

## Supplementary Materials for

# SpaDAMA: Spatial Transcriptomics Deconvolution Using Domain-Adversarial Masked Autoencoder

### Supplementary Notes

#### 1. Evaluation metrics

We use Pearson correlation coefficient (PCC), structural similarity index measure (SSIM), root mean square error (RMSE), Jensen-Shannon divergence (JS/JSD), Accuracy score (AS) and Area Under the Receiver Operating Characteristic Curve (AUC) to evaluate the proposed method against baselines.

$$\text{PCC} = \frac{\text{cov}(x_i, \hat{x}_i)}{\sigma_i \hat{\sigma}_i} \quad (1)$$

where  $x_i$  is the ground truth cell type composition for cell type  $i$ ,  $\sigma_i$  is its standard deviation, and  $\hat{x}_i$  and  $\hat{\sigma}_i$  are the predicted values.

$$\text{SSIM} = \frac{(2\hat{u}_i u_i + C_1)(2\text{cov}(x_i, \hat{x}_i) + C_2)}{(\hat{u}_i^2 + u_i^2 + C_1)(\hat{\sigma}_i^2 + \sigma_i^2 + C_2)} \quad (2)$$

where  $\mu_i$  is the average ground truth cell type composition for cell type  $i$ ,  $\hat{\mu}_i$  is the predicted average, and  $C_1$  and  $C_2$  are constants set to 0.01 and 0.03, respectively.

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{j=1}^M (x_{ij} - \hat{x}_{ij})^2} \quad (3)$$

where  $x_{ij}$  is the ground truth cell type composition for cell type  $i$  in spot  $j$ , and  $\hat{x}_{ij}$  is the predicted value.

$$\text{JS} = \frac{1}{2} \text{KL} \left( P_i, \frac{\hat{P}_i + P_i}{2} \right) + \frac{1}{2} \text{KL} \left( \hat{P}_i, \frac{\hat{P}_i + P_i}{2} \right) \quad (4)$$

where  $P_i$  and  $\hat{P}_i$  are the spatial distributions of cell type  $i$  in the ground truth and prediction, respectively.

$$\text{AS} = \frac{1}{4} (\text{RANK}_{\text{PCC}} + \text{RANK}_{\text{SSIM}} + \text{RANK}_{\text{RMSE}} + \text{RANK}_{\text{JS}}) \quad (5)$$

where the average PCC/SSIM and RMSE/JS of all deconvolution methods are ranked in ascending and descending order, respectively, to obtain  $\text{RANK}_{\text{PCC}}$ ,  $\text{RANK}_{\text{SSIM}}$ ,  $\text{RANK}_{\text{RMSE}}$ , and  $\text{RANK}_{\text{JS}}$ .

$$\text{AUC} = \int_0^1 \text{TPR}(t) d\text{FPR}(t) \quad (6)$$

where  $\text{TPR}(t)$  is the True Positive Rate and  $\text{FPR}(t)$  is the False Positive Rate at a given threshold  $t$ . The area under the curve (AUC) quantifies the model's ability to distinguish between positive and negative classes.

#### 2. Baselines

To assess the performance of SpaDAMA, we compared it with eight other state-of-the-art methods. The methods compared include the following:

- Spoint [1]: Spoint utilizes deep learning for feature extraction in spatial transcriptomics data and applies PCA for dimensionality reduction, enhancing cell type deconvolution accuracy.
- Tangram [2]: Tangram employs deep learning to map scRNA-seq data onto ST data, predicting the spatial distribution of cell types.
- Cell2location [3]: Cell2location performs cell-type deconvolution of ST data using a hierarchical Bayesian framework to achieve accurate results.
- DestVI [4]: DestVI utilizes variational inference and latent variable models to estimate the proportions of different cell types.
- scpDeconv [5]: scpDeconv combines scRNA-seq data with proteomic data using a domain adversarial autoencoder to improve the precision of identifying cell types.
- Stereoscope [6]: Stereoscope integrates scRNA-seq and ST data to spatially map cell types by leveraging probabilistic modeling techniques.
- SpaOTsc [7]: SpaOTsc leverages optimal transport theory to match single-cell RNA-seq data with spatial transcriptomics (ST) data for accurate cell-type deconvolution.
- NovoSpaRc [8]: NovoSpaRc also employs optimal transport to link single-cell RNA-seq data with spatial transcriptomics, but it focuses on reconstructing the spatial distribution of cell types across the tissue.

### 3. Ablation studies

To evaluate the contributions of the masking and adversarial modules to the SpaDAMA model, we conducted a series of ablation experiments (Table S2 and Figure S1). We specifically aimed to determine if these modules enhance model performance and whether reconstruction loss should be computed on the masked region or the entire ST data, as well as the impact of different masking rates. The experiments included a baseline model with only the encoder and predictor, a model with a masking module (reconstruction loss computed only for the masked region), a model with full reconstruction loss, and a model with an adversarial module, with masking rates  $\rho = 0.2, 0.3, 0.4$ , and  $0.5$ . The ablation experiments confirm the critical role of the masking and adversarial modules in enhancing the accuracy of cell type proportion predictions, with the impact of the masking rate on the model's performance being relatively minor. In particular, the model performed best when the masking rate was  $\rho = 0.3$ .

### 4. Downstream analysis

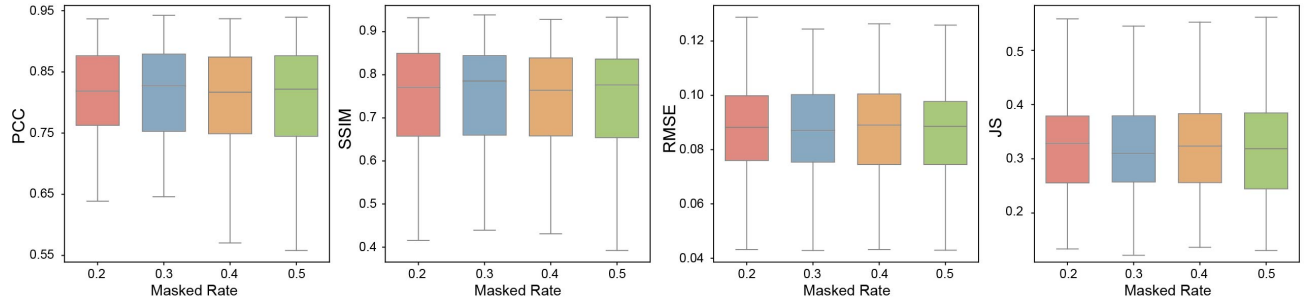
We performed downstream analysis on the deconvolution results of the real dataset. In the zebrafish embryo dataset, certain cell types exhibit spatial colocalization (Figure S4A). For instance, Erythroid Lineage Cells and Pronephros Cells both originate from the mesoderm during early embryonic development [9, 10]. These two cell types are located in the lower regions of the embryo and are positioned adjacent to each other. Erythroid lineage cells contribute to early hematopoiesis, while pronephros cells play a crucial role in the development of the initial kidney structures. Their spatial proximity likely reflects the interaction and coordination of mesodermal tissues during embryogenesis. Additionally, Hatching Gland Cells are situated in the head region of the embryo, closely associated with the distribution of Integument Cells [11, 12]. As the embryo develops, integument cells form a protective barrier on the surface, while hatching gland cells secrete enzymes to facilitate the hatching process. The spatial proximity and complementary functions of these two cell types suggest that they may coordinate and regulate each other's activity during the hatching process [13].

The cell composition predicted by SpaDAMA clearly demonstrates the synergistic interactions between different cell types during immune responses in the mouse lymph node dataset (Figure S4B). For example, the interaction between the regulatory immune cells Tregs and Migratory DCs plays a crucial role in immune regulation [14]. Specifically, Tregs inhibit the expression of co-stimulatory molecules (such as CD80/CD86) on dendritic cells, thereby reducing their ability to activate effector T cells. This effectively modulates the intensity of immune responses, preventing excessive activation of the immune system. Meanwhile, Cxcl9-high monocytes and cDC2s play important roles in immune surveillance and antigen presentation [15]. During inflammatory responses, Cxcl9-high monocytes secrete the chemokine CXCL9, which recruits cDC2s to the site of inflammation, enhancing antigen presentation and T cell activation, thereby strengthening the immune response. Lastly, Ly6-high monocytes and NK cells collaborate in responding to pathogens and tumor cells. Ly6-high monocytes secrete cytokines such as IL-12 and IL-18, which promote the activation, proliferation, and cytotoxic function of NK cells. This interaction enables Ly6-high monocytes and NK cells to work closely together in antiviral and antitumor immune responses, effectively counteracting pathogen threats [16]. Through the finely tuned regulation and cooperation of these cells, the immune system maintains balance across different immune response scenarios, ensuring an adequate response to pathogens while preventing immune overactivation that could lead to tissue damage [17, 18].

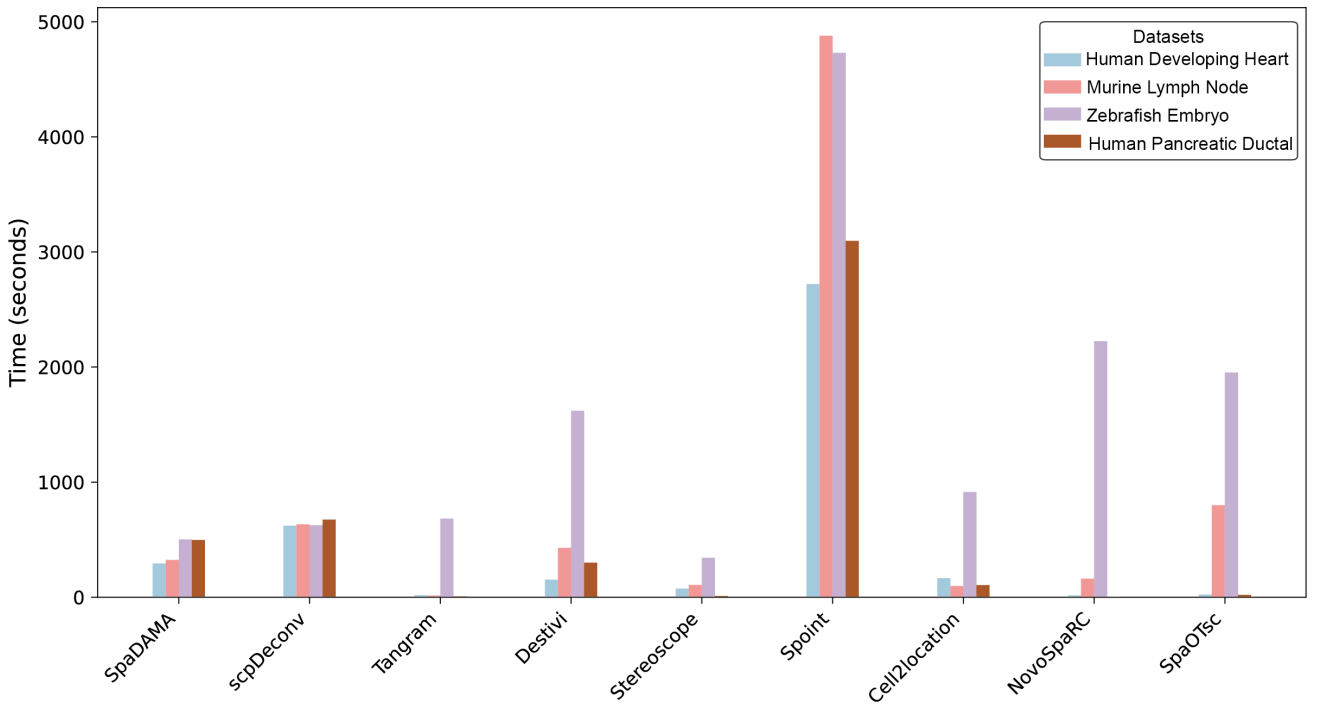
The predictions made by SpaDAMA in the human developing heart dataset reveal the spatial co-localization of certain cell types, highlighting the synergistic interactions between different cell types during cardiac development (Figure S4C). For instance, Epicardium-derived cells secrete various growth factors (such as FGF, BMP) that regulate the proliferation, differentiation, and function of *Myoz2-enriched\_cardiomyocytes* cells [19, 20]. The interaction between these two cell types plays a crucial role in the early stages of heart development, driving angiogenesis and promoting the thickening of the heart wall [21, 22]. Additionally, *Endothelium\_pericytes\_adventitia* cells and *Fibroblast\_like\_2* cells collaborate closely in the process of vascular formation [23, 24]. *Endothelium\_pericytes\_adventitia* cells primarily contribute to the formation, stabilization, and maintenance of blood vessels, thus participating in the construction of the cardiac vasculature [21, 23]. Meanwhile, *Fibroblast\_like\_2* cells support these structures by synthesizing extracellular matrix components and plays an essential role in the structural and functional remodeling of the heart [24]. Moreover, *Erythrocytes\_1* cells and *Erythrocytes\_2* cells represent distinct erythrocyte subpopulations at different stages of cardiac development [25, 26]. *Erythrocytes\_1* cells are early-stage erythrocytes in the embryo, primarily responsible for providing initial oxygen transport, while *Erythrocytes\_2* cells represent mature erythrocytes with adult-type hemoglobin, capable of more efficient oxygen transport. The gradual transition between these two erythrocyte subpopulations reflects adaptive changes in blood supply during heart development, ensuring that the heart meets its physiological demands at different developmental stages [25, 26].

The predictions made by SpaDAMA in the human pancreatic ductal adenocarcinoma dataset provide insight into the cell interactions at various stages of tumor development, including early invasion, malignant transformation, and subsequent immune evasion mechanisms (Figure S4D). In the early stages of PDAC, the interaction between *Ductal\_CRISP3\_high-centroacinar\_like* and *Ductal\_terminal\_ductal\_like* cells likely plays a critical role in the formation of the tumor microenvironment and its invasive characteristics [27]. Specifically, during Acinar-to-Ductal Metaplasia (ADM), *Acinar\_cells* undergo transformation into ductal-like cells and interact with *Endocrine\_cells*, driving tumorigenesis and progression [28]. This process is closely linked to early tumor initiation, proliferation, and local metastatic formation, potentially by modifying the composition of the tumor microenvironment, which fosters tumor malignancy. Moreover, *Endothelial\_cells* and *mDCs\_B* each play significant roles in immune surveillance, but in the PDAC tumor microenvironment, these cells' functions may be suppressed by tumor cell-mediated immune evasion mechanisms [29]. Specifically, *Endothelial\_cells* contribute to tumor angiogenesis by secreting immune-suppressive factors, such as VEGF, while potentially promoting immune tolerance to inhibit *mDCs\_B* cells' anti-tumor immune functions. This mechanism may diminish the ability of *mDCs\_B* cells to activate effector T cells, thus allowing the tumor to evade immune surveillance and promoting immune escape [30].

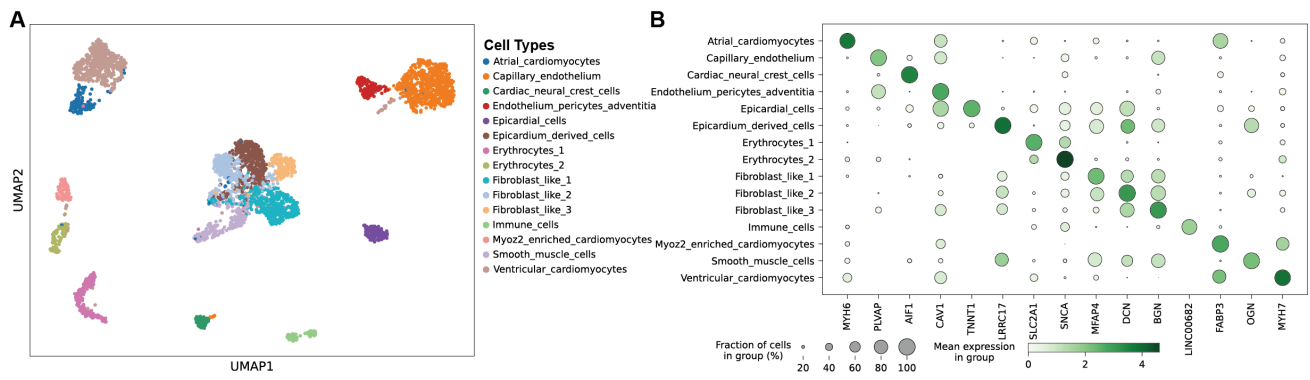
## Supplementary Figures



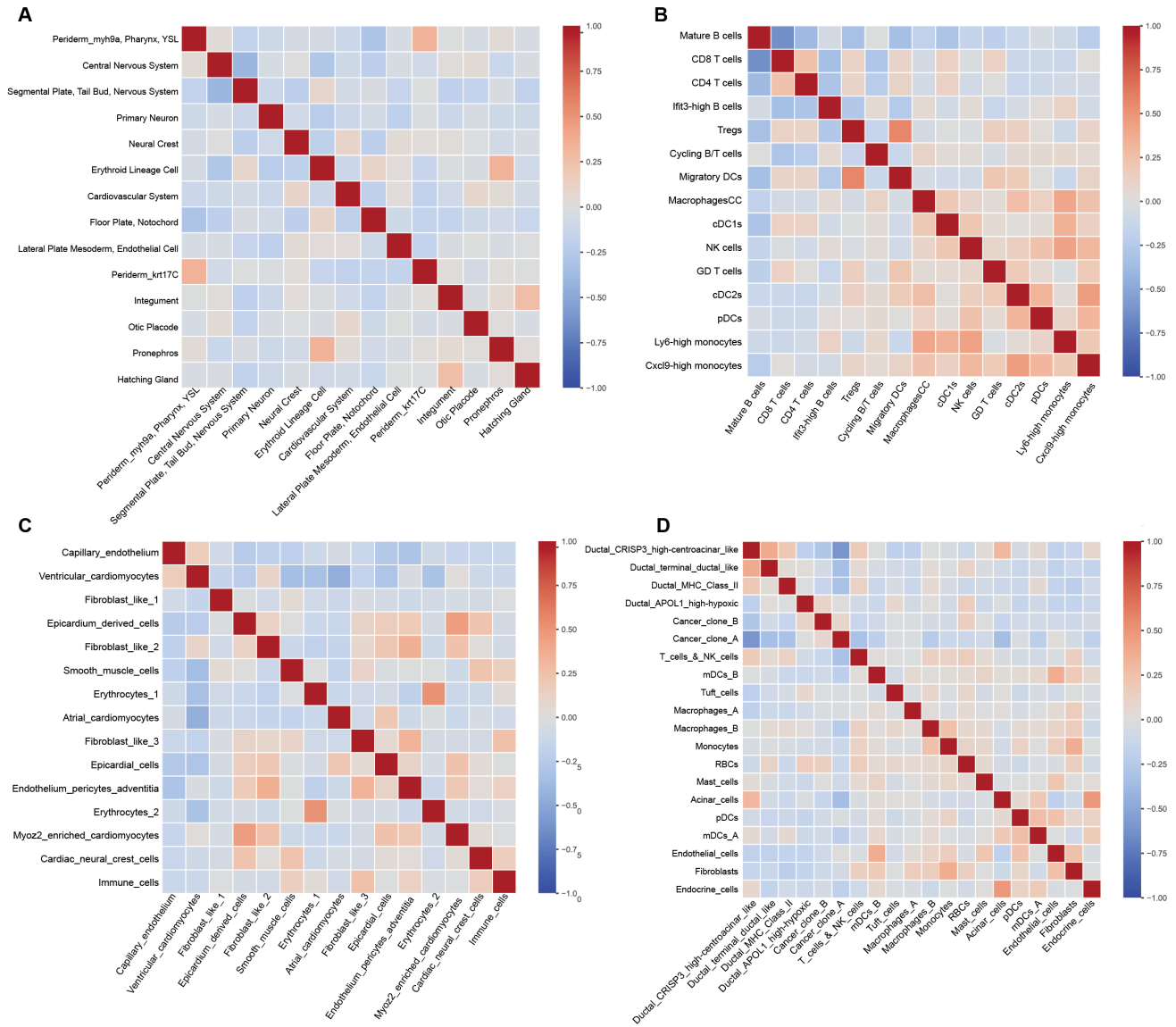
**Figure S1.** We evaluated the PCC, SSIM, RMSE, and JS values of SpaDAMA at different mask rates  $\rho$  in the 32 simulated datasets.



**Figure S2.** The runtime of nine methods on four real-world datasets is presented. When the dataset size is small (with fewer than 2000 spots), Tangram runs very quickly, typically in just a few seconds. Similarly, novoSpaRC and SpaOTsc also complete computations rapidly when the dataset size is even smaller (with fewer than 1000 spots), usually within a few seconds. As a result, the execution times of these methods are difficult to clearly depict in the bar chart.



**Figure S3.** SpaDAMA analyzes cell types in the Human Developing Heart (HDH) dataset. **(A)** Clustering results of scRNA-seq data based on cell types from the same tissue. The UMAP clustering plot reveals distinct clusters corresponding to different cell types, with significant differences in gene expression profiles. Similar cell subtypes (e.g., the three fibroblast subgroups) cluster together, suggesting potential functional or developmental relationships, possibly reflecting shared biological processes or signaling pathways. **(B)** Marker genes selected for each cell type. The marker genes show the highest expression levels and relative abundance within their respective cell populations, further supporting the distinct gene expression characteristics between different cell types.



**Figure S4.** Correlations in cell-type proportions across spatial locations between pairs of cell types inferred by SpaDAMA. The color scale represents the correlation values. (A) shows the Zebrafish Embryo (ZE) dataset. (B) shows the mouse lymph node (MLN) dataset. (C) shows the human developing heart (HDH) dataset. (D) shows the human pancreatic ductal adenocarcinoma (PDAC) dataset.

## Supplementary Tables

**Table S1.** Detailed information of each dataset. The scRNA-seq data in Dataset1 to Dataset32 are real data, and the matching ST data are simulated, while the scRNA-seq and ST data in Dataset33 to Dataset36 are both real data.

Datasets	Tissue	data_type		source_data				preproce_data					
		SC	ST	Number of Cells/Spots		Number of Genes		Number of Cells/Spots		Number of Genes		Droupout rate	
				SC (Cells)	ST (Spots)	SC	ST	SC (Cells)	ST (Spots)	SC	ST	SC	ST
Dataset1	Human-Brain[31]	Real	Simulated	10000	1000	33691	46732	10000	1000	4456	4456	81.29%	61.01%
Dataset2	Mouse-Brain[31]	Real	Simulated	10000	1000	43324	24683	10000	1000	4481	4481	61.05%	76.30%
Dataset3	Human-Liver[32]	Real	Simulated	3821	1000	18328	19850	3821	1000	1186	1186	90.06%	94.83%
Dataset4	Human-Liver[33]	Real	Simulated	6948	1000	20007	26160	6948	1000	1171	1171	92.08%	86.78%
Dataset5	Human-Lung[34]	Real	Simulated	10000	1000	25734	38150	10000	1000	2303	2303	88.12%	95.32%
Dataset6	Human-Lung[35]	Real	Simulated	10000	1000	25734	26828	10000	1000	1639	1639	84.16%	89.17%
Dataset7	Human-Lung[36]	Real	Simulated	10000	1000	22066	25199	10000	1000	2831	2831	88.54%	90.01%
Dataset8	Human-Kidney[37]	Real	Simulated	10000	1000	27345	31489	10000	1000	1531	1531	90.28%	95.15%
Dataset9	Mouse-Kidney[38]	Real	Simulated	10000	1000	24965	29244	10000	1000	1120	1120	96.76%	95.51%
Dataset10	Human-Heart[39]	Real	Simulated	10000	1000	17926	29484	10000	1000	1426	1426	82.13%	95.66%
Dataset11	Human-Heart[39]	Real	Simulated	10000	1000	17926	31580	10000	1000	1426	1426	82.75%	92.18%
Dataset12	Human-Pancreas[40]	Real	Simulated	2282	1000	21198	17499	2282	1000	1077	1077	79.84%	85.17%
Dataset13	Human-Pancreas[40]	Real	Simulated	1040	1000	21625	17499	1040	1000	1427	1427	64.85%	85.48%
Dataset14	Human-Pancreas[41]	Real	Simulated	943	1000	21625	21198	943	1000	1135	1135	66.08%	45.05%
Dataset15	Mouse-Pancreas[40]	Real	Simulated	1382	1000	19745	14860	1382	1000	1504	1504	67.44%	84.52%
Dataset16	Mouse-Trachea[42]	Real	Simulated	6937	1000	27084	18388	6937	1000	1034	1034	87.83%	87.74%
Dataset17	Human-Brain[31]	Real	Simulated	10000	1000	46732	33691	10000	1000	4942	4942	85.99%	79.82%
Dataset18	Mouse-Brain[31]	Real	Simulated	9999	1000	24683	43324	9999	1000	3227	3227	78.03%	37.39%
Dataset19	Human-Liver[43]	Real	Simulated	10000	1000	19850	18328	10000	1000	1183	1183	95.88%	87.03%
Dataset20	Human-Liver[44]	Real	Simulated	8785	1000	26160	20007	8785	1000	1101	1101	90.33%	90.97%
Dataset21	Human-Lung[45]	Real	Simulated	10000	1000	38150	25734	10000	1000	1813	1813	96.07%	55.29%
Dataset22	Human-Lung[46]	Real	Simulated	10000	1000	26828	25734	10000	1000	1502	1502	90.63%	44.66%
Dataset23	Human-Lung[47]	Real	Simulated	10000	1000	25199	22066	10000	1000	2535	2535	92.00%	86.07%
Dataset24	Human-Kidney[48]	Real	Simulated	10000	1000	31489	27345	10000	1000	1218	1218	95.91%	87.56%
Dataset25	Mouse-Kidney[49]	Real	Simulated	10000	1000	29244	24965	10000	1000	885	885	96.63%	94.99%
Dataset26	Human-Heart[50]	Real	Simulated	10000	1000	29484	17926	10000	1000	1077	1077	95.42%	75.15%
Dataset27	Human-Heart[51]	Real	Simulated	10000	1000	31580	17926	10000	1000	1358	1358	93.58%	77.14%
Dataset28	Human-Pancreas[41]	Real	Simulated	7944	1000	17499	21198	7944	1000	1148	1148	87.47%	41.59%
Dataset29	Human-Pancreas[31]	Real	Simulated	8494	1000	17499	21625	8494	1000	1559	1559	87.80%	32.99%
Dataset30	Human-Pancreas[31]	Real	Simulated	2282	1000	21198	21625	2282	1000	1080	1080	81.47%	38.31%
Dataset31	Mouse-Pancreas[52]	Real	Simulated	1827	1000	14860	19745	1827	1000	1143	1143	87.02%	24.93%
Dataset32	Mouse-Trachea[53]	Real	Simulated	7128	1000	18388	27084	7128	1000	1006	1006	88.17%	82.86%
Dataset33	Human Developing Heart[54]	Real	Real	3777	210	15323	38936	3777	209	2373	2373	81.00%	94.01%
Dataset34	Murine Lymph Node[4]	Real	Real	14989	1092	12854	13948	14989	1092	1870	1870	83.41%	61.09%
Dataset35	Zebrafish Embryo[55]	Real	Real	9903	3048	23986	23110	9903	3048	2396	2396	0.00%	0.00%
Dataset36	Human Pancreatic Ductal[56]	Real	Real	1926	426	14121	14121	1926	426	1575	1575	0.00%	0.00%

**Table S2.** Ablation experiments on the 32 simulated datasets in terms of the AS metric. AL: Adversarial Learning, mask: Reconstruction loss computed only for the masked points. Full\_rec: Full reconstruction loss. SpaDAMA (w/o mask, w/o AL): Baseline model with only the encoder and predictor, no mask; SpaDAMA (w/ mask, w/o AL): Baseline model with only the encoder and predictor, with mask, loss computed only for masked points; SpaDAMA (w/o mask, w/ AL): Baseline model with all modules, no mask, adversarial learning added; SpaDAMA (w/ Full\_rec, w/o AL): Baseline model with only the encoder and predictor, with mask, full reconstruction loss.

AS( $\uparrow$ )	Dataset1	Dataset2	Dataset3	Dataset4	Dataset5	Dataset6	dataset7	dataset8
SpaDAMA (w/o mask, w/o AL)	0.638 $\pm$ 0.054	0.563 $\pm$ 0.042	0.538 $\pm$ 0.082	0.363 $\pm$ 0.021	0.425 $\pm$ 0.018	0.438 $\pm$ 0.052	0.513 $\pm$ 0.036	0.413 $\pm$ 0.036
SpaDAMA (w/ mask, w/o AL)	0.425 $\pm$ 0.049	0.813 $\pm$ 0.017	0.313 $\pm$ 0.051	0.425 $\pm$ 0.014	0.475 $\pm$ 0.031	0.625 $\pm$ 0.008	0.425 $\pm$ 0.024	0.413 $\pm$ 0.009
SpaDAMA (w/o mask, w/ AL)	0.700 $\pm$ 0.008	0.363 $\pm$ 0.021	0.725 $\pm$ 0.018	0.800 $\pm$ 0.007	0.525 $\pm$ 0.011	0.650 $\pm$ 0.035	0.638 $\pm$ 0.026	0.800 $\pm$ 0.012
SpaDAMA (w/ Full_rec, w/o AL)	0.513 $\pm$ 0.011	0.375 $\pm$ 0.018	0.625 $\pm$ 0.034	0.550 $\pm$ 0.065	0.575 $\pm$ 0.058	0.375 $\pm$ 0.016	0.463 $\pm$ 0.041	0.475 $\pm$ 0.048
SpaDAMA	0.725 $\pm$ 0.068	0.888 $\pm$ 0.007	0.800 $\pm$ 0.015	0.863 $\pm$ 0.024	0.998 $\pm$ 0.001	0.913 $\pm$ 0.011	0.963 $\pm$ 0.006	0.900 $\pm$ 0.015
AS( $\uparrow$ )	Dataset9	Dataset10	Dataset11	Dataset12	Dataset13	Dataset14	Dataset15	Dataset16
SpaDAMA (w/o mask, w/o AL)	0.550 $\pm$ 0.022	0.425 $\pm$ 0.028	0.400 $\pm$ 0.008	0.450 $\pm$ 0.022	0.400 $\pm$ 0.012	0.400 $\pm$ 0.062	0.363 $\pm$ 0.027	0.513 $\pm$ 0.051
SpaDAMA (w/ mask, w/o AL)	0.625 $\pm$ 0.016	0.375 $\pm$ 0.038	0.388 $\pm$ 0.012	0.588 $\pm$ 0.037	0.763 $\pm$ 0.009	0.650 $\pm$ 0.007	0.425 $\pm$ 0.001	0.400 $\pm$ 0.028
SpaDAMA (w/o mask, w/ AL)	0.700 $\pm$ 0.008	0.825 $\pm$ 0.004	0.775 $\pm$ 0.011	0.663 $\pm$ 0.052	0.638 $\pm$ 0.027	0.588 $\pm$ 0.006	0.750 $\pm$ 0.028	0.563 $\pm$ 0.004
SpaDAMA (w/ Full_rec, w/o AL)	0.325 $\pm$ 0.001	0.425 $\pm$ 0.011	0.463 $\pm$ 0.037	0.475 $\pm$ 0.084	0.325 $\pm$ 0.004	0.363 $\pm$ 0.012	0.550 $\pm$ 0.035	0.588 $\pm$ 0.021
SpaDAMA	0.800 $\pm$ 0.035	0.950 $\pm$ 0.005	0.975 $\pm$ 0.003	0.825 $\pm$ 0.038	0.875 $\pm$ 0.008	1.000 $\pm$ 0.000	0.913 $\pm$ 0.004	0.938 $\pm$ 0.006
AS( $\uparrow$ )	Dataset17	Dataset18	Dataset19	Dataset20	Dataset21	Dataset22	Dataset23	Dataset24
SpaDAMA (w/o mask, w/o AL)	0.425 $\pm$ 0.024	0.750 $\pm$ 0.032	0.375 $\pm$ 0.018	0.363 $\pm$ 0.036	0.588 $\pm$ 0.007	0.525 $\pm$ 0.041	0.450 $\pm$ 0.015	0.425 $\pm$ 0.041
SpaDAMA (w/ mask, w/o AL)	0.488 $\pm$ 0.074	0.400 $\pm$ 0.032	0.575 $\pm$ 0.091	0.325 $\pm$ 0.023	0.650 $\pm$ 0.002	0.488 $\pm$ 0.024	0.525 $\pm$ 0.034	0.888 $\pm$ 0.021
SpaDAMA (w/o mask, w/ AL)	0.638 $\pm$ 0.012	0.438 $\pm$ 0.036	0.650 $\pm$ 0.055	0.775 $\pm$ 0.003	0.650 $\pm$ 0.045	0.763 $\pm$ 0.026	0.500 $\pm$ 0.012	0.338 $\pm$ 0.002
SpaDAMA (w/ Full_rec, w/o AL)	0.550 $\pm$ 0.052	0.488 $\pm$ 0.046	0.650 $\pm$ 0.062	0.638 $\pm$ 0.037	0.238 $\pm$ 0.006	0.400 $\pm$ 0.073	0.525 $\pm$ 0.034	0.563 $\pm$ 0.037
SpaDAMA	0.900 $\pm$ 0.015	0.925 $\pm$ 0.003	0.750 $\pm$ 0.138	0.900 $\pm$ 0.028	0.875 $\pm$ 0.014	0.825 $\pm$ 0.038	0.998 $\pm$ 0.001	0.788 $\pm$ 0.022
AS( $\uparrow$ )	Dataset25	Dataset26	Dataset27	Dataset28	Dataset29	Dataset30	Dataset31	Dataset32
SpaDAMA (w/o mask, w/o AL)	0.463 $\pm$ 0.026	0.413 $\pm$ 0.057	0.488 $\pm$ 0.002	0.388 $\pm$ 0.022	0.613 $\pm$ 0.036	0.550 $\pm$ 0.047	0.588 $\pm$ 0.052	0.538 $\pm$ 0.056
SpaDAMA (w/ mask, w/o AL)	0.738 $\pm$ 0.017	0.300 $\pm$ 0.008	0.513 $\pm$ 0.022	0.500 $\pm$ 0.028	0.800 $\pm$ 0.028	0.638 $\pm$ 0.029	0.425 $\pm$ 0.043	0.313 $\pm$ 0.011
SpaDAMA (w/o mask, w/ AL)	0.613 $\pm$ 0.026	0.763 $\pm$ 0.027	0.750 $\pm$ 0.052	0.888 $\pm$ 0.011	0.488 $\pm$ 0.022	0.475 $\pm$ 0.008	0.663 $\pm$ 0.012	0.525 $\pm$ 0.043
SpaDAMA (w/ Full_rec, w/o AL)	0.275 $\pm$ 0.009	0.613 $\pm$ 0.042	0.325 $\pm$ 0.004	0.450 $\pm$ 0.042	0.400 $\pm$ 0.022	0.513 $\pm$ 0.012	0.363 $\pm$ 0.021	0.738 $\pm$ 0.021
SpaDAMA	0.913 $\pm$ 0.002	0.913 $\pm$ 0.011	0.925 $\pm$ 0.004	0.775 $\pm$ 0.069	0.700 $\pm$ 0.105	0.825 $\pm$ 0.014	0.963 $\pm$ 0.002	0.888 $\pm$ 0.037

## References

1. H. Xu and et al. Spacel: deep learning-based characterization of spatial transcriptome architectures. *Nature Communications*, 14:7603, 2023.
2. T. Biancalani and et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with tangram. *Nature Methods*, 18:1352–1362, 2021.
3. V. Kleshchevnikov and et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nature Biotechnology*, 40:661–671, 2022.
4. R. and et al. Lopez. Destvi identifies continuums of cell types in spatial transcriptomics data. *Nature Biotechnology*, 40:1360–1369, 2022.
5. F. Wang and et al. Deep domain adversarial neural network for the deconvolution of cell type mixtures in tissue proteome profiling. *Nature Machine Intelligence*, 5:1236–1249, 2023.
6. A. Andersson, J. Bergenstr hle, et al. Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Communications Biology*, 3:565, 2020.
7. Z. Cang and Q. Nie. Inferring Spatial and Signaling Relationships Between Cells from Single-Cell Transcriptomic Data. *Nature Communications*, 11:2084, 2020.
8. N. Moriel et al. NovoSpaRc: Flexible Spatial Reconstruction of Single-Cell Gene Expression with Optimal Transport. *Nature Protocols*, 16:4177–4200, 2021.
9. C. B. Kimmel et al. Stages of embryonic development of the zebrafish. *Developmental Dynamics*, 203(3):253–310, 1995.
10. J. H. Gutzman and H. L. Sive. The role of the periderm in zebrafish embryogenesis. *Developmental Biology*, 332(2):359–370, 2009.
11. B. Thisse and C. Thisse. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nature Protocols*, 3(1):59–69, 2008.
12. C. P. Heisenberg and C. N sslein-Volhard. *Zebrafish: A practical approach*. Oxford University Press, 2009.
13. M. L. Martin et al. Role of the hatching gland in zebrafish embryonic development. *Developmental Biology*, 271(1):1–10, 2004.
14. Y. Liu et al. Treg cells and dendritic cells: The role of reciprocal interaction in maintaining tolerance. *Immunological Reviews*, 259(1):159–169, 2014.
15. A. Schlitzer et al. Human monocyte subsets and their differentiation in health and disease. *Immunology*, 146(2):126–138, 2015.
16. C. Fauriat et al. Regulation of human nk-cell cytokine production by target cell interaction and its impact on nk-cell-mediated killing. *Blood*, 115(3):651–658, 2010.
17. C. R. Mackay et al. The role of immune cells in the induction and regulation of immune responses. *Immunity*, 39(4):581–593, 2013.
18. J. Banchereau et al. The immune system and cancer: A review. *Cell*, 148(3):449–461, 2012.
19. B. Zhou et al. Epicardial progenitors contribute to the cardiomyocyte lineage in the developing heart. *Nature*, 469(7330):301–305, 2011.
20. K. N. Ivey et al. Transcriptional regulation of epicardial to myocardial signaling in heart development. *Nature Reviews Molecular Cell Biology*, 19(4):224–240, 2018.
21. H. Gerhardt and C. Betsholtz. Endothelial–pericyte interactions in angiogenesis. *Cell and Tissue Research*, 314(1):15–23, 2003.
22. Y. Cao et al. Vascular remodeling during heart development. *Current Opinion in Hematology*, 22(1):49–54, 2015.
23. K. J. Lavine et al. Fibroblast growth factor signaling controls the patterning of the proepicardium and the development of the coronary vasculature. *Developmental Cell*, 8(4):551–563, 2005.
24. C. A. Souders et al. Fibroblasts and the remodeling of the heart. *Circulation Research*, 105(1):51–62, 2009.
25. J. Palis et al. Hematopoiesis in the embryo and adult: insights from the zebrafish model. *Developmental Biology*, 348(1):278–288, 2010.
26. J. Hort et al. Development of erythrocytes in the early vertebrate embryo. *Developmental Biology*, 289(1):1–16, 2006.
27. S. Yadav et al. Acinar-to-ductal metaplasia and pancreatic cancer: Insights into mechanisms of tumor initiation and progression. *Frontiers in Oncology*, 10:872, 2020.
28. F. C. Martins et al. Endothelial cell modulation of the immune microenvironment in pancreatic cancer. *Cancer Research*, 80(18):3983–3994, 2020.
29. M. A. Gon alves et al. The role of dendritic cells in pancreatic cancer: Immune tolerance and immune escape mechanisms. *Frontiers in Immunology*, 13:725062, 2022.
30. M. Wang et al. Immune escape mechanisms in pancreatic cancer and their implications for immunotherapy. *Journal of Hematology & Oncology*, 14(1):87, 2021.
31. B. Li, W. Zhang, and et al. Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. *Nature Methods*, 19:662–670, 2022.
32. L. et al. Ma. Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell*, 36(4):418–430.e6, 2019.
33. S. A. et al. MacParland. Single cell rna sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nature Communications*, 9(1):4383, 2018.
34. F. et al. Fan. Elevated mast cell abundance is associated with enrichment of ccr2+ cytotoxic t cells and favorable prognosis in lung adenocarcinoma. *Cancer Research*, 83(16):2690–2703, Aug 2023.
35. Z. et al. Wang. Single-cell transcriptomic analysis reveals an immunosuppressive network between postn cafs and ackr1 ecs in tki-resistant lung cancer. *Cancer Genomics & Proteomics*, 21(1):65–78, Jan–Feb 2024.



36. X. et al. Jin. Dissection of the cell communication interactions in lung adenocarcinoma identified a prognostic model with immunotherapy efficacy assessment and a potential therapeutic candidate gene *itgb1*. *Heliyon*, 10(17):e36599, Sep 2024.
37. B. B. et al. Lake. A single-nucleus rna-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nature Communications*, 10(1):2832, Jun 2019.
38. J. et al. Cao. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science*, 361(6409):1380–1385, Sep 2018.
39. C. B. et al. Sim. Sex-specific control of human heart maturation by the progesterone receptor. *Circulation*, 143(16):1614–1628, Apr 2021.
40. M. et al. Baron. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Systems*, 3(4):346–360.e4, Oct 2016.
41. M. et al. Enge. Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. *Cell*, 171(2):321–330.e14, Oct 2017.
42. D. T. et al. Montoro. A revised airway epithelial hierarchy includes *cftr*-expressing ionocytes. *Nature*, 560(7718):319–324, Aug 2018.
43. A. et al. Sharma. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell*, 183(2):377–394.e21, Oct 2020.
44. N. et al. Aizarani. A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature*, 572(7768):199–204, Aug 2019.
45. R. et al. Zilionis. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity*, 50(5):1317–1334.e10, May 2019.
46. N. et al. Kim. Single-cell rna sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nature Communications*, 11(1):2285, May 2020.
47. Synapse. Dataset: syn21041850, 2024. Accessed: 2024-12-27.
48. Y. et al. Muto. Defining cellular complexity in human autosomal dominant polycystic kidney disease by multimodal single cell analysis. *Nature Communications*, 13(1):6497, Oct 2022.
49. Y. et al. Muto. Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nature Communications*, 12(1):2190, Apr 2021.
50. N. R. et al. Tucker. Transcriptional and cellular diversity of the human heart. *Nature*, 2024. In press.
51. A. et al. Daoud. Integrative single-cell analysis of cardiac and pulmonary sarcoidosis using publicly available cardiac and bronchoalveolar lavage fluid sequencing datasets. *Frontiers in Cardiovascular Medicine*, 10:101227818, 2023.
52. Tabula Muris Consortium et al. Single-cell transcriptomics of 20 mouse organs creates a tabula muris. *Nature*, 562(7727):367–372, Oct 2018.
53. L. W. et al. Plasschaert. A single-cell atlas of the airway epithelium reveals the *cftr*-rich pulmonary ionocyte. *Nature*, 560(7718):377–381, Aug 2018.
54. M. Asp, S. Giacomello, and et al. A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell*, 179(7):1647–1660.e19, 2019.
55. H. Li, J. Zhou, Z. Li, et al. A Comprehensive Benchmarking with Practical Guidelines for Cellular Deconvolution of Spatial Transcriptomics. *Nature Communications*, 14:1548, 2023.
56. Y. Ma and X. Zhou. Spatially Informed Cell-Type Deconvolution for Spatial Transcriptomics. *Nature Biotechnology*, 40(10):1349–1359, 2022.