresponding sections for region-wise mapping. The signal extraction and segmentation in 2D slices are generally easier in comparison with the task in 3D whole-brain due to the higher resolution and clear imaging condition. For the 2D histological images with stained somata from Allen, we follow the semi-automated annotated pipeline for 3D cubes (Fig. 3B). Then a weak-supervised network is used for the segmentation of somata in Allen's 2D histological slices. For the 2D slices imaged by confocal microscopes, we follow the automated annotation pipeline in D-LMBmap. Then a 2D nnU-Net model is trained for the automated axon segmentation. In the 2D registration task, we employed the ANTs for the rigid, affine, and b-spline registration between the 2D slice and the corresponding section in 3D atlas. If users require more accurate results, the

2.5D registration can be used, where we trained a light-sheet based registration model in three views (coronal, horizontal, and sagittal).

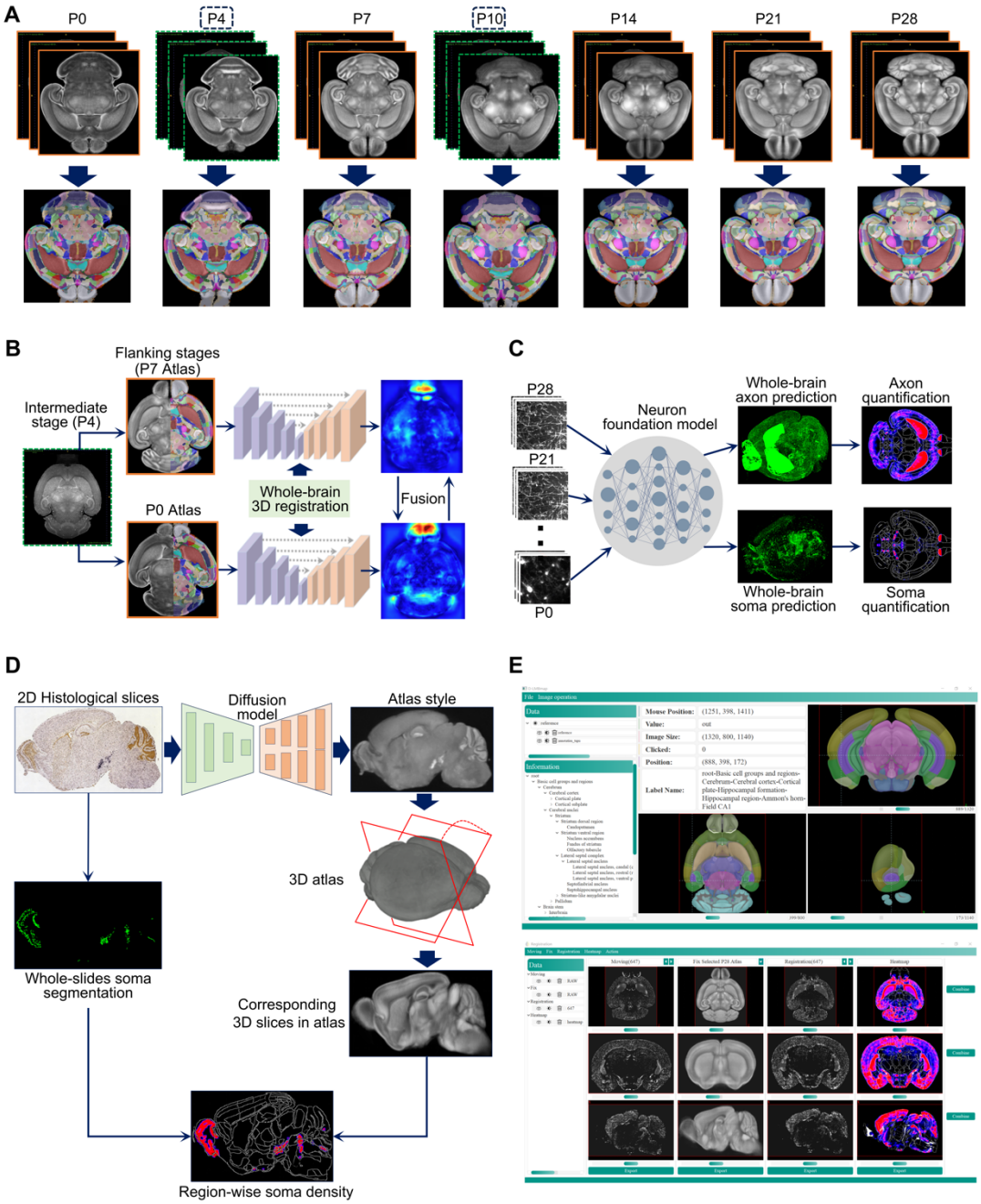


**Figure 1.** Overview of D-LMBmapX.

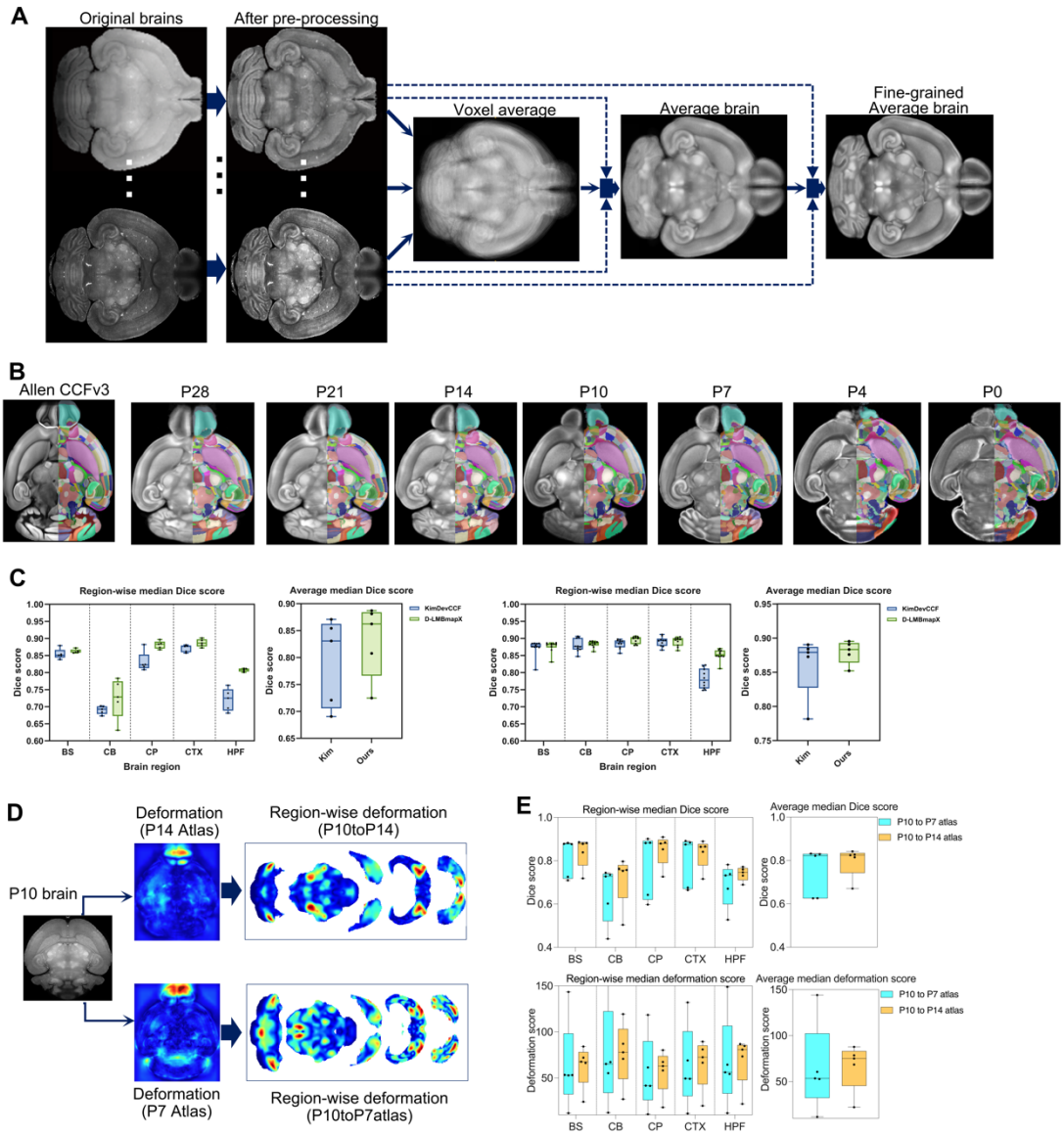


**Figure 2.** Construction of 3D mouse brain atlas for any postnatal stage analysis.

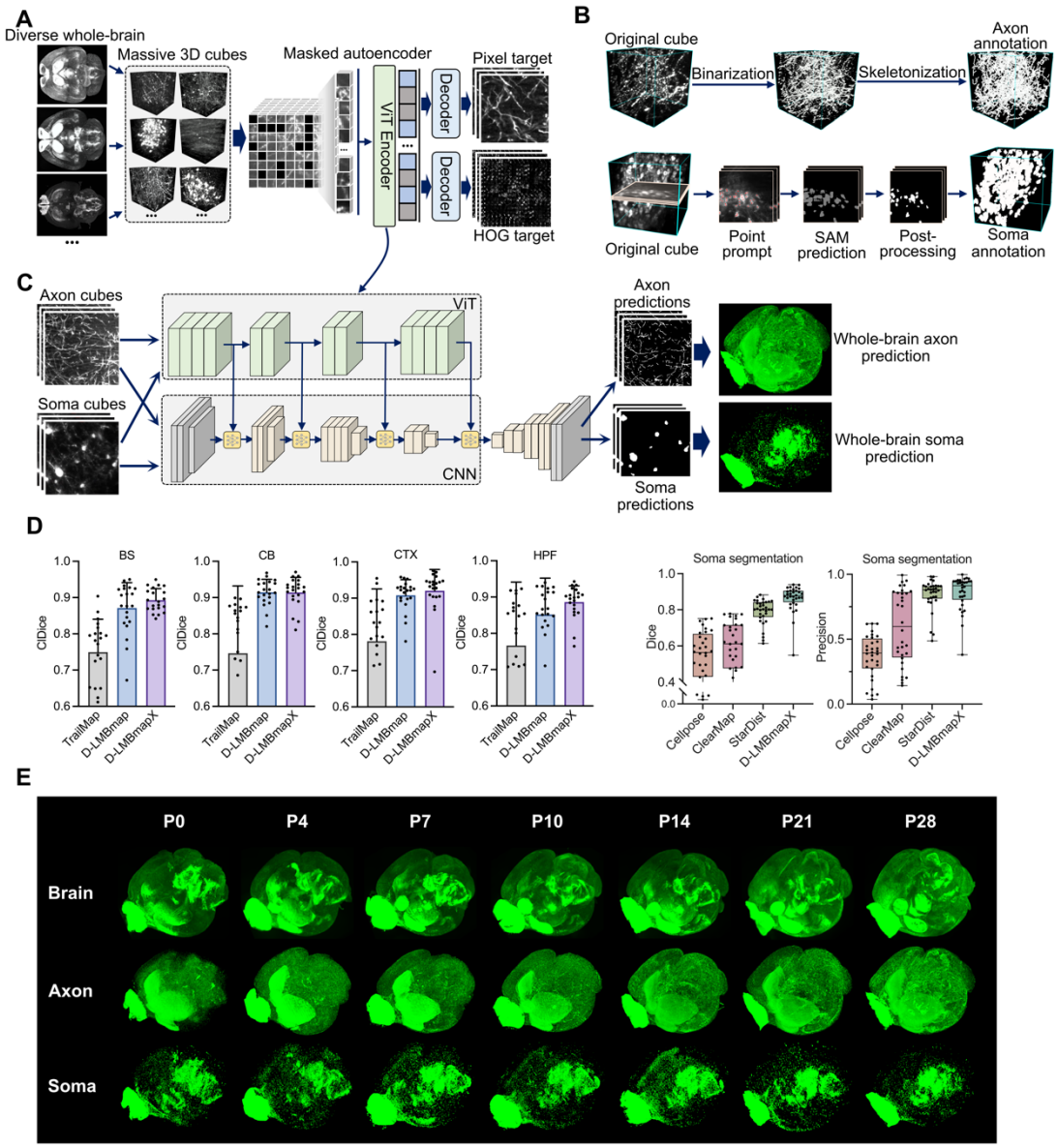


**Figure 3.** Generalized whole-brain projection quantification by implementing neuron foundation model.

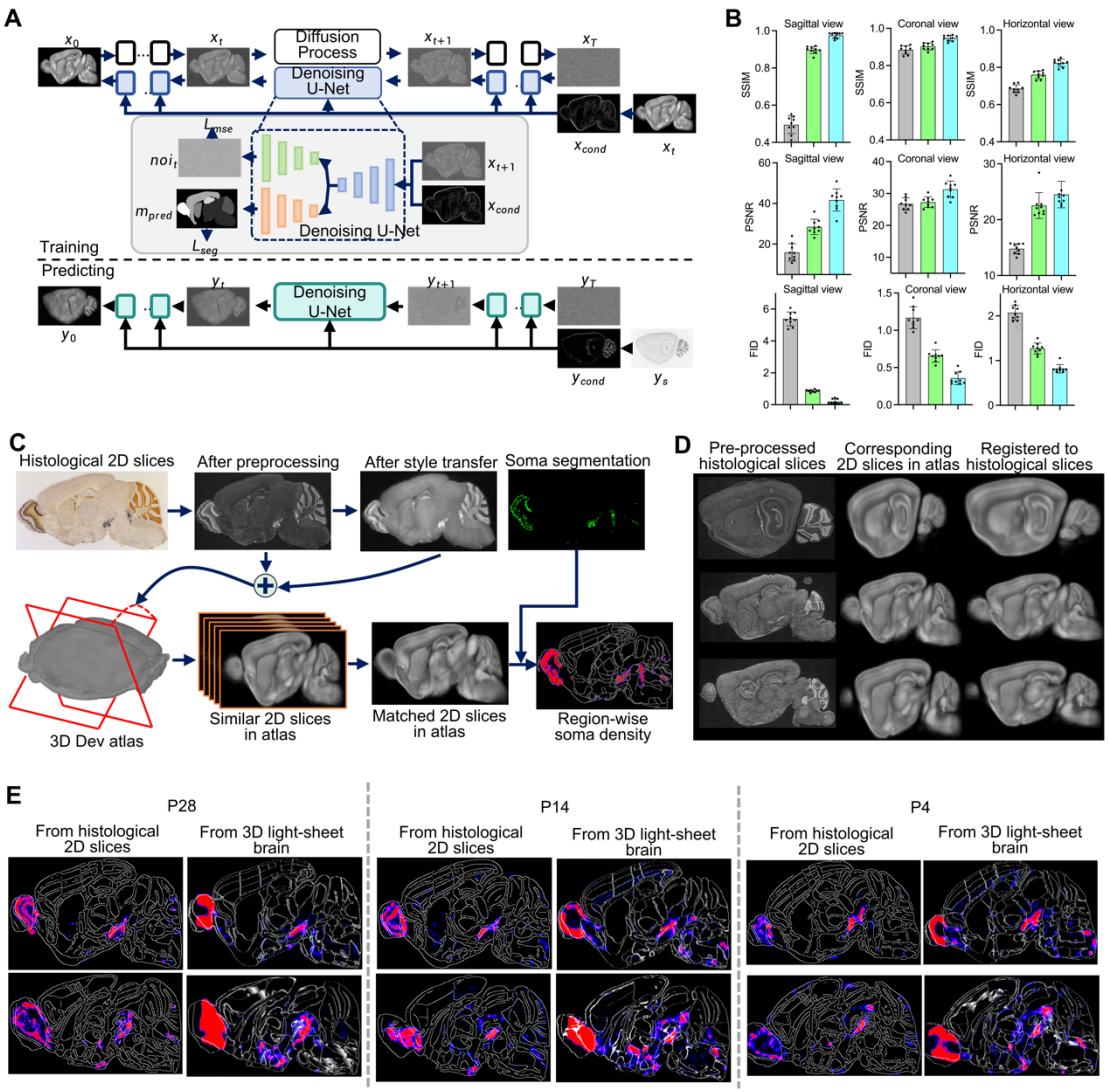


**Figure 4.** Cross-modality and cross-dimension whole-brain mapping and validation across various developmental stages.