Generative cerebral vasculature visualization using spatial transcriptomic data

Ingrid Berg[1], Jiqing Wu[2,*], Viktor H. Koelzer[1,2,3,*]

[1]Department of Pathology and Molecular Pathology, University Hospital Zurich, University of Zurich, Switzerland

[2]Department of Biomedical Engineering, University of Basel, Switzerland

[3]Institute of Medical Genetics and Pathology, University Hospital Basel, Switzerland

*Corresponding Authors: Jiqing Wu, Viktor H. Koelzer

Email: Jiqing.Wu@unibas.ch, Viktor.Koelzer@usb.ch

Author Contributions: I.B., J.W. and V.H.K. conceived the research idea. J.W. implemented the algorithm and carried out the experiments. I.B., J.W. and V.H.K. analyzed the results. I.B. drafted the manuscript. J.W. and V.H.K. critically reviewed and edited the manuscript. J.W. and V.H.K. supervised the project.

Competing Interest Statement: I.B. and J.W. declare no competing interests. V.H.K. declares project-based research funding from Roche and the Image Analysis Group outside of the submitted work. V.H.K. is on an advisory board of Takeda, has served as an invited speaker on behalf of Indica Labs and for Sharing Progress in Cancer Care, an independent nonprofit organization, outside of the submitted work.

Classification: Major: Biological Sciences, Minor: Neuroscience

Keywords: Spatial transcriptomics, generative modeling, cerebral vasculature

This PDF file includes:

Main Text
Figure 1

Abstract

The brain is sustained by an intricate vascular network that provides a continuous supply of nutrients and oxygen essential for its function. Understanding the structural variability and region-specific function of this neurovascular system is essential for interpreting neurological alterations in disease models. Leveraging the spatial mRNA-guided generative model Tera-MIND, we propose to predict the spatial organization of brain vasculature by the co-expression patterns of *Cldn5* and *Acta2* genes and their learned attention map at cell-level resolution. These findings showcase the modeling capability of generative AI trained on single cell spatial mRNA data for higher order structures and underscores its potential for high-throughput investigation of critical vascular functions and their dysregulation in neurological disorders.

Significance Statement

The brain's vascular network is complex and functionally essential. Accurate reconstruction of this network at single-cell transcriptomic resolution is key to understanding neurovascular disorders, where vascular gene expression disruption contributes to pathogenesis. This study applies spatial transcriptomic and histological data to generate representations of brain vasculature using Tera-MIND, providing a framework that offers a scalable approach for understanding vascular organization.

Main Text

Introduction

Neurons rely on continuous blood flow due to their high energy requirements and low capacity for energy storage. Even small disruptions in blood flow can impair neuronal function and survival. Neuronal activity, in turn, modulates blood flow, and several neurological diseases involve vascular dysfunction[1]. Mapping the cerebrovascular network is challenging due to the large variety of blood vessel calibers, from capillaries to large vessels, in a complex 3D structure. However, brain-wide vascular maps are fundamental for understanding how vascular network organization supports neural function and how changes in this system contribute to disease[1]. Recent developments in generative artificial intelligence (GenAI) have supported the in-silico synthesis of realistic biological structures spanning different spatial scales. With complex datasets such as biomedical images and spatial transcriptomics as inputs, GenAI models can learn to reconstruct tissue microanatomy, simulate morphological transitions, and create "digital twins" of organs[2]. In the recent Tera-MIND study, the authors proposed a novel GenAI approach that can capture molecular-to-morphology spatial associations, exemplified by key pathways involved in glutamatergic and dopaminergic neuronal systems[3]. In this application study, we further investigate the potential of generative AI to model biologically meaningful higher-order structures using single-cell spatial transcriptomics data.

Results

With Tera-MIND, we can visualize an integrated model of gene expression in the neurovascular context. Specifically, we apply the Tera-MIND approach (Fig. 1A and 1B) to generate high-resolution, spatially-resolved visualizations of cerebral microvascular structures, creating a virtual representation of murine brains (exemplified in Fig. 1C and 1D) with projected microvasculature. This enables the spatial analysis of gene expression patterns in relation to vascular architecture and neurovascular function (Fig. 1E-H). In this application study, we focus on two vasculature-associated genes for which spatial transcriptomic information was available in all three mouse brain data sets: *Cldn5* and *Acta2*. Both *Cldn5* and *Acta2* expression play key roles in the function of the

neurovascular unit, and dysfunction of these genes is associated with various neurological diseases.

Claudin-5

Cldn5, a gene encoding the endothelial tight-junction protein claudin-5, is critical for blood-brain barrier (BBB) integrity and is known to exhibit highest expression in brain capillaries and small venules[4]. Using the Tera-MIND model, we predicted *Cldn5* expression across the brain and observed spatial patterns that correspond closely to known vascular anatomy (Fig. 1F). Notably, regions with longitudinally sectioned vessels, such as those extending from the cortex into deeper brain layers[5], exhibited strong and continuous predicted *Cldn5* signal (Fig. 1H, representative merged image of four adjacent sections), consistent with the expected localization of capillary-rich zones. These findings demonstrate that the model successfully captures biologically meaningful vascular features relevant to BBB-associated gene expression.

Alpha smooth muscle actin

Acta2, which encodes alpha smooth muscle actin, is predominantly expressed in vascular smooth muscle cells and marks arterial structures, particularly larger arteries[1]. Tera-MIND predictions of *Acta2* expression localized to regions consistent with major arterial territories, including the ventral surface of the brain where larger arteries are anatomically expected[6] (Fig. 1E, 1H). To explore spatial interactions between vascular components, we jointly analyzed *Acta2* and *Cldn5* expression using gene-gene attention mapping. This analysis revealed co-localized expression within shared vascular segments (Fig. 1G), suggesting that the model accurately captures complementary features of the vascular hierarchy—from endothelial markers of the blood-brain barrier to smooth muscle-associated arterial identity.

Vascular network reconstruction

To reconstruct the vascular architecture of the healthy mouse brain, we applied the Tera-MIND model to consecutive 2D spatial transcriptomics sections, focusing on genes associated with the neurovascular unit. The resulting 3D-integrated predictions enabled visualization of vascular continuity and organization across anatomical regions. Fig. 1H illustrates the spatial distribution of vascular features, with predicted *Cldn5* and *Acta2* expression highlighting endothelial and smooth muscle components, respectively. These predictions align with known vascular structures and support downstream analyses of spatial gene expression in the context of vascular integrity.

Discussion

In this study, we present a generative, gene-informed digital twin of the murine cerebral vasculature, reconstructed at teravoxel scale from single-cell spatial transcriptomics data. By leveraging vascular marker genes, the model enables the formation of coherent, spatially-continuous vascular structures that reflect known anatomical and molecular features. This virtual representation supports high-resolution spatial analysis of gene expression in relation to vascular architecture, offering a new framework to study neurovascular organization and function in healthy tissue. Importantly, this work establishes a scalable approach for modeling higher-order tissue structures from single-cell spatial data within their native spatial context, enabling the investigation of gene expression at cellular resolution.

Looking ahead, this modeling framework could be extended to incorporate additional modalities and conditions, enabling more comprehensive 3D simulations of tissue organization and dynamics. Integration with complementary histological, imaging, and genomic datasets may enhance the model's capacity to capture morphological and functional variability. While this study focuses on

healthy brain tissue, the generative approach lays the foundation for future in silico exploration of disease-associated perturbations—such as blood-brain barrier dysfunction—once appropriate spatial transcriptomic datasets from disease models become available.

A key strength of our approach lies in its multi-scale ability for interpretable image prediction, enabling detailed visualization of vascular structures from gene expression data. However, a limitation of the current study is that 3D reconstruction is based on ground truth data with non-contiguous 2D sections with 200 micron gaps[7]. This inherent sparsity in the ground truth data introduces challenges for accurate 3D vascular network reconstruction, potentially leading to discontinuities or artifacts. One promising strategy to address this limitation has been shown by Turos *et al.* with X-Pression, which combines single 2D spatial transcriptomics sections with micro-CT imaging to enable full 3D reconstruction of gene expression patterns at the tissue level—eliminating the need for serial sectioning of entire organs[8].

Guided by the spatial co-expression of vasculature-relevant genes *Cldn5* and *Acta2*, we demonstrated the utility of generative modeling Tera-MIND for vasculature reconstruction using spatial transcriptomics data. This multi-modal approach provides an ethically responsible and cost-effective platform for probing vascular heterogeneity and thereby paves the way for systematic, high-throughput analysis of vascular dynamics across diverse biological contexts.

Materials and Methods

Tera-MIND generative framework

The Tera-MIND model is a novel generative framework designed for tera-scale biomedical data. It leverages paired spatial transcriptomics arrays and co-registered histological images to simulate mouse brains in space. Tera-MIND employs a patch-wise and boundary-aware training strategy, enabling the model to learn from smaller sub-volumes. This design allows for efficient training and inference on high-volume datasets, making whole-brain reconstruction at cell-level resolution computationally feasible under mild hardware requirements[3].

Spatial transcriptomics data

Recent efforts by a network of researchers has resulted in the generation of a detailed cellular atlas of the mouse brain[7,9–11]. The mouse brain data set of coronal sections with high-resolution images of DAPI (4',6-diamidino-2-phenylindole) and Poly-T (mRNA) stained tissue with the corresponding spatially indexed transcriptomic readout used for development of Tera-MIND was published by Yao *et al*.[7]. Training of Tera-MIND on this data set is described in more detail in the previous study[3]. In short, by training a patch-based and boundary-aware diffusion model on gene expression input and corresponding histological images, it predicts the spatial associations between molecular gene expression markers and morphology.

As described previously[3], data sets for three different mouse brains were available[7,9]. The primary generation results are based on the data set of a female mouse brain (P56), and the other two available data sets (one female and one male mouse brain, P56) were used as training data. We refer interested readers to the previous study[3] for a more detailed comparison of ground truth and generated images. Here, we include one example of a ground truth (Fig. 1C) and a corresponding generated (Fig. 1D) DAPI image of a coronal mouse brain section.

Validation and analysis

After training the model based on the two training data sets with corresponding transcriptomic and histological data, the model of the brain vasculature based on *Acta2* and *Cldn5* expression was generated on the unseen third data set (Fig. 1A and 1B, model training and test). The reconstructed

vasculature was validated against available imaging data and known anatomical landmarks from the Xiong *et al*. study[6].

Data Availability

Spatial transcriptomic and histological data (DAPI and Poly-T) for all three mouse brain atlases can be accessed via https://doi.org/10.35077/g.610.
The processed spatial transcriptomic and histological data can be accessed via https://doi.org/10.35077/g.1176.
The code and details on model training and testing are made available at the repository github.com/CTPLab/Tera-MIND.

Acknowledgments

This study was funded by the Promedica Foundation (F-87701-41-01) and by core funding to the Computational and Translational Pathology Lab led by V.H.K at the University of Basel and University of Zurich.

References

1. Kirst, C. *et al.* Mapping the Fine-Scale Organization and Plasticity of the Brain Vasculature. *Cell* **180**, 780-795.e25 (2020).
2. Wu, J. & Koelzer, V. H. Towards generative digital twins in biomedical research. *Comput. Struct. Biotechnol. J.* **23**, 3481–3488 (2024).
3. Wu, J., Berg, I., Li, Y., Konukoglu, E. & Koelzer, V. H. Tera-MIND: Tera-scale mouse brain simulation via spatial mRNA-guided diffusion. Preprint at https://doi.org/10.48550/ARXIV.2503.01220 (2025).
4. Greene, C., Hanley, N. & Campbell, M. Claudin-5: gatekeeper of neurological function. *Fluids Barriers CNS* **16**, 3 (2019).
5. Miyawaki, T. *et al.* Visualization and molecular characterization of whole-brain vascular networks with capillary resolution. *Nat. Commun.* **11**, 1104 (2020).
6. Xiong, B. *et al.* Precise Cerebral Vascular Atlas in Stereotaxic Coordinates of Whole Mouse Brain. *Front. Neuroanat.* **11**, 128 (2017).
7. Yao, Z. *et al.* A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **624**, 317–332 (2023).
8. Túrós, D. *et al.* Deep learning-based 3D spatial transcriptomics with X-Pression. Preprint at https://doi.org/10.1101/2025.03.21.644627 (2025).
9. Zhang, M. *et al.* Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature* **624**, 343–354 (2023).
10. Shi, H. *et al.* Spatial atlas of the mouse central nervous system at molecular resolution. *Nature* **622**, 552–561 (2023).
11. Langlieb, J. *et al.* The molecular cytoarchitecture of the adult mouse brain. *Nature* **624**, 333–342 (2023).

Figure

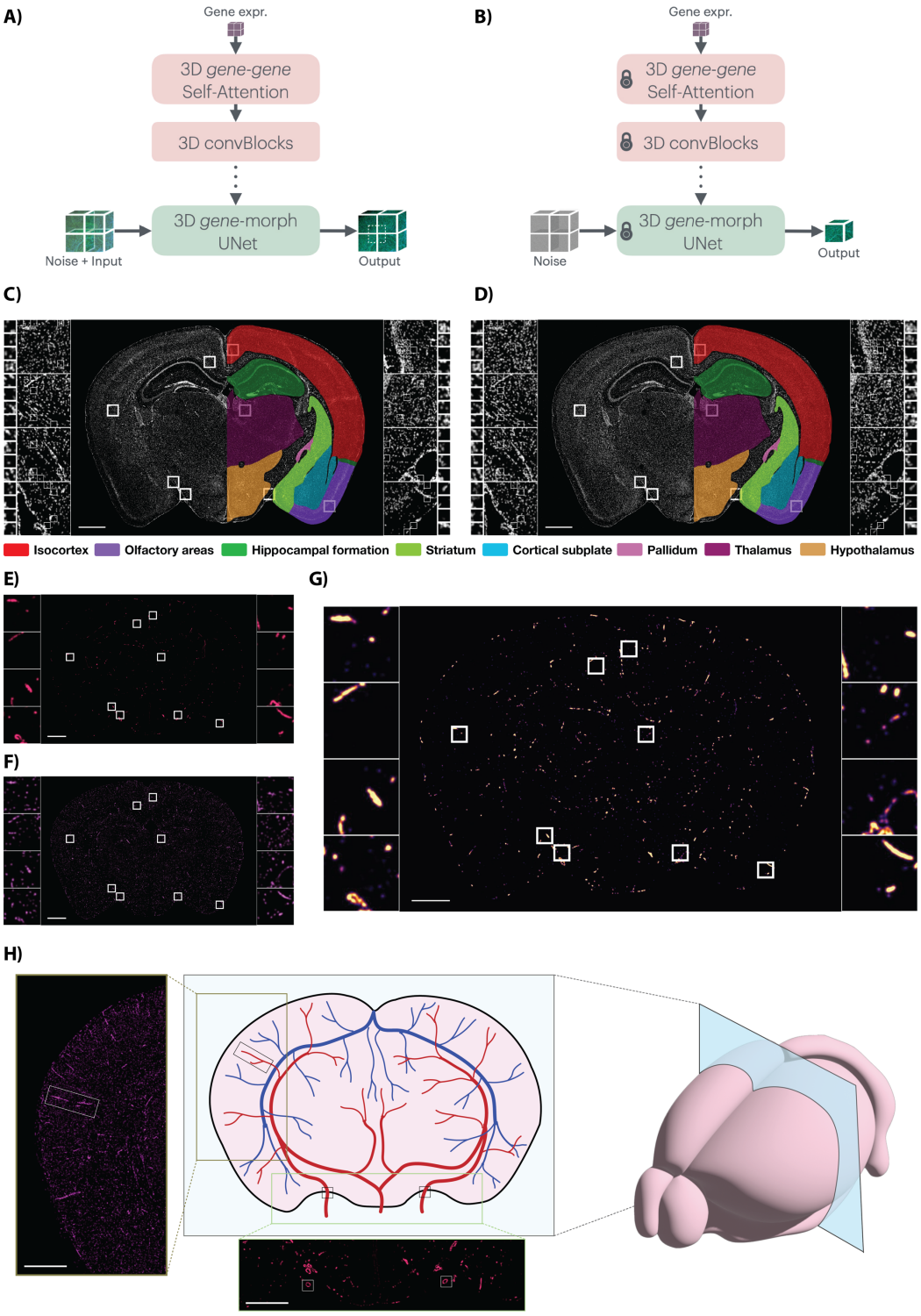
A)
Gene expr.
3D gene-gene Self-Attention
3D convBlocks
3D gene-morph UNet
Noise + Input
Output
B)
Gene expr.
3D gene-gene Self-Attention
3D convBlocks
3D gene-morph UNet
Noise
Output
C)
D)
Isocortex
Olfactory areas
Hippocampal formation
Striatum
Cortical subplate
Pallidum
Thalamus
Hypothalamus
E)
F)
G)
H)

Figure 1. A) Illustration of patch-based diffusion model training. **B)** Illustration of model test/inference. **C)** Image showing the ground truth of a DAPI staining of a coronal section of the mouse brain, including color labels of major brain regions. **D)** Image showing the generated version of the DAPI staining, including color labels of major brain regions. **E)** Generated image showing *Acta2* expression in a coronal brain section. **F)** Generated image showing *Cldn5* expression in a coronal brain section. **G)** Visualization of *Cldn5*-*Acta2* attention. **H)** Schematic illustration of vasculature distribution, including examples of *Cldn5 (left panel)* and *Acta2 (bottom panel)* expression (merged image of 4 consecutive slides) with corresponding structures highlighted with black and white boxes. **C-H)** Length of scale bar (white, bottom left corner) is 1mm.