SpaTM: Topic Models for Inferring Spatially Informed Transcriptional Programs

Adrien Osakwe,[1,*] Wenqi Dong,[2] Qihuang Zhang,[1,3] Robert Sladek[1,4,*] and Yue Li[1,2,*]

[1]Quantitative Life Sciences Program, [2]School of Computer Science, [3]Department of Epidemiology, Biostatistics and Occupational Health and [4]Department of Human Genetics, McGill University, 3480 Rue University, Montréal, H3A 2A7, QC, Canada
[*]Corresponding authors: adrien.osakwe@mail.mcgill.ca rob.sladek@mcgill.ca, yueli@cs.mcgill.ca

Abstract

Spatial transcriptomics has revolutionized our ability to characterize tissues and diseases by contextualizing gene expression with spatial organization. Available methods require researchers to either train a model using histology-based annotations or use annotation-free clustering approaches to uncover spatial domains. However, few methods provide researchers with a way to jointly analyze spatial data from both annotation-free and annotation-guided perspectives using consistent inductive biases and levels of interpretability. A single framework with consistent inductive biases ensures coherence and transferability across tasks, reducing the risks of conflicting assumptions. To this end, we propose the Spatial Topic Model (SpaTM), a topic-modeling framework capable of annotation-guided and annotation-free analysis of spatial transcriptomics data. SpaTM can be used to learn gene programs that represent histology-based annotations while providing researchers with the ability to infer spatial domains with an annotation-free approach if manual annotations are limited or noisy. We demonstrate SpaTM's interpretability with its use of topic mixtures to represent cell states and transcriptional programs and how its intuitive framework facilitates the integration of annotation-guided and annotation-free analyses of spatial data with downstream analyses such as cell type deconvolution. Finally, we demonstrate how both approaches can be used to extend the analysis of large-scale snRNA-seq atlases with the inference of cell proximity and spatial annotations in human brains with Major Depressive Disorder.

Key words: Spatial Transcriptomics, Topic Modeling, Bayesian Inference, Spatial Segmentation, Neighbor Prediction

Introduction

Advances in Spatial Transcriptomics (ST) have made it possible to reliably profile spatially oriented gene expression throughout the transcriptome. The resulting datasets have enabled novel discoveries by helping researchers identify transcriptionally distinct spatial regions in different tissues and diseases (Moses and Pachter, 2022; Maynard et al., 2021). Many spatial transcriptomics analyses have been done in brain and cancer datasets, where researchers have made extensive use of spatial information to compare changes in spatial regions, cell type composition and cell-cell interactions (Zhang et al., 2023; Raghavan et al., 2024; Miller et al., 2022; Bonev et al., 2024).

One of the main analyses done for spatial transcriptomics is region segmentation (Moses and Pachter, 2022). In particular, researchers face the challenge of exploring domain relationships and identity changes at bordering regions. Segmentation can be done by training a model on histology- and cytoarchitecture-based manual annotations (CeLEry) or through spatial clustering approaches (BayesSpace, SpaGCN) (Zhang et al., 2023; Zhao et al., 2021a; Hu et al., 2021). Although these methods have demonstrated state-of-the-art performance in their respective tasks, there is still a need for methods which can supplement these tasks with intuitive, easy-to-interpret results. For example, BayesSpace and SpaGCN identify spatially-resolved markers by first generating spatial clusters and running a differential expression analysis on the clusters. As previously demonstrated by Neufeld et al. (2024), such approaches are prone to generating false positives during marker gene detection. Thus, it is preferable to identify methods which jointly learn spatial domains and their markers. Finally, most methods are designed to perform only one of these tasks, requiring researchers to integrate results from different methods to perform more comprehensive analyses. This raises new complications, as methods vary in their inductive biases and may not provide results with similar levels of interpretability.

In addition, the limited availability of spatial datasets for specific tissues and diseases requires researchers to find ways to impute spatial information into single cell atlases to perform large-scale spatial analyses (Tian et al., 2023). These imputation tasks facilitate our ability to link the findings made in single cell studies with spatial studies, leading to more nuanced hypothesis generation for experimental efforts. For example, the brains of patients suffering from Major Depressive Disorder (MDD) have been shown to experience distinct morphological changes (Stockmeier et al., 2004). Although large-scale single cell analyses of MDD have revealed transcriptional changes in MDD onset, there is still a need to integrate these findings with spatial information to characterize pathogenetic mechanisms (Maitra et al., 2023). Thus, finding methods that can reliably model spatial transcriptomics data and project it onto single cell datasets is crucial to integrate the findings from both modalities.

Currently, there is a lack of a computational framework that can easily be adjusted for different types of spatial analyses and imputing spatial information while ensuring the seamless and intuitive integration of downstream findings. Such a framework will greatly aid the analysis of single cell disease atlases. To this end, we present the Spatial Topic Model (SpaTM), a topic modeling framework capable of spatial deconvolution, spatial label prediction, spot proximity prediction, and de novo inference of spatial gene signatures (Fig. 1). By using

the interpretable and flexible framework of the topic model as its foundation, SpaTM is able to easily undertake different forms of spatial analyses with the same inductive bias. This ensures a more practical analysis framework for users while retaining competitive performance in different tasks. We have also implemented SpaTM into R and extended the SingleCellExperiment and SpatialExperiment object classes to facilitate the integration of our tool into large-scale analyses undertaken by life science researchers (Righelli et al., 2022). Finally, we demonstrate how the SpaTM framework can be used to impute spatial information into a publicly available MDD snRNA-seq dataset to investigate spatially oriented changes in MDD and uncover novel pathogenetic insights.

Methods

SpaTM

SpaTM is built around topic models, an interpretable and principled machine learning framework for discovering latent topics from an unlabeled corpus of documents treated as bags of words (Fig. 1) (Blei et al., 2003). The topics are distinct sets of multinomial distributions over the entire vocabulary. In genomics, the topics can be interpreted as modules of co-expressed genes. Due to the natural probabilistic interpretation and the existence of efficient inference algorithms, topic models have a broad and long history of success in many domains that involve sparse and high dimensional data including scRNA-seq and ST data (Zhao et al., 2021b; Swapna et al., 2023; Miller et al., 2022). In our context, "documents" are spots in ST (or cells in scRNA-seq); "words" represent genes detected by sequencing; "tokens" are individual transcript reads of the same gene and "topics" are biological processes. The value of topic modeling has previously been demonstrated with a Guided-Topic Model for deconvolution (GTM-Decon) tasks, which enabled researchers to ensure the topics learned are anchored to cell types of interest (Swapna et al., 2023). SpaTM extends GTM-Decon (hereby referred to as SpaTM-G) for spatial label prediction with the supervised topic model (SpaTM-S) (Mcauliffe and Blei, 2007) and cell proximity inference via the relational topic model (SpaTM-R) (Fig. 1C) (Chang and Blei, 2009).

SpaTM-G —— Guided Topic Model

Our base topic model follows the collapsed-variational Bayes latent Dirichlet allocation (LDA) (Teh et al., 2006). Here, we assume that each spot $s \in \{1, ..., S\}$ is a mixture of K topics. This mixture of topics θ_s is in turn sampled from a K-dimensional Dirichlet distribution $\theta_s \sim Dir(\alpha_{\mathbf{s}})$ with a spot-topic prior α_s. While α_s is considered as a constant across spots by default, our guided topic model framework (SpaTM-G) exploits the fact that α_s can be anchored to cell types, enabling cell type deconvolution on unseen samples (Swapna et al., 2023). For each spot s, with a library size N_s, the topic assignment of each read, $z_{i,s}$, for read $i \in \{1, ..., N_s\}$ is randomly sampled from one of the K topics based on the corresponding topic mixture by $z_{i,s} \sim Categorical(\theta_s)$. With the topic assignment $z_{i,s} = k$, we then map the ith read to a gene x based on the distribution of genes for the sampled topic $x_{i,s} \sim Categorical(\phi_k)$ where ϕ_k is sampled from a Dirichlet distribution over all genes with $\phi_k \sim Dir(\beta)$. Thus, the key objective of the topic model is to learn a spots-by-topics matrix θ containing the mixture of topics representing each spot and a genes-by-topics matrix ϕ which represents the gene composition of each topic. Using the Dirichlet-Multinomial conjugacy, we can obtain simplified posteriors for the topic and gene assignments:

$$p(z|\alpha) = \prod_s \frac{\Gamma(\sum_k \alpha_{s,k})}{\prod_k \Gamma(\alpha_{s,k})} \frac{\prod_k \Gamma(\alpha_{s,k} + n_{.,s,k})}{\Gamma(\sum_k \alpha_{s,k} + n_{.,s,k})},$$

$$p(x|z) = \prod_k \frac{\Gamma(G\beta)}{\prod_g \Gamma(\beta)} \frac{\prod_g \Gamma(\beta + n_{g,.,k})}{\Gamma(G\beta + \sum_g n_{g,.,k})},$$

where G is the total number of genes in the dataset and the sufficient statistics $n_{.,s.k}$ and $n_{g,.,k}$ represent the number of reads assigned to topic k for spot s and gene g respectively. This also implies that we can calculate the expected values of θ_s and ϕ_k with the posterior estimates of our topic assignments as they are proportional to the sufficient statistics. The conditional distribution of the topic assignment for read i in spot s, while fixing the assignment for all other tokens can be represented in closed-form:

$$p\left(z_{i,s} = k | x_{i,s} = g, z^{-(i,s)}, x^{-(i,s)}\right)$$

$$\propto p\left(z_{i,s} = k, z^{-(i,s)} | \alpha_{s,k}\right) p\left(x_{i,s} = g, x^{-(i,s)} | z_{i,s} = k, z^{-(i,s)}\right)$$

$$\propto \left(\alpha_{s,k} + n^{(-i)}_{.,s,k}\right) \left(\frac{\beta + n^{-(i,s)}_{g,.,k}}{G\beta + \sum_g n^{-(i,s)}_{g,.,k}}\right).$$

Topic inference can be achieved by collapsed variational Bayes (CVB), where we approximate the posterior topic assignment distribution $p(z_{i,s}|x, \alpha)$ as $q(z_{i,s}|x_{i,s} = g) = \prod_k \gamma_{i,s,k}^{[z_{i,s}=k]}$, where

$$\gamma_{i,s,k} \propto \left(\alpha_{s,k} + n^{(-i)}_{.,s,k}\right) \left(\frac{\beta + n^{-(i,s)}_{g,.,k}}{G\beta + \sum_g n^{-(i,s)}_{g,.,k}}\right).$$

Here the sufficient statistics are calculated using the variational parameters as topic assignments:

$$n^{(-i)}_{.,s,k} = \sum_{i' \neq i}^{N_s} \gamma_{i',s,k},$$

$$n^{(-i)}_{g,.,k} = \sum_{i' \neq i}^{N_s} [x_{i',s'} = g] \gamma_{i',s,k}.$$

By assuming that all reads for a given gene g in a spot s have the same topic assignment posterior distribution, we can calculate the variational parameters at the level of genes instead of reads, thereby improving computational efficiency:

$$\gamma_{g,s,k} \propto \left(\alpha_{s,k} + n^{(-g)}_{.,s,k}\right) \left(\frac{\beta + n^{-(s)}_{g,.,k}}{G\beta + \sum_g n^{-(s)}_{g,.,k}}\right).$$

The model parameters γ, θ, ϕ are inferred via the Expectation-Maximization (EM) algorithm, where we alternate between estimating the variational parameter γ in the E-step and updating the sufficient statistics $n_{.,s,k}$ and $n_{g,.,k}$ in the M-step. This iterative approach maximizes the evidence lower bound (ELBO):

$$ELBO = E_q\left[\log p\left(x|z\right)\right] + E_q[\log p(z|\alpha)] - E_q[\log q\left(z\right)],$$

which is used to assess model convergence. Upon convergence, topic mixtures θ_s and ϕ_k can be estimated as the expectations

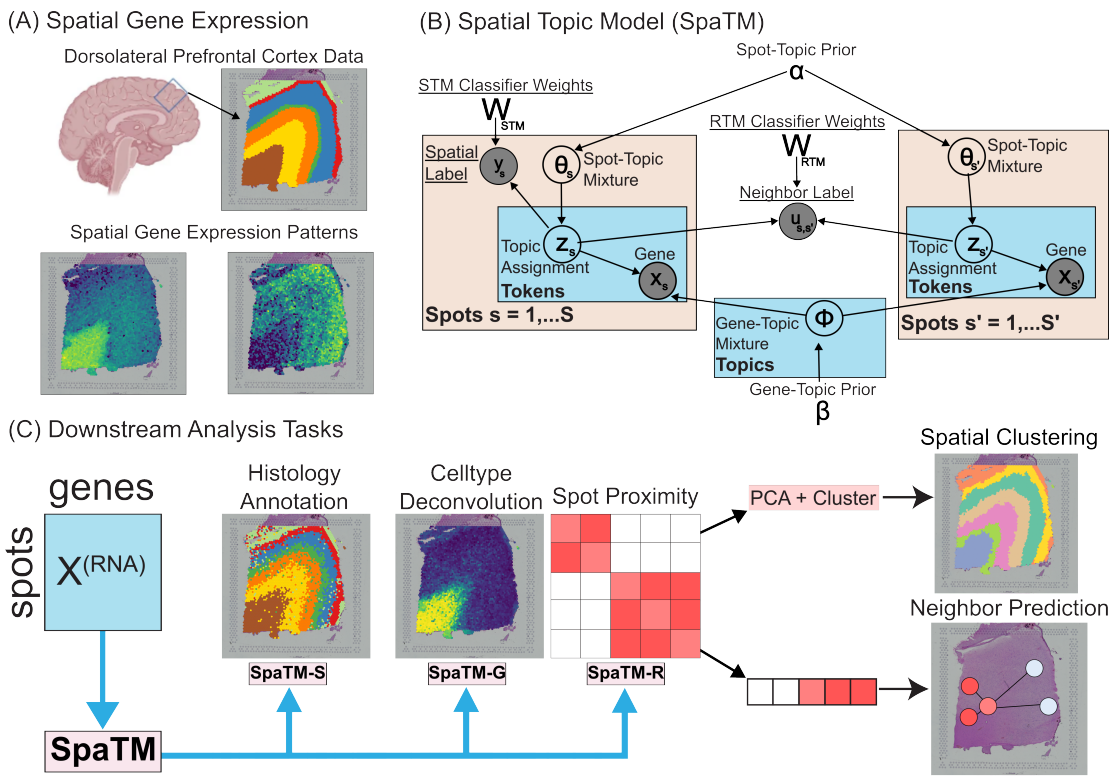

Fig. 1. SpaTM for a unified analysis of Spatial Transcriptomics data. A) SpatialLIBD 10X Visium dataset for dorsolateral prefrontal cortex slices containing manual histology-based annotations and spatially-driven gene expression. B) Spatial Topic Model (SpaTM) modeling framework for spatial transcriptomics data depicted as a probabilistic graphical model. The framework extends the GTM-Decon tool to include a supervised (STM) modeling component as well as a spot proximity (RTM) component. C) SpaTM provides a unified framework for ST tasks through its use of the STM (SpaTM-S), GTM (SpaTM-G) and RTM (SpaTM-R) components.

of a Dirichlet-Multinomial:

$$\mathbb{E}_q\left[\theta_{s,k}\right] = \frac{\alpha_{s,k} + n_{.,s,k}}{\sum_k \alpha_{s,k} + n_{.,s,k}} \equiv \widehat{\theta}_{\mathrm{s,k}},$$

$$\mathbb{E}_q\left[\phi_{\mathrm{g,k}}\right] = \frac{\beta + n_{g,.,k}}{G\beta + \sum_g n_{g,.,k}} \equiv \widehat{\phi}_{\mathrm{g,k}}.$$

Inference on unseen observations can be achieved by fixing the gene-topic component of the variational parameter estimate to be $\hat{\phi}$ such that

$$\gamma_{g,s,k} \propto \left(\alpha_{s,k} + n^{(-g)}_{.,s,k}\right)\hat{\phi}_{g,k}.$$

Running the E-M algorithm using the above estimate until convergence can then enable us to infer $\hat{\theta}$ for unseen observations as

$$\mathbb{E}_q\left[\theta_{s,k}\right] = \frac{\alpha + n_{.,s,k}}{K\alpha + n_{.,s,k}}.$$

SpaTM-S —— Supervised Topic Model

By using topic mixtures learned by a topic model as input variables for predicting a spatial region y, we can reduce high-dimensional and noisy gene expression data into a few interpretable topics that represent spatially-oriented gene programs. Although this can be done by sequentially training a basic LDA topic model with a prediction algorithm, it does not guarantee that the topics learned will be optimal for explaining and predicting spatial regions. Thus, we developed a supervised topic model (SpaTM-S), an extension of LDA that jointly trains a topic model with a prediction algorithm (Mcauliffe and Blei, 2007). This joint training ensures that we learn the underlying gene programs (topics) that predict spatial regions. We assume that the spatial label $y_s \in \{1, ..., L\}$ is generated from a categorical distribution:

$$y_s \sim Categorical(\pi_s) = \prod_l \pi_{s,l}^{[y_s=l]},$$

where π_s represents a $1 \times L$ vector of probabilities generated by the softmax outputs of a logistic classifier:

$$p(y_s = l|\overline{z}_s, W^{STM}) = \prod_{j=1}^{L} \pi_{s,j}^{[y_s=j]} = \frac{\exp(\overline{z}_s W_l^{STM})}{\sum_{j=1}^{L}\exp(\overline{z}_s W_j^{STM})}.$$

Model weights are represented by a $K \times L$ matrix W^{STM} and $\overline{z}_s$ is a $1 \times K$ vector that represents the average topic assignment for spot s, with the kth entry of $\overline{z}_s$ given by

$$\overline{z}_{s,k} = \frac{1}{N_s}\sum_{i=1}^{N_s}[z_{i,s} = k],$$

where N_s represents the total number of reads for spot s. We can then provide a posterior for the spatial label that fixes all

topic assignments but $z_{i,s}$

$$p(y_s = l|z_s^{-(i)}, z_{i,s} = k, W^{STM}) = \prod_{j=1}^{L} \pi_{s,j}^{[z_{i,s}=k][y_s=j]}$$
$$= \frac{\exp(\overline{z}_s^{[z_{i,s}=k]} W_l^{STM})}{\sum_{j=1}^{L} \exp(\overline{z}_s^{[z_{i,s}=k]} W_j^{STM})},$$

where

$$\overline{z}_s^{[z_{i,s}=k]} = \frac{1}{N_s} \left(\sum_{t \neq i}^{N_s} [z_{t,s} = k] + [z_{i,s} = k] \right).$$

Since we know the ground-truth label for each sample, we can rewrite the above probability such that it represents our confidence in each topic assignment's ability to predict the label y_s

$$p(y_s = l|z_s^{-(i)}, z_{i,s} = k, W^{STM}) = \frac{\pi_{s,l}(\overline{z}_s^{[z_{i,s}=k]})}{\sum_{j=1}^{K} \pi_{s,l}(\overline{z}_s^{(z_{i,s}=j)})}.$$

We can thus update the variational parameter estimate to account for spatial label predictions as

$$\gamma_{i,s,k} \propto (\alpha + n_{s,k}^{-(i)}) \left(\frac{\beta + n_{x_{(i,s)},k}^{-i,s}}{\sum_w \beta + n_{w,k}^{-i,s}} \right) \left(\frac{\pi_{s,l}(\overline{z}_s^{[z_{i,s}=k]})}{\sum_{t=1}^{K} \pi_{s,l}(\overline{z}_s^{[z_{i,s}=t]})} \right).$$

As with SpaTM-G, SpaTM-S can be trained using expectation-maximization by estimating the variational parameter in the E-step and optimizing the classifier weights using updated average topic assignments in the M-step. The model's convergence is determined using an updated evidence lower-bound

$$ELBO = \mathbb{E}_{q(z)}[log\ p(z|\alpha)] + \mathbb{E}_{q(z)}[log\ p(x|z)]$$
$$+ \mathbb{E}_{q(z)}[log\ p(y_s = l|z, W^{STM})] - \mathbb{E}_{q(z)}[log\ q(z)].$$

We can also use a multi-layer perceptron (MLP) to predict the labels:

$$p(y_s = l|z_s, W^{STM}) = \pi_{s,y_s} = \frac{exp(g(\overline{z}_s, W^{STM})^{(l)})}{\sum_{j=1}^{L} exp(g(\overline{z}_s, W^{STM})^{(j)})},$$

where $g(\overline{z}_s, W^{STM})^{(l)}$ is the output for class l for an MLP with parameter weights W^{STM}. After training, we can make predictions on unseen cells by using their topic mixture $\hat{\theta}_s$ as input to the trained classifier, where $\hat{\theta}_s$ is inferred by the topic model with the gene-topic component fixed to be the estimate $\hat{\phi}$ from the training data.

Spatially Informed Gene Expression Correction

The gene-topic mixtures $\hat{\phi}$ learned by SpaTM-S can be used to reconstruct the expression profiles of unseen samples to reduce sparsity and noise. We can infer the spot-topic mixtures $\hat{\theta}$ for an unseen slice and compute the matrix product with $\hat{\phi}$ to create a corrected spot-by-gene expression matrix $\hat{X}$. We scale the corrected expression profile of each spot by their library size, making them compatible with existing analysis pipelines.We evaluate the effectiveness of the correction by comparing the spatial autocorrelation (Moran's I) of layer markers when using log-normalized counts from the uncorrected and SpaTM-S-corrected profiles.

SpaTM-R —— Relational Topic Model

The relational topic model (RTM) is a variant of STM that was designed to predict links between documents (Chang and Blei, 2009). Here, we propose to adapt RTM to predict spot proximity and neighbors in ST data (SpaTM-R). We assume that for a pair of spots s, s', we have a binary neighbor indicator $u_{s,s'}$ such that

$$u_{s,s'} \sim Bernoulli(g(.|\overline{z}_s \circ \overline{z}_{s'})),$$

where $g()$ represents a link probability function that defines the probability of spot s' being a neighbor of spot s and $\overline{z}_s \circ \overline{z}_{s'}$ represents the element-wise product of the average topic assignment for each spot. Thus, $g()$ enables SpaTM-R to predict neighboring spots based on the similarity (or dissimilarity) of their expression profiles. We fix our topic assignments and set our link probability function to be a sigmoid function, where

$$g(.|\overline{z}_s^{(z_{s,i}=k)} \circ \overline{z}_{s'}) = \frac{1}{1 + \exp(-W^{RTM}(\overline{z}_s^{(z_{s,i}=k)} \circ \overline{z}_{s'}))}.$$

The Bernoulli likelihood for the neighbor predictions of a given spot s is then defined as:

$$\eta_{s,k} = \sum_{s' \neq s} g(.|\overline{z}_s^{(z_{s,i}=k)} \circ \overline{z}_{s'})^{(u_{s,s'})}$$
$$\times (1 - g(.|\overline{z}_s^{(z_{s,i}=k)} \circ \overline{z}_{s'}))^{(1-u_{s,s'})}.$$

Finally, the topic inference is defined by the joint likelihood of the transcriptome and spatial components of the data:

$$\gamma_{i,s,k} \propto (\alpha + n_{s,k}^{-(i)}) \left(\frac{\beta + n_{x_{(i,s)},k}^{-(i,s)}}{\sum_w \beta + n_{w,k}^{-(i,s)}} \right) \eta_{s,k}^{-(i)}.$$

We use expectation-maximization to update the variational parameters and optimize a logistic regression model. The model's convergence will be evaluated with the following evidence lower-bound

$$ELBO = \mathbb{E}_{q(z)}[\log\ p(z|\alpha)] + \mathbb{E}_{q(z)}[\log\ p(x|z)]$$
$$+ \mathbb{E}_{q(z)}[\log\ p(y_{s,s'} = 1|z, \eta, \upsilon)] - \mathbb{E}_{q(z)}[\log\ q(z)].$$

The inferred $\hat{\theta}_s$ and $\hat{\theta}_{s'}$ can then be used as input to the classifier for predictions on unseen pairs of spots.

Neighbor Prediction

To train SpaTM-R for spot proximity and neighbor prediction, we define spots as neighbors if they are within a user-defined Euclidean distance from spot s. This results in a large class imbalance (i.e., most spots are not a neighbor of any given spot). To address this, we sub-sample the non-neighbor spots, assigning sampling probabilities that are proportional to their distance from the current spot.

Spatial clustering

We evaluated whether SpaTM-R can recapitulate spatial information from gene expression. We first ran PCA on the spot proximity matrix predicted by SpaTM-R followed by the Leiden community detection algorithm to generate clusters. If the number of regions is known, we fine-tune the Leiden resolution parameter to produce the desired number of clusters. For the ease of visualization, we smooth the cluster labels by updating the label of any spot based on the label of its neighboring spots using majority voting.

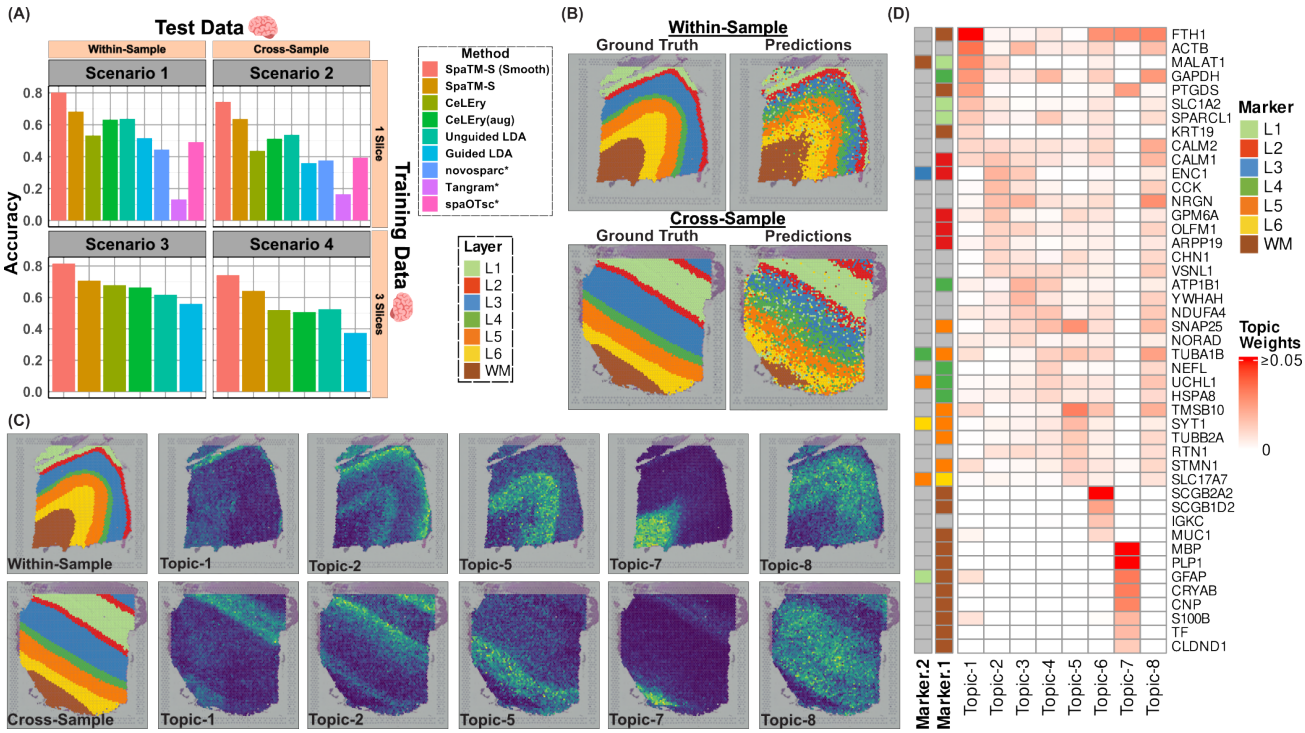

Fig. 2. SpaTM-S enables accurate and interpretable histology annotation predictions. A) Within-sample and cross-sample DLPFC layer predictions. Models were trained on one slice to predict on another from the same patient (Scenario 1) and different patient (Scenario 2), and were separately trained on three slices to predict on the same test slices (Scenarios 3 and 4). Asterisks (*) denote methods which can only be trained on a single tissue slice. Results for SpaTM-S's label smoothing (Smooth) and CeLEry's data augmentation procedures (aug) were also included. B) Visualization of SpaTM-S predictions from scenarios 3 and 4. C) Visualization of the inferred spot-topic mixtures from scenarios 3 and 4. D) Gene topics learned by SpaTM-S for scenario 3 and 4. For each topic, the top 10 genes are displayed. The color intensity is proportional to the probabilities, with values greater than or equal to 0.05 represented as the maximum intensity. Known layer-specific marker genes are annotated on the left. The Marker.1 (Marker.2) annotation represents the layer with the highest (second highest) differential expression (if any) for the corresponding gene as reported by Maynard et al. (2021).

Results

Benchmarking SpaTM-S

We evaluated SpaTM-S in predicting spatial cortical layers (L1-L6 and White Matter) in human dorsolateral prefrontal cortex (DLPFC) brain slices using the SpatialLIBD 10x Visium dataset (Maynard et al., 2021). Benchmarking against state-of-the-art methods involved four scenarios: training on one slice and predicting on another from the same patient (scenario 1) or a different patient (scenario 2), and training on three slices to predict on the same test slices (scenarios 3 and 4 respectively). We compared SpaTM-S's performance to the results reported by Zhang et al. (2023) for the same scenarios as well as sequentially training a topic model (i.e., Cell-type Guided with single nuclei training data from Russell et al. (2024) and Unguided as the standard topic model) and a post-hoc classifier (Fig. 2A). SpaTM-S accurately predicts spatial labels from spot-based gene expression profiles and outperformed existing methods for within-sample and cross-sample prediction (Fig. 2A,B) (Zhang et al., 2023; Cang and Nie, 2020; Biancalani et al., 2021; Moriel et al., 2021). Updating SpaTM-S's spatial label predictions based on neighboring spots via majority voting (Smooth) further improved its predictive performance. These results demonstrate the reliability of SpaTM-S and the added value of the STM training procedure compared to sequentially training a topic model and classifier.

SpaTM-S Interpretability

A key advantage of the SpaTM-S framework is the ability to identify spatial gene programs using the inferred topic matrices $\hat{\theta}$ and $\hat{\phi}$. Visualizing the spot-topic mixtures $\hat{\theta}_s$ for each spot from unseen histology slices revealed spatial patterns that strongly align with the ground-truth layer regions (Fig. 2C). These patterns were present in both the within and cross-sample scenarios demonstrating the generalizability of the inferred topics. Next, we selected the top genes for each inferred topic in $\hat{phi}$ and compared them to the layer-specific markers identified in the SpatialLIBD study. We found considerable overlap, further supporting the accuracy of the spatial gene programs learned by SpaTM-S (Fig. 2D) (Maynard et al., 2021).

Spatially Informed Gene Expression Correction

Sparsity and noise in ST data can impact the visibility of spatial gene expression patterns when defining region markers (Maynard et al., 2021). A solution is to apply pseudo-bulking to spots in different spatial regions to reduce noise in marker expression. However, the loss of spot-level resolution limits our ability to explore how marker expression changes across region borders. Instead, we propose to use the spatially-informed topic mixtures learned by SpaTM-S to correct spatial expression profiles (Fig. 3A). Taking the matrix product of the inferred spot-topic $\hat{\theta}$ and gene-topic $\hat{\phi}$ mixtures yields a spot-by-gene matrix $\hat{X}$ that represents the reconstructed gene expression of a given sample. We trained SpaTM-S on three

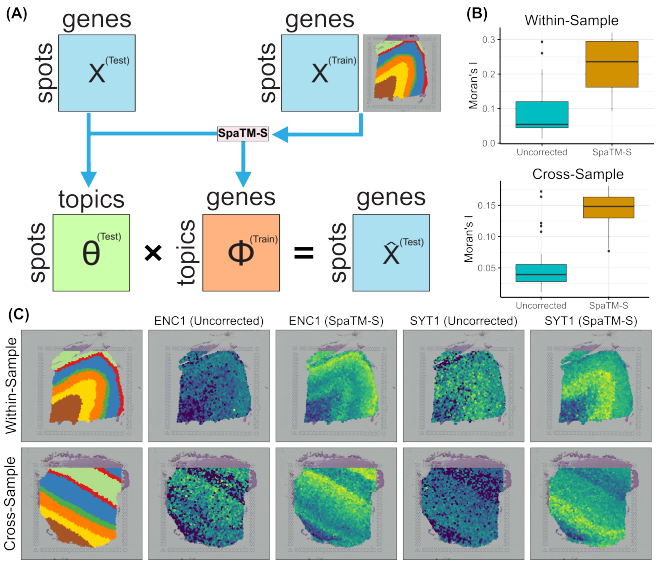

Fig. 3. SpaTM-S enables spatially informed gene expression correction. A) Spot-level gene expression can be corrected with SpaTM-S by taking the matrix product of the inferred spot-topic and gene-topic mixtures. B) Comparing the spatial autocorrelation (Moran's I) estimates of previously reported DLPFC markers using log-normalized counts from the uncorrected and SpaTM-S-corrected expression profiles for a within-sample and cross-sample slice. C) Visualization of spatial expression patterns for DLPFC markers using log-normalized counts from the uncorrected and SpaTM-S-corrected expression profiles.

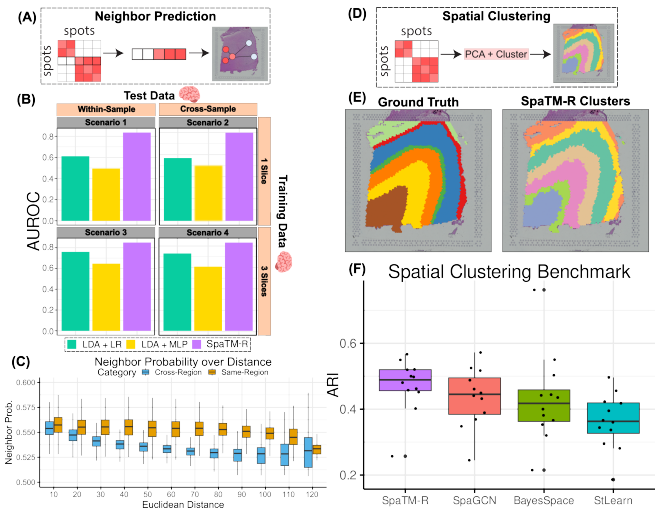

Fig. 4. Training SpaTM-R for neighbor prediction yields reliable performance and high-quality spatial clustering. A) We train SpaTM-R to predict neighbors such that we can infer a spot proximity matrix and neighbors in unseen samples. B) Benchmarking results for neighbor prediction for SpaTM-R and independently training the LDA and classifier components. We train on one slice and predict on another from the same patient (Scenario 1) and different patient (Scenario 2), and separately train on three slices to predict on the same test slices (Scenarios 3 and 4). C) Comparison of neighbor predictions for pairs of spots in the same and different regions over different Euclidean distances. D) SpaTM-R uses the predicted spot proximity to generate spatial clusters. E) Visualization of smoothened SpaTM-R spatial clustering results for one DLPFC slice. F) ARI benchmarking for SpaTM-R against state-of-the-art methods on the SpatialLIBD dataset.

slices and inferred $\hat{\theta}$ for within and cross-sample test slices. After reconstructing the spot-by-gene matrix, we scaled the expression profile of each spot by its library size and applied log-normalization. The corrected expression profiles improved the spatial autocorrelation (Moran's I) of layer markers (identified by pseudo-bulking in Maynard et al. (2021)) in comparison to the uncorrected counts (Fig. 3B). The reconstructed profiles also exhibited much clearer spatial patterns (Fig. 3C).

Benchmarking SpaTM-R

SpaTM-R complements SpaTM-S by enabling the prediction of a spot proximity matrix for neighbor prediction 4A). To evaluate neighbor prediction, we trained SpaTM-R in four scenarios: training on one slice and predicting on another from the same patient (scenario 1) or a different patient (scenario 2), and training on three slices to predict on the same test slices (scenarios 3 and 4 respectively). Neighbors were defined as any spot within a given Euclidean distance of a given spot s. For benchmarking purposes, we used a distance threshold, which provided approximately 10-12 neighbors for each spot. To ensure a balanced training and testing dataset, we then sampled an equal number of spots beyond the threshold distance to use as negative examples. We set our sampling procedure to prioritize spots that were further from the current spot s. We then compared SpaTM-R's performance to running a logistic regression and 1-layer MLP on the topic mixtures generated from a standard topic model (LDA). Neighbor prediction evaluation on the generated positive and negative examples was done using AUROC, which revealed the superior performance of SpaTM-R (Fig. 4B). To determine how SpaTM-R handles predicting neighbors within and across spatial regions, we compared the neighbor prediction for within-region and cross-region pairs of spots over different Euclidean distances (Fig. 4C). This revealed that SpaTM-R was more inclined to predict same-region pairs of spots as neighbors than cross-region pairs as their distance increased, suggesting that the neighbor predictions could reveal global spatial patterns in the sample.

Thus, we evaluated the ability for SpaTM-R to generate spatial clusters without using manual annotations (Fig. 4D). We trained SpaTM-R to predict neighbors on each slice from the SpatialLIBD dataset and used the predicted spot proximity matrices to run PCA and Leiden Community Detection to generate clusters. A visual inspection of the generated clusters reveals their close resemblance to the ground-truth layer labels (Fig. 4E). We then compared these results to previously reported benchmarks for state-of-the-art methods on the same 12 DLPFC slices (Hu et al., 2021). SpaTM-R demonstrated competitive performance to existing methods while retaining its interpretable framework (Fig. 4F).

Integration of SpaTM Results

A crucial advantage of SpaTM is its ability to perform multiple tasks without requiring changes in our inductive biases. This makes the integration of results from its different tasks more reliable and intuitive than using a pipeline of independent methods. We demonstrate this using one DLPFC slice from the SpatialLIBD database to integrate results found using the SpaTM-S, SpaTM-R and SpaTM-G components. This revealed that white matter spots (as determined by SpaTM-S) were represented by two clusters found by SpaTM-R (cluster 2 and 4) (Fig. 5A). We found that cluster 4 represented white matter spots bordering layer 6 whereas cluster 2 represented the core white matter region. Comparing the SpaTM-S-derived topic

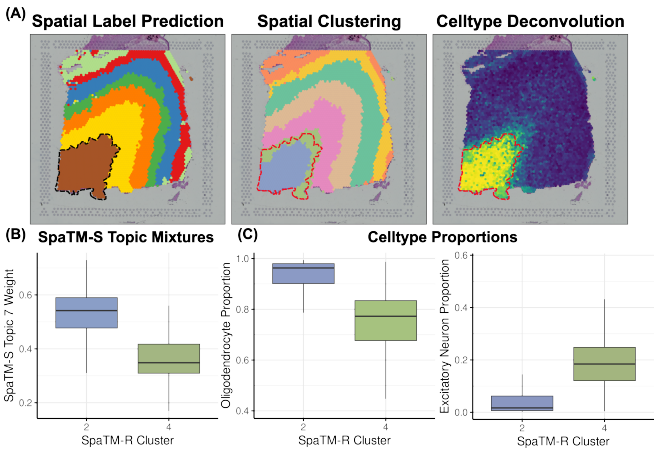

Fig. 5. SpaTM enables efficient and intuitive integration of spatial analysis results. A) Results from the SpaTM suite for one DLPFC slice. SpaTM-S (left) provides the spatial label predictions, SpaTM-R (center) provides spatial clusters and SpaTM-G (right) provides cell type proportions (visualization shows oligodendrocytes). The dotted line delineates the predicted white matter spots. B) Comparison of SpaTM-S-derived topic proportions across white matter sub-regions defined by SpaTM-R. C) Cell-type Deconvolution results (SpaTM-G) comparing cell type proportions in SpaTM-R white matter sub-regions using SpaTM-G.

mixtures between the white matter spots in both clusters shows a clear decrease in the white matter-associated topic weight (Fig. 5B). Finally, running a cell-type deconvolution analysis by training SpaTM-G on healthy DLPFC samples from a reference snRNA-seq dataset revealed that the core white matter cluster had a higher proportion of oligodendrocytes whereas the border cluster revealed the appearance of excitatory neurons (Fig. 5C) (Maitra et al., 2023). Together, our proposed SpaTM suite has greatly facilitated integrating the results from all three tasks and has identified that the predicted white matter region (as predicted by SpaTM-S) is composed of two domains (as predicted by SpaTM-R) which harbor distinct cell type proportions (as predicted by SpaTM-G).

Imputing Spatial Features for MDD snRNA-seq

In addition to providing a flexible analysis toolkit for ST datasets, SpaTM can be used to impute spatial information into single cell atlases, where for many diseases, large-scale ST datasets are yet to be developed. We acquired a public snRNA-seq dataset from the DLPFC of post-mortem human brains diagnosed with MDD or otherwise healthy (Maitra et al., 2023). The dataset comprised of 160k nuclei from 71 patients and had collected male and female samples in separate batches. Here, we applied both the SpaTM-S and SpaTM-R models, trained on the SpatialLIBD data to impute spatial labels and cell proximity information within the atlas (Fig. 6A). Analyzing the imputed cell proximity matrix revealed strong clustering by both cell types and imputed layers demonstrating the reliability and coherence between the SpaTM-S and SpaTM-R imputations (Fig. 6B). In addition, jointly ordering the cell proximity matrix by layers and cell type reveals that imputed layers are largely composed of biologically relevant cell types and retain intra-layer cell-type-specific patterns. As male and female samples were collected in different batches, we ran a sex-specific analysis to compare differential layer proportions at a cell-type-specific level (Fig. 6C). In females, we found that MDD cases had a higher proportion of oligodendrocyte progenitor cells (OPC) in layer 1 (L1) and a shift of inhibitory neurons (InN) from layer 3 (L3) to layer 2 (L2). We also observed that both males and females with MDD exhibited a higher proportion of microglia (Mic) in layer 1 than control samples. A Mann-Whitney U test on all layer-cell-type combinations revealed the above results to be significant. In general, MDD samples exhibited a higher level of variability in layer proportions compared to healthy controls.

Discussion

Through our analysis, we have demonstrated that SpaTM provides a competitive solution for a suite a spatial transcriptomics analysis tasks. The inferred spot-topic and gene-topic mixtures improve inference on spatial labels and spatial clustering by providing interpretable spatially-oriented gene programs. We also show how topic mixtures can be used to reconstruct spatially informed expression profiles that resolve noise and sparsity issues in ST data, avoiding the need for pseudo-bulking and retaining spot-level resolution for marker exploration. In addition, the proposed SpaTM suite provides a unified topic model framework to tackle different tasks to facilitate an integrative ST analysis. In contrast, existing pipelines often require running multiple methods with limited transferability due to their distinct model assumptions. As demonstrated by our analysis of the SpatialLIBD DLPFC data, we are able to characterize the white matter sub-regions by integrating the predictions and topic mixtures inferred from the three topic modeling components provided by SpaTM (Fig. 4). Finally, our analysis of the MDD snRNA-seq atlas reveals how spatial imputation can extend single-cell analyses. The existing literature suggests that the observed shifts in the proportions of the cell type-specific layer may be related to neuronal damage, with depletion of OPCs suggested to contribute to depressive-like behavior (Fernandez-Castaneda and Gaultier, 2016). These findings demonstrate how spatial imputation can provide a clearer explanation for MDD by providing a way to integrate spatially driven changes with gene expression.

Conclusion

SpaTM is a flexible and interpretable analysis framework for spatial transcriptome analyses. It enables cell-type deconvolution, spatial label prediction, expression correction, cell proximity inference and spatial clustering with a consistent inductive bias. In addition to its quantitative performance, SpaTM surpasses previous methods by enabling rapid inference and interpretation of spatially driven gene programs. In future work, SpaTM can be extended to infer cell-cell communication from the predicted cell-cell proximity using methods such as CellAgentChat (Raghavan et al., 2024), which would help identify spatial subdomains with distinct interaction patterns in the cellular microenvironment.

Code availability

Package: `https://github.com/li-lab-mcgill/SpaTM`
Analysis: `https://github.com/aosakwe/SpaTM_Analysis`

Competing interests

No competing interest is declared.

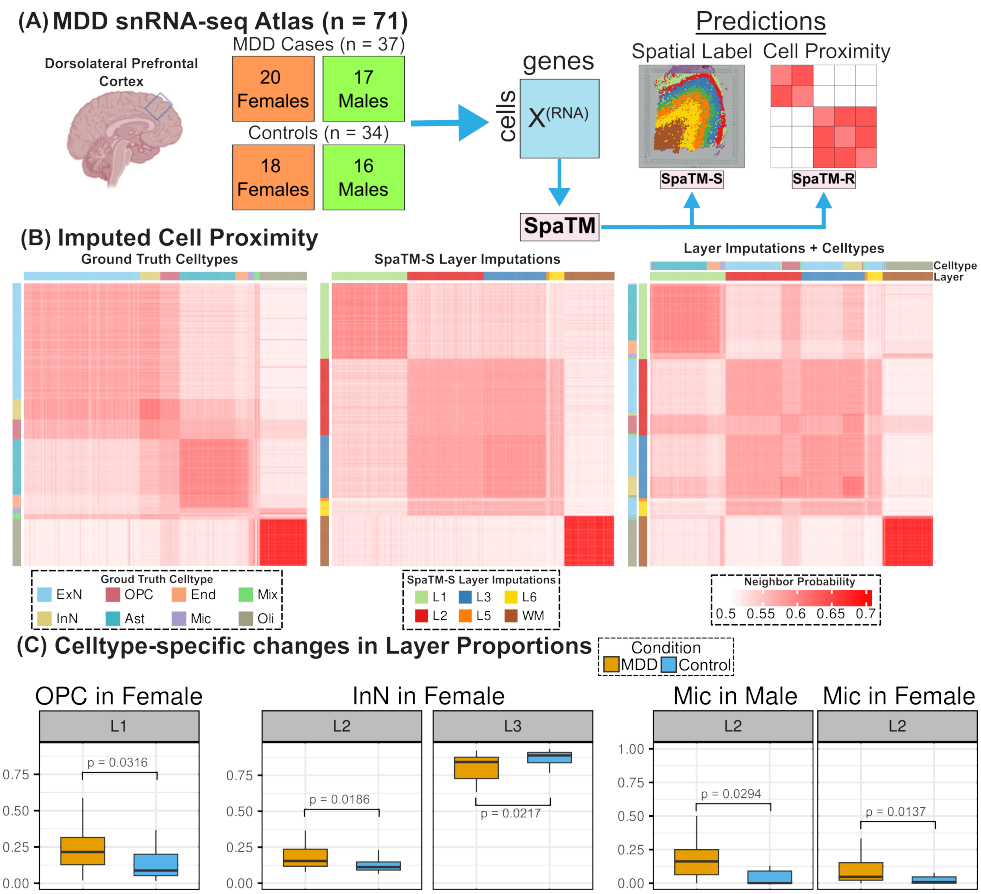

Fig. 6. Sex-specific spatial analysis of snRNA-seq MDD using SpaTM spatial imputations. A) Large-scale DLPFC snRNA-seq atlas for MDD. We use SpaTM to impute spatial labels (SpaTM-S) and cell proximity (SpaTM-R). B) Heatmap of cell proximity matrix imputed by SpaTM-R for one healthy female sample. The cells in the heatmap are ordered by celltype (left) SpaTM-S layer predictions (center) and celltype nested within layer predictions (right). C) Assessing sex-specific differences in celltype proportions between healthy and MDD samples in different cortical layers using SpaTM-S predictions.

Acknowledgments

Y.L. is supported by Canada Research Chair (Tier 2) in Machine Learning for Genomics and Healthcare (CRC-2021-00547) and Natural Sciences and Engineering Research Council (NSERC) Discovery Grant (RGPIN-2016-05174). Q.Z is supported by Fonds de recherche du Québec Scholar (Junior 1), NSERC and the Fonds de recherche du Québec - Santé (FRQS) program. A.O. is supported by training scholarships from the Canadian Institutes of Health Research (CIHR) and FRQS. We thank Doruk Cakmakci and Anjali Chawla for their helpful feedback on data visualizations and the MDD analysis.

References

T. Biancalani, G. Scalia, L. Buffoni, R. Avasthi, Z. Lu, A. Sanger, N. Tokcan, C. R. Vanderburg, Segerstolpe, M. Zhang, I. Avraham-Davidi, S. Vickovic, M. Nitzan, S. Ma, A. Subramanian, M. Lipinski, J. Buenrostro, N. B. Brown, D. Fanelli, X. Zhuang, E. Z. Macosko, and A. Regev. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nature Methods*, 18(11): 1352–1362, Nov. 2021. ISSN 1548-7105. doi: 10.1038/s41592-021-01264-7. URL `https://www.nature.com/articles/s41592-021-01264-7`. Publisher: Nature Publishing Group.

D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3 (null):993–1022, Mar. 2003. ISSN 1532-4435.

B. Bonev, G. Castelo-Branco, F. Chen, S. Codeluppi, M. R. Corces, J. Fan, M. Heiman, K. Harris, F. Inoue, M. Kellis, A. Levine, M. Lotfollahi, C. Luo, K. R. Maynard, M. Nitzan, V. Ramani, R. Satijia, L. Schirmer, Y. Shen, N. Sun, G. S. Green, F. Theis, X. Wang, J. D. Welch, O. Gokce, G. Konopka, S. Liddelow, E. Macosko, O. Ali Bayraktar, N. Habib, and T. J. Nowakowski. Opportunities and challenges of single-cell and spatially resolved genomics methods for neuroscience discovery. *Nature Neuroscience*, 27(12):2292–2309, Dec. 2024. ISSN 1546-1726. doi: 10.1038/s41593-024-01806-0. URL `https://www.nature.com/articles/s41593-024-01806-0`. Publisher: Nature Publishing Group.

Z. Cang and Q. Nie. Inferring spatial and signaling relationships between cells from single cell transcriptomic data. *Nature Communications*, 11(1):2084, Apr. 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-15968-5. URL `https://www.nature.com/articles/s41467-020-15968-5`. Publisher: Nature Publishing Group.

J. Chang and D. Blei. Relational Topic Models for Document Networks. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, pages 81–88. PMLR, Apr. 2009. URL `https://proceedings.mlr.`

press/v5/chang09a.html. ISSN: 1938-7228.

A. Fernandez-Castaneda and A. Gaultier. Adult oligodendrocyte progenitor cells - multifaceted regulators of the CNS in health and disease. *Brain, behavior, and immunity*, 57:1–7, Oct. 2016. ISSN 0889-1591. doi: 10.1016/j.bbi.2016.01.005. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4940337/.

J. Hu, X. Li, K. Coleman, A. Schroeder, N. Ma, D. J. Irwin, E. B. Lee, R. T. Shinohara, and M. Li. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nature Methods*, 18(11): 1342–1351, Nov. 2021. ISSN 1548-7105. doi: 10.1038/s41592-021-01255-8. URL https://www.nature.com/articles/s41592-021-01255-8. Publisher: Nature Publishing Group.

M. Maitra, H. Mitsuhashi, R. Rahimian, A. Chawla, J. Yang, L. M. Fiori, M. A. Davoli, K. Perlman, Z. Aouabed, D. C. Mash, M. Suderman, N. Mechawar, G. Turecki, and C. Nagy. Cell type specific transcriptomic differences in depression show similar patterns between males and females but implicate distinct cell types and genes. *Nature Communications*, 14(1):2912, May 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-38530-5. URL https://www.nature.com/articles/s41467-023-38530-5. Publisher: Nature Publishing Group.

K. R. Maynard, L. Collado-Torres, L. M. Weber, C. Uytingco, B. K. Barry, S. R. Williams, J. L. Catallini, M. N. Tran, Z. Besich, M. Tippani, J. Chew, Y. Yin, J. E. Kleinman, T. M. Hyde, N. Rao, S. C. Hicks, K. Martinowich, and A. E. Jaffe. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nature Neuroscience*, 24(3):425–436, Mar. 2021. ISSN 1546-1726. doi: 10.1038/s41593-020-00787-0. URL https://www.nature.com/articles/s41593-020-00787-0. Publisher: Nature Publishing Group.

J. Mcauliffe and D. Blei. Supervised Topic Models. In *Advances in Neural Information Processing Systems*, volume 20. Curran Associates, Inc., 2007. URL https://proceedings.neurips.cc/paper_files/paper/2007/hash/d56b9fc4b0f1be8871f5e1c40c0067e7-Abstract.html.

B. F. Miller, F. Huang, L. Atta, A. Sahoo, and J. Fan. Reference-free cell type deconvolution of multi-cellular pixel-resolution spatially resolved transcriptomics data. *Nature Communications*, 13(1):2339, Apr. 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-30033-z. URL https://www.nature.com/articles/s41467-022-30033-z. Publisher: Nature Publishing Group.

N. Moriel, E. Senel, N. Friedman, N. Rajewsky, N. Karaiskos, and M. Nitzan. NovoSpaRc: flexible spatial reconstruction of single-cell gene expression with optimal transport. *Nature Protocols*, 16(9):4177–4200, Sept. 2021. ISSN 1750-2799. doi: 10.1038/s41596-021-00573-7.

L. Moses and L. Pachter. Museum of spatial transcriptomics. *Nature Methods*, 19(5):534–546, May 2022. ISSN 1548-7105. doi: 10.1038/s41592-022-01409-2. URL https://www.nature.com/articles/s41592-022-01409-2. Publisher: Nature Publishing Group.

A. Neufeld, L. L. Gao, J. Popp, A. Battle, and D. Witten. Inference after latent variable estimation for single-cell RNA sequencing data. *Biostatistics*, 25(1):270–287, Jan. 2024. ISSN 1465-4644. doi: 10.1093/biostatistics/kxac047. URL https://doi.org/10.1093/biostatistics/kxac047.

V. Raghavan, Y. Li, and J. Ding. Harnessing Agent-Based Modeling in CellAgentChat to Unravel Cell-Cell Interactions from Single-Cell Data. Mar. 2024. doi: 10.1101/2023.08.23.554489. URL https://www.biorxiv.org/content/10.1101/2023.08.23.554489v2. Pages: 2023.08.23.554489 Section: New Results.

D. Righelli, L. M. Weber, H. L. Crowell, B. Pardo, L. Collado-Torres, S. Ghazanfar, A. T. L. Lun, S. C. Hicks, and D. Risso. SpatialExperiment: infrastructure for spatially-resolved transcriptomics data in R using Bioconductor. *Bioinformatics (Oxford, England)*, 38(11):3128–3131, May 2022. ISSN 1367-4811. doi: 10.1093/bioinformatics/btac299.

A. J. C. Russell, J. A. Weir, N. M. Nadaf, M. Shabet, V. Kumar, S. Kambhampati, R. Raichur, G. J. Marrero, S. Liu, K. S. Balderrama, C. R. Vanderburg, V. Shanmugam, L. Tian, J. B. Iorgulescu, C. H. Yoon, C. J. Wu, E. Z. Macosko, and F. Chen. Slide-tags enables single-nucleus barcoding for multimodal spatial genomics. *Nature*, 625 (7993):101–109, Jan. 2024. ISSN 1476-4687. doi: 10.1038/s41586-023-06837-4. URL https://www.nature.com/articles/s41586-023-06837-4. Publisher: Nature Publishing Group.

C. A. Stockmeier, G. J. Mahajan, L. C. Konick, J. C. Overholser, G. J. Jurjus, H. Y. Meltzer, H. B. M. Uylings, L. Friedman, and G. Rajkowska. Cellular changes in the postmortem hippocampus in major depression. *Biological Psychiatry*, 56(9):640–650, Nov. 2004. ISSN 0006-3223. doi: 10.1016/j.biopsych.2004.08.022.

L. S. Swapna, M. Huang, and Y. Li. GTM-decon: guided-topic modeling of single-cell transcriptomes enables sub-cell-type and disease-subtype deconvolution of bulk transcriptomes. *Genome Biology*, 24(1):190, Aug. 2023. ISSN 1474-760X. doi: 10.1186/s13059-023-03034-4. URL https://doi.org/10.1186/s13059-023-03034-4.

Y. Teh, D. Newman, and M. Welling. A collapsed variational bayesian inference algorithm for latent dirichlet allocation. *Advances in neural information processing systems*, 19, 2006.

L. Tian, F. Chen, and E. Z. Macosko. The expanding vistas of spatial transcriptomics. *Nature Biotechnology*, 41 (6):773–782, June 2023. ISSN 1546-1696. doi: 10.1038/s41587-022-01448-2. URL https://www.nature.com/articles/s41587-022-01448-2. Publisher: Nature Publishing Group.

Q. Zhang, S. Jiang, A. Schroeder, J. Hu, K. Li, B. Zhang, D. Dai, E. B. Lee, R. Xiao, and M. Li. Leveraging spatial transcriptomics data to recover cell locations in single-cell RNA-seq with CeLEry. *Nature Communications*, 14 (1):4050, July 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-39895-3. URL https://www.nature.com/articles/s41467-023-39895-3. Publisher: Nature Publishing Group.

E. Zhao, M. R. Stone, X. Ren, J. Guenthoer, K. S. Smythe, T. Pulliam, S. R. Williams, C. R. Uytingco, S. E. B. Taylor, P. Nghiem, J. H. Bielas, and R. Gottardo. Spatial transcriptomics at subspot resolution with BayesSpace. *Nature Biotechnology*, 39(11):1375–1384, Nov. 2021a. ISSN 1546-1696. doi: 10.1038/s41587-021-00935-2. URL https://www.nature.com/articles/s41587-021-00935-2. Publisher: Nature Publishing Group.

Y. Zhao, H. Cai, Z. Zhang, J. Tang, and Y. Li. Learning interpretable cellular and gene signature embeddings from single-cell transcriptomic data. *Nature Communications*, 12(1):5261, Dec. 2021b. ISSN 2041-1723. doi: 10.1038/s41467-021-25534-2. URL https://www.nature.com/articles/s41467-021-25534-2.