BioFlow: A Biologically Valid Support-Preserving Flow for Histology-Conditioned Spatial Transcriptomics Prediction

Haoran Xu[1], Yang Liu[2], Wei Yuan[1], and Xiao Han[1⋆]

[1] College of Biomedical Engineering, Sichuan University, Chengdu, China
xhr2023@stu.scu.edu.cn
yuanw@stu.scu.edu.cn
xiao_han@scu.edu.cn

[2] Department of Computer Science, Hong Kong Baptist University, Hong Kong SAR, China
yangliu1214@hkbu.edu.hk

Abstract. Spatial transcriptomics (ST) prediction from histology requires modeling sparse, zero-inflated, and intrinsically non-negative gene expression. However, existing generative models evolve in unconstrained real-valued spaces, allowing trajectories to leave the biological valid support of count-based data and degrading spatial structure modeling. We propose BioFlow, a support-preserving flow-matching framework that embeds non-negativity directly into the generative dynamics via velocity reparameterization. By constraining the generative dynamics rather than correcting invalid outputs post hoc, BioFlow maintains trajectory-level support consistency and aligns generation with the intrinsic geometry of sparse expression data. Across multiple public ST datasets, BioFlow achieves state-of-the-art spatial correlation performance while improving training stability and computational efficiency. These results underscore the importance of support-consistent generative dynamics for reliable modeling of sparse biomedical count data. Our source code is available at https://github.com/hrterry/BioFLow.

Keywords: Spatial Transcriptomics · Generative Model · Constrained Flow Matching.

1 Introduction

Spatial transcriptomics (ST) [18] enables spatially resolved molecular profiling of tissue sections, preserving both gene expression and spatial context [5,20,11,3,17,19]. By linking molecular patterns with histopathological morphology [2], ST provides clinically meaningful insight into tumor heterogeneity and tissue organization. However, high-throughput ST sequencing remains costly and limited in availability, whereas hematoxylin and eosin (H&E) whole-slide images (WSIs)

⋆ Corresponding author: xiao_han@scu.edu.cn

are routinely acquired in clinical workflows. Predicting spatial gene expression directly from WSIs therefore presents a scalable alternative for molecular inference [19,8,15,21]. For such predictions to be reliable, models must respect not only spatial dependencies but also the intrinsic geometric structure of transcriptomic data.

Existing approaches to histology-conditioned gene expression prediction include regression-based models (e.g., ST-Net [8], HisToGene [15], TRIPLEX [4]), graph-based spatial modeling (e.g., HisToST [22], MERGE [7]), and more recently generative approaches such as diffusion [9] and flow matching [13] (e.g., STFlow [10], STEM [23]). However, recent generative models typically evolve in unconstrained Euclidean spaces and may generate negative intermediate states, without explicitly respecting the admissible domain of gene expression.

This mismatch is particularly relevant for zero-inflated ST expression data, where many values lie near the boundary of the non-negative orthant. Unconstrained trajectories may enter invalid regions, causing the learned vector field to be optimized on biologically implausible states and thereby biasing trajectory evolution. Existing remedies enforce validity only at the distributional [10] or output level [4]: positive-support priors do not prevent off-support trajectories, while post-hoc projection modifies only the final prediction and cannot fully recover spatial pattern fidelity after the trajectory has evolved (Fig. 1b). These observations motivate embedding support constraints directly into the generative dynamics.

To address this issue, we propose **BioFlow**, a support-preserving flow-matching framework that reparameterizes the velocity field to keep continuous-time trajectories within the non-negative expression domain. Rather than correcting invalid outputs after generation, BioFlow constrains the evolution itself. We implement the constrained dynamics with a lightweight token-based Transformer [16,6] that captures spatial interactions. Across multiple public ST datasets, BioFlow achieves state-of-the-art spatial correlation with favorable accuracy-efficiency trade-offs.

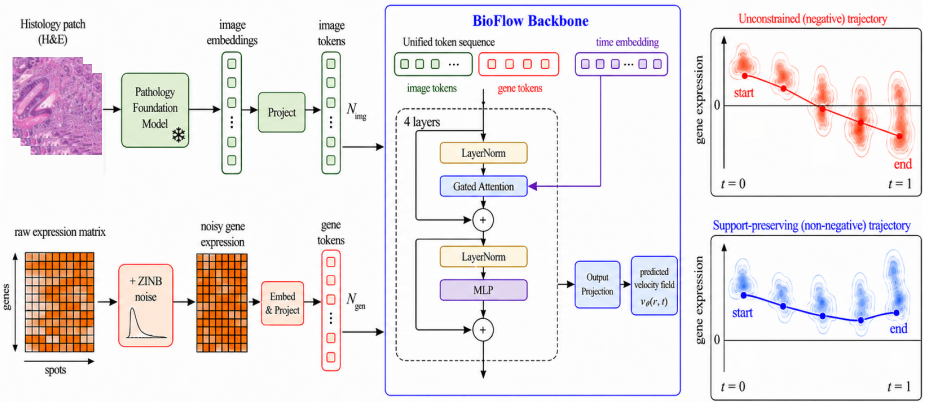
2 Methods

2.1 Problem Formulation

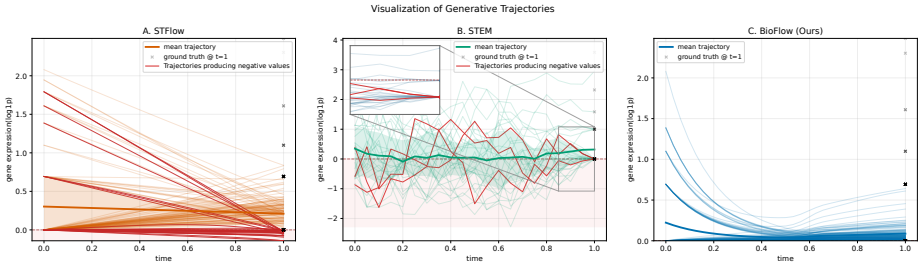
We consider spot-level paired data $\mathcal{D} = \{(I_i, G_i)\}_{i=1}^{N}$, from a tissue slide with N spots, where $I_i \in \mathbb{R}^{H\times W\times 3}$ denotes a WSI histology patch of spatial size $H \times W$ and $G_i \in \mathbb{R}^{M}_{\geq 0}$ is the corresponding gene expression vector with M genes. Our goal is to learn the conditional distribution $p_\theta(G \mid I)$.

2.2 Conditional Flow Matching

To model $p_\theta(G \mid I)$, we adopt Conditional Flow Matching (CFM) [13], which parameterizes generation through a time-dependent velocity field $v_\theta(x, t \mid I)$ in the expression space. Rather than directly regressing the final expression vector,



(a) Architecture of BioFlow.



(b) Empirical generative trajectories across methods.

Fig. 1: **Overview of BioFlow.** (a) BioFlow predicts the velocity field from histology features, noisy gene-expression states, and time embeddings, while preserving non-negative sampled trajectories. (b) Empirical trajectories in log1p expression space show that flow matching (STFlow) and diffusion (STEM) may enter the negative region during sampling, whereas support-preserving flow matching (BioFlow) keeps sampled trajectories within the non-negative domain.

CFM learns dynamics that transport samples from a base distribution $p_0(x)$ to the target distribution $p_1(x \mid I)$ along a continuous trajectory $\{x_t\}_{t\in[0,1]}$. Given a training pair $(I, G) \sim \mathcal{D}$, we denote the endpoint as $x_1 \triangleq G$ and sample $x_0 \sim p_0(x)$ from a ZINB prior. A linear interpolation path is constructed as $x_t = (1-t)x_0 + tx_1, t \sim \text{Uniform}(0,1)$. The corresponding target velocity is $u^*(x_0, x_1) = \frac{dx_t}{dt} = x_1 - x_0$. The velocity network v_θ is trained by minimizing

$$\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_{x_0,x_1,t}\left[\|v_\theta(x_t, t \mid I) - u^*(x_0, x_1)\|_2^2\right], \tag{1}$$

where θ denotes the learnable parameters of the vector field v_θ. At inference, gene expression is generated by numerically integrating the learned flow:

$$x_{t=1} = \Phi_\theta(x_0, I) = x_0 + \int_0^1 v_\theta(x_t, t \mid I)\, dt. \tag{2}$$

Although the interpolation path is non-negative when both endpoints are non-negative, the learned velocity field is unconstrained. During sampling, numerical integration of an unconstrained field may therefore move intermediate states outside the non-negative admissible domain.

2.3 Support-preserving Flow Dynamics

Gene expression values are inherently non-negative. Rather than enforcing this property only at the final output, BioFlow constrains the generative dynamics so that the non-negative orthant is invariant along the trajectory: $x_t \in \mathbb{R}^M_{\geq 0}, \forall t \in [0,1]$. Specifically, we reparameterize the velocity field as

$$\hat{v}_\theta\left(x_t, t\right) = -\frac{x_t}{\Delta} + \phi\left(x_t, t\right), \quad \phi(\cdot) \geq 0, \tag{3}$$

where $\Delta > 0$ is a decay time scale and ϕ_θ is implemented with a softplus activation. The decay term allows expression values to decrease in the interior of the non-negative domain, while the non-negative residual term restores expression values according to the histology-conditioned dynamics.

This parameterization yields a continuous-time non-negativity guarantee. For each gene dimension j, the dynamics satisfy

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\Delta} + \phi_{\theta,j}(x(t), t \mid I). \tag{4}$$

Solving this linear differential equation gives

$$x_j(t) = e^{-t/\Delta} x_j(0) + \int_0^t e^{-(t-s)/\Delta} \phi_{\theta,j}(x(s), s \mid I)\, ds \geq 0. \tag{5}$$

Therefore, if $x(0) \in \mathbb{R}^M_{\geq 0}$, then $x(t) \in \mathbb{R}^M_{\geq 0}$ for all $t \in [0,1]$ under the continuous BioFlow dynamics. In practice, we use forward Euler integration with step size $h > 0$ during sampling

$$x_{t+h} = x_t + h\hat{v}_\theta(x_t, t \mid I) = \left(1 - \frac{h}{\Delta}\right) x_t + h\phi_\theta(x_t, t \mid I). \tag{6}$$

When $0 < h \leq \Delta$, both coefficients are non-negative, and thus $x_t \geq 0$ implies $x_{t+h} \geq 0$. This gives a discretization guarantee for the Euler sampler used in our implementation. We set $h = 0.05$ and $\Delta = 0.1$ in all experiments, satisfying the positivity condition.

BioFlow is trained with the same CFM objective, replacing the unconstrained velocity v_θ with the constrained velocity $\hat{v}_\theta$

$$\mathcal{L}_{\text{BioFlow}}(\theta) = \mathbb{E}_{x_0, x_1, t}\left[\left\|\hat{v}_\theta(x_t, t \mid I) - u^*(x_0, x_1)\right\|_2^2\right]. \tag{7}$$

Unlike post-hoc clamping or output-level positivity transforms, BioFlow constrains the vector field itself. This makes non-negativity a property of the trajectory dynamics rather than a correction applied after invalid states have occurred. The constraint preserves the non-negative admissible domain of expression values, while exact sparsity or zero-pattern preservation is not explicitly enforced.

Table 1: Comparison of spatial gene expression prediction across datasets. PCC(H/M/All) denote Pearson correlations computed over the high-, medium-, and all predicted genes, respectively. Subscript numbers denote standard deviation for PCC(All). BioFlow outperforms competing methods across all datasets and evaluation levels.

Method	HER2ST			PRAD			READ		
	PCC(H)↑	PCC(M)↑	PCC(All)↑	PCC(H)↑	PCC(M)↑	PCC(All)↑	PCC(H)↑	PCC(M)↑	PCC(All)↑
MERGE	0.495	0.361	$0.209_{.104}$	0.447	0.386	$0.295_{.032}$	0.336	0.301	$0.150_{.067}$
TRIPLEX	0.467	0.349	$0.177_{.089}$	0.535	0.461	$0.267_{.055}$	0.272	0.180	$0.006_{.017}$
STFlow	0.451	0.340	$0.226_{.183}$	0.503	0.430	$0.289_{.013}$	0.457	0.363	$0.199_{.053}$
BioFlow	**0.512**	**0.412**	$\mathbf{0.287}_{.207}$	**0.571**	**0.529**	$\mathbf{0.401}_{.004}$	**0.517**	**0.420**	$\mathbf{0.254}_{.041}$

2.4 Velocity Field Prediction Network

We parameterize the conditional velocity field with a transformer backbone. As illustrated in Fig. 1a, histology patches are first encoded using a pathology foundation model (UNI) to obtain patch-level embeddings. These embeddings are projected into image tokens and concatenated with noisy gene tokens, forming a unified time-conditioned token sequence. Gated-attention [14] enables global token interaction, capturing spatial dependencies without predefined graphs. The output head maps gene-token representations to per-gene velocity predictions at each time step, parameterizing $\phi_\theta(x_t, t, I)$ in our support-preserving flow. The architecture can be replaced by other expressive backbones without altering the formulation.

3 Experiments and Results

Experimental Setup. We evaluate BioFlow on HER2-positive breast cancer (HER2ST) [1], prostate adenocarcinoma (PRAD), and rectal adenocarcinoma (READ) from HEST-1k [12], using patient-wise splits; dataset statistics are reported in Table 3. Genes are selected from each training fold by intersecting the top-ranked genes by mean expression and variance, yielding 431 genes for HER2ST and 50 genes for PRAD and READ. We report Pearson correlation coefficient (PCC) across spots for high-ranked, medium-ranked, and all selected genes, denoted PCC(H), PCC(M), and PCC(All), respectively. We compare BioFlow with STFlow, MERGE, TRIPLEX, and STEM. STFlow and TRIPLEX use the same UNI encoder as BioFlow, whereas MERGE follows its original feature pipeline; STEM is evaluated separately due to its computational cost.

3.1 Quantitative Results

Overall Performance. Table 1 summarizes quantitative results across three datasets. BioFlow achieves the highest PCC across all gene strata and datasets. While baseline methods exhibit noticeable degradation from PCC(H) to PCC(All), BioFlow maintains more stable accuracy under increasing difficulty. In terms of

Table 2: Comparison with diffusion-based STEM on PRAD and HER2ST. Results follow the original STEM setup (UNI+CONCH encoder, preprocessing, gene panels, PCC-10/50/200 for PRAD and PCC-10/100/400 for HER2ST). BioFlow attains higher accuracy with significantly lower computation.

	PRAD				HER2ST		
Method	**GPUHours ↓**	**PCC-10 ↑**	**PCC-50 ↑**	**PCC-200 ↑**	**PCC-10 ↑**	**PCC-100 ↑**	**PCC-400 ↑**
STEM	4000	0.610	0.532	0.383	0.829	0.772	0.598
BioFlow	**0.3**	**0.660**	**0.620**	**0.526**	**0.872**	**0.815**	**0.732**

PCC(All), BioFlow improves over competing methods by +27–62% on HER2ST, +36–50% on PRAD, and +27–32% on READ, indicating improved robustness in sparse and heterogeneous regimes.

Accuracy-Efficiency Trade-off. Fig. 2 analyzes predictive performance versus training cost, measured on the NVIDIA L40S (48GB) GPU under identical hardware settings. BioFlow attains superior PCC with substantially lower cost. On HER2ST, BioFlow requires 122 minutes, compared to 396 and 631 minutes for MERGE and STFlow, and 2800 minutes for TRIPLEX (3–23× faster). Similar trends are observed on PRAD and READ (18 and 7 minutes, respectively), demonstrating favorable scalability across dataset sizes.

Comparison with Diffusion Models. We additionally compare BioFlow with the diffusion-based model STEM (Table 2). Due to iterative denoising, STEM requires substantial computational overhead, approximately 4,000 GPU hours on PRAD (Nvidia A100 GPU), whereas BioFlow completes training in 18 minutes on the same hardware. For a fair and reproducible comparison, we adopt the official STEM configuration with aligned encoder (UNI+CONCH) and preprocessing. Under this matched setup, BioFlow achieves higher predictive accuracy across gene strata while requiring orders-of-magnitude less computation.

Visualization of Marker Gene Expression. Fig. 3 visualizes *AMD1* expression in PRAD, where BioFlow best recovers the ground-truth spatial pattern and attains the highest gene-level PCC (0.639).

3.2 Ablation Study

Positivity-by-Design Fixes. We evaluate whether simple output-level positivity corrections are sufficient. Using STFlow as a controlled backbone, we compare post-hoc clamping and softplus reparameterization on PRAD (Table 4). Although clamping reduces MSE by masking negative deviations in zero-inflated data, it substantially degrades spatial structure, with PCC(H) dropping from 0.503 to 0.329 ($\approx 35\%$). Softplus exhibits similar degradation. In contrast, BioFlow achieves the highest PCC and lowest MAE without such trade-offs. These results indicate that output-level positivity fixes cannot recover structural distortions introduced during trajectory evolution, and that enforcing support constraints at the dynamics level is necessary for stable spatial modeling.

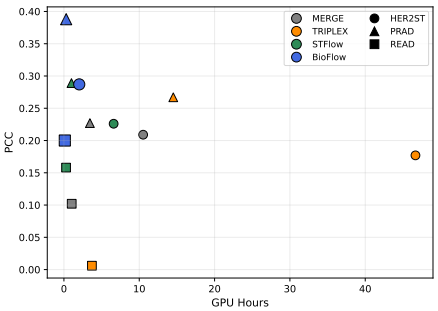


Fig. 2: PCC(All) vs. training GPU hours for MERGE, TRIPLEX, STFlow, and BioFlow on HER2ST, PRAD, and READ.

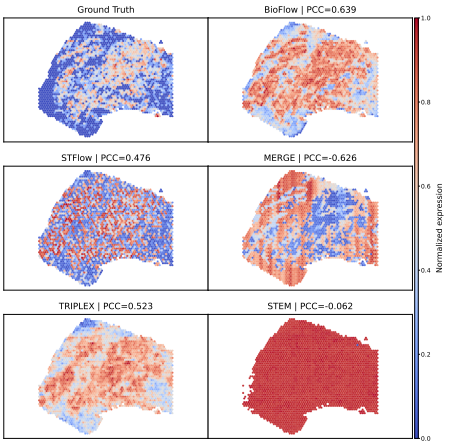


Fig. 3: Representative spatial visualization of AMD1 expression in PRAD dataset.

Table 3: Dataset statistics. Split: patient-wise train-test folds; Samples: total WSI-ST pairs; AvgSpots: mean number of spatial spots per slide.

Dataset	avgspots	Patients	Samples	Genes	Splits
HER2ST	378	8	24	431	8
PRAD	2726	2	23	50	2
READ	1480	2	4	50	2

Table 4: Effect of output-level positivity fixes on PRAD. STFlow is evaluated with post-hoc clamping and softplus reparameterization. PCC(H/M/All), MSE, and MAE are reported. TRIPLEX uses built-in output clamping.

Method	PCC(H)↑	PCC(M)↑	PCC(L)↑	MSE↓	MAE↓
TRIPLEX (clamp)	0.535	0.461	0.267	4.934	1.594
STFlow	0.503	0.430	0.289	1.935	1.265
STFlow + clamp	0.329	0.283	0.196	**1.532**	1.037
STFlow + softplus	0.308	0.260	0.172	1.730	1.022
BioFlow	**0.571**	**0.529**	**0.401**	1.645	**0.995**

Non-negative Constraint. Enforcing the support constraint improves both efficiency and accuracy (Table 5), reducing training time by 20–38% and increasing PCC across gene subsets. Moreover, PCC remains stable across a range of integration settings satisfying $0 < h < \Delta$. Taken together, these results indicate that support-preserving dynamics enhance stability and generalization without architectural changes.

Velocity Prediction Network. To disentangle architectural capacity from support preservation, we replace the STFlow velocity network with our Transformer design while disabling the support constraint for fairness. This modification yields higher PCC across gene strata, particularly on PCC(All) (Table 6), indicating stronger spatial modeling capacity.

Sensitivity to Δ and h. Table 7 shows that BioFlow maintains comparable performance over feasible parameter pairs satisfying $h \leq \Delta$, whereas violating

Table 5: Ablation on the support constraint. w/o constraint removes the non-negative support formulation and predicts the velocity field directly, while non-neg enforces the support preservation. All other settings follow Table 1.

Dataset	Method	PCC(H)↑	PCC(M)↑	PCC(All)↑	Time (min)↓
HER2ST	w/o Constraint	0.506	0.386	0.260	136
	Non-neg.	**0.512**	**0.412**	**0.287**	**109**
PRAD	w/o Constraint	0.505	0.469	0.360	29
	Non-neg.	**0.571**	**0.529**	**0.401**	**18**

Table 6: Velocity prediction network ablation. Under identical settings (support constraint disabled), BioFlow achieves higher PCC than STFlow across datasets.

Dataset	Model	PCC(H)↑	PCC(M)↑	PCC(All)↑
HER2ST	STFlow	0.453	0.342	0.227
	BioFlow	**0.506**	**0.386**	**0.260**
PRAD	STFlow	0.503	0.430	0.289
	BioFlow	**0.505**	**0.469**	**0.360**

Table 7: Sensitivity of PCC(All) on READ to Δ and Euler step size h. Bold entries violate the positivity condition $h \leq \Delta$.

$\Delta \backslash h$	0.02 (50 steps)	0.05 (20 steps)	0.10 (10 steps)
0.20	0.2355	0.2354	0.2393
0.10	0.2555	0.2544	0.2565
0.05	0.2469	0.2557	**0.1623**
0.02	0.2370	**0.0111**	**−0.0020**

this condition causes substantial degradation. We use $\Delta = 0.1$ and $h = 0.05$ in all experiments.

4 Conclusion

We presented BioFlow, a support-preserving flow-matching framework that enforces non-negative constraints at the level of generative dynamics rather than post hoc correction. By aligning trajectory evolution with the intrinsic support geometry of sparse count data, BioFlow improves structural fidelity, stability, and computational efficiency. Beyond spatial transcriptomics, our formulation highlights the importance of geometry-aware generative dynamics for modeling constrained biomedical data.

Acknowledgments. This work was supported by the Fundamental Research Funds for the Central Universities (1082204112M91) and the Sichuan Provincial Natural Science Foundation (Grant No. 2026NSFSC0661).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Andersson, A., Larsson, L., Stenbeck, L., Salmén, F., Ehinger, A., Wu, S.Z., Al-Eryani, G., Roden, D., Swarbrick, A., Borg, Å., et al.: Spatial deconvolution of her2-

positive breast cancer delineates tumor-associated cell type interactions. Nature communications **12**(1), 6012 (2021)
2. Bunne, C., Roohani, Y., Rosen, Y., Gupta, A., Zhang, X., Roed, M., Alexandrov, T., AlQuraishi, M., Brennan, P., Burkhardt, D.B., et al.: How to build the virtual cell with artificial intelligence: Priorities and opportunities. Cell **187**(25), 7045–7063 (2024)
3. Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., Zhuang, X.: Spatially resolved, highly multiplexed rna profiling in single cells. Science **348**(6233), aaa6090 (2015)
4. Chung, Y., Ha, J.H., Im, K.C., Lee, J.S.: Accurate spatial gene expression prediction by integrating multi-resolution features. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11591–11600 (2024)
5. Eng, C.H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.C., et al.: Transcriptome-scale super-resolved imaging in tissues by rna seqfish+. Nature **568**(7751), 235–239 (2019)
6. Esser, P., Kulal, S., Blattmann, A., Entezari, R., Müller, J., Saini, H., Levi, Y., Lorenz, D., Sauer, A., Boesel, F., et al.: Scaling rectified flow transformers for high-resolution image synthesis. In: Forty-first international conference on machine learning (2024)
7. Ganguly, A., Chatterjee, D., Huang, W., Zhang, J., Yurovsky, A., Johnson, T.S., Chen, C.: Merge: Multi-faceted hierarchical graph-based gnn for gene expression prediction from whole slide histopathology images. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 15611–15620 (2025)
8. He, B., Bergenstråhle, L., Stenbeck, L., Abid, A., Andersson, A., Borg, Å., Maaskola, J., Lundeberg, J., Zou, J.: Integrating spatial gene expression and breast tumour morphology via deep learning. Nature biomedical engineering **4**(8), 827–834 (2020)
9. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)
10. Huang, T., Liu, T., Babadi, M., Jin, W., Ying, R.: Scalable generation of spatial transcriptomics from histology images via whole-slide flow matching. arXiv preprint arXiv:2506.05361 (2025)
11. Janesick, A., Shelansky, R., Gottscho, A., Wagner, F., Rouault, M., Beliakoff, G., de Oliveira, M.F., Kohlway, A., Abousoud, J., Morrison, C., et al.: High resolution mapping of the breast cancer tumor microenvironment using integrated single cell, spatial and in situ analysis of ffpe tissue. Biorxiv pp. 2022–10 (2022)
12. Jaume, G., Doucet, P., Song, A., Lu, M.Y., Almagro Pérez, C., Wagner, S., Vaidya, A., Chen, R., Williamson, D., Kim, A., et al.: Hest-1k: A dataset for spatial transcriptomics and histology image analysis. Advances in Neural Information Processing Systems **37**, 53798–53833 (2024)
13. Lipman, Y., Chen, R.T., Ben-Hamu, H., Nickel, M., Le, M.: Flow matching for generative modeling. arXiv preprint arXiv:2210.02747 (2022)
14. Ma, N., Goldstein, M., Albergo, M.S., Boffi, N.M., Vanden-Eijnden, E., Xie, S.: Sit: Exploring flow and diffusion-based generative models with scalable interpolant transformers. In: European Conference on Computer Vision. pp. 23–40. Springer (2024)
15. Pang, M., Su, K., Li, M.: Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. BioRxiv pp. 2021–11 (2021)
16. Peebles, W., Xie, S.: Scalable diffusion models with transformers. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 4195–4205 (2023)

17. Raj, A., Van Den Bogaard, P., Rifkin, S.A., Van Oudenaarden, A., Tyagi, S.: Imaging individual mrna molecules using multiple singly labeled probes. Nature methods **5**(10), 877–879 (2008)
18. Rao, A., Barkley, D., França, G.S., Yanai, I.: Exploring tissue architecture using spatial transcriptomics. Nature **596**(7871), 211–220 (2021)
19. Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al.: Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science **353**(6294), 78–82 (2016)
20. Wang, X., Allen, W.E., Wright, M.A., Sylwestrak, E.L., Samusik, N., Vesuna, S., Evans, K., Liu, C., Ramakrishnan, C., Liu, J., et al.: Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science **361**(6400), eaat5691 (2018)
21. Xie, R., Pang, K., Chung, S., Perciani, C., MacParland, S., Wang, B., Bader, G.: Spatially resolved gene expression prediction from histology images via bi-modal contrastive learning. Advances in Neural Information Processing Systems **36**, 70626–70637 (2023)
22. Zeng, Y., Wei, Z., Yu, W., Yin, R., Yuan, Y., Li, B., Tang, Z., Lu, Y., Yang, Y.: Spatial transcriptomics prediction from histology jointly through transformer and graph neural networks. Briefings in Bioinformatics **23**(5) (2022)
23. Zhu, S., Zhu, Y., Tao, M., Qiu, P.: Diffusion generative modeling for spatially resolved gene expression inference from histology images. arXiv preprint arXiv:2501.15598 (2025)