

# Unsupervised Spatially Embedded Deep Representation of Spatial Transcriptomics

Jinmiao Chen (✉ [Chen\\_Jinmiao@immunol.a-star.edu.sg](mailto:Chen_Jinmiao@immunol.a-star.edu.sg))

Singapore Immunology Network <https://orcid.org/0000-0001-7547-6423>

**Huazhu Fu**

Inception Institute of Artificial Intelligence <https://orcid.org/0000-0002-9702-5524>

**Hang XU**

**Kelvin Chong**

Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A\*STAR), 8A Biomedical Grove, Immunos Building, 138648, Singapore <https://orcid.org/0000-0002-4473-6453>

**Mengwei Li**

Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A\*STAR), 8A Biomedical Grove, Immunos Building, 138648, Singapore

**Kok Siong Ang**

Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A\*STAR), 8A Biomedical Grove, Immunos Building, 138648, Singapore

**Hong Kai Lee**

Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A\*STAR), 8A Biomedical Grove, Immunos Building, 138648, Singapore

**Jingjing Ling**

Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A\*STAR), 8A Biomedical Grove, Immunos Building, 138648, Singapore

**Ao Chen**

BGI-Shenzhen

**Ling Shao**

Inception Institute of Artificial Intelligence, Abu Dhabi, United Arab Emirates

**Longqi Liu**

BGI-Shenzhen, Shenzhen, China

---

## Article

**Keywords:** spatial transcriptomics, graph convolutional network, gene expression, deep learning

**Posted Date:** July 6th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-665505/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

1 **Unsupervised Spatially Embedded Deep Representation of Spatial**  
2 **Transcriptomics**

3

4 Huazhu Fu<sup>1,\*</sup>, Hang Xu<sup>2,\*</sup>, Kelvin Chong<sup>2</sup>, Mengwei Li<sup>2</sup>, Kok Siong Ang<sup>2</sup>, Hong Kai Lee<sup>2</sup>,  
5 Jingjing Ling<sup>2</sup>, Ao Chen<sup>3</sup>, Ling Shao<sup>1</sup>, Longqi Liu<sup>3</sup>, Jinmiao Chen<sup>2,†</sup>

6

7 <sup>1</sup> Inception Institute of Artificial Intelligence, Abu Dhabi, United Arab Emirates

8 <sup>2</sup> Singapore Immunology Network (SIgN), Agency for Science, Technology and Research  
9 (A\*STAR), 8A Biomedical Grove, Immunos Building, 138648, Singapore

10 <sup>3</sup> BGI-ShenZhen, Shenzhen 518103, China.

11

12 \*These authors contributed equally to this work.

13 †Corresponding author. Email: chen\_jinmiao@immunol.a-star.edu.sg

14

15

16 **Key words:** spatial transcriptomics; graph convolutional network; gene expression; deep  
17 learning

18

19

## 20 Abstract

21 Spatial transcriptomics enable us to dissect tissue heterogeneity and map out inter-cellular  
22 communications. Optimal integration of transcriptomics data and associated spatial  
23 information is essential towards fully exploiting the data. We present SEDR, an unsupervised  
24 spatially embedded deep representation of both transcript and spatial information. The SEDR  
25 pipeline uses a deep autoencoder to construct a low-dimensional latent representation of gene  
26 expression, which is then simultaneously embedded with the corresponding spatial  
27 information through a variational graph autoencoder. We applied SEDR on human dorsolateral  
28 prefrontal cortex data and achieved better clustering accuracy, and correctly retraced the  
29 prenatal cortex development order with trajectory analysis. We also found the SEDR  
30 representation to be eminently suited for batch integration. Applying SEDR to human breast  
31 cancer data, we discerned heterogeneous sub-regions within a visually homogenous tumor  
32 region, identifying a tumor core with pro-inflammatory microenvironment and an outer ring  
33 region enriched with tumor associated macrophages which drives an immune suppressive  
34 microenvironment.

35

36

37

38

## 39 Introduction

40 Single-cell omics technologies enable measurements at single-cell resolution, and have led to  
41 discoveries of new subpopulations across various tissues, in both healthy and diseased states.  
42 However, the dissociation of tissue into single cells prior to high throughput omics data  
43 acquisition leads to cellular spatial information being lost, hindering our ability to dissect the  
44 spatial organization and intercellular interactions of individual cells. While computational tools  
45 have been developed to predict cell-cell interactions from ligand and receptor expression, they  
46 require validation using immunohistochemistry (IHC) or immunofluorescence (IF). Emerging  
47 spatial omics technologies overcome these limitations through the simultaneous  
48 measurement of gene/protein expression and spatial locations of cells. Such spatially resolved  
49 transcriptomes of histological tissues enable the reconstruction of tissue architecture and cell-  
50 cell interactions.<sup>1,2,3,4,5,6,7,8,9</sup> This approach has proven valuable in many applications including  
51 studies on brain disorders,<sup>2,10</sup> tumour microenvironments,<sup>3,11</sup> and embryonic development.<sup>12</sup>

52 Among currently available spatial transcriptomics approaches, *in situ* capturing-based  
53 technologies such as 10x Genomics Visium and Nanostring GeoMX DSP have gained  
54 popularity owing to their accessibility and ability to profile a large number of mRNA targets  
55 within each spot. In principle, a histological section from a tissue sample is permeabilized and  
56 the released mRNA is captured by either spatially arrayed oligos on slide surfaces or by pre-  
57 hybridized RNA-target barcodes in manually defined regions of interest (ROIs). However, both  
58 technologies suffer from mRNA capture area limitations, with the smallest diameter typically  
59 being ~50µm, which is larger than a single cell. To overcome this, several computational  
60 methods have been developed to deconvolve the cell mixture of the spatial  
61 spot.<sup>13,14,15,16,17,18,19,20</sup> Recently, improvements in mRNA capture methods have led to smaller  
62 subcellular capture areas that are ~1-10µm in diameter. These high-resolution spatial  
63 transcriptomics methods can obtain spatially resolved transcriptomes with increased spatial  
64 fidelity, without compromising the number of genes captured. They include Slide-seq,<sup>8</sup> DBiT-  
65 seq,<sup>9</sup> Stereo-seq,<sup>5</sup> PIXEL-seq,<sup>6</sup> and Seq-Scope,<sup>7</sup> with the highest resolution (~1µm) thus far

66 obtained by the latter three. These submicrometer-resolution methods usually require voxel  
67 binning or cell segmentation to produce a gene-by-cell expression matrix for downstream  
68 analysis. Capture area sizes have also improved and thus increased the overall cell  
69 throughput, necessitating new computational methods that can handle big spatial data.

70         When analyzing spatial transcriptomics data, combining both gene expression and  
71 spatial information to learn a discriminative representation for each cell or spot is crucial.  
72 However, established workflows, e.g., Seurat,<sup>21</sup> still employ pipelines designed for single-cell  
73 RNA-seq analysis, which primarily focus on gene expression data and ignore the structural  
74 relationship of the spatial neighborhood. Recently, several new methods have been developed  
75 for spatial transcriptomics to overcome this limitation. For example, BayesSpace<sup>22</sup> starts from  
76 a Markov random field (MRF) prior which hypothesizes that spots belonging to the same cell  
77 type should be closer to one another, and updates the model with a Bayesian approach.  
78 Giotto<sup>23</sup> implements a hidden Markov random field (HMRF) model to detect domains with  
79 coherent patterns by comparing gene expression between cells and their neighbors.  
80 SpaGCN<sup>24</sup> combines spatial distances and histological dissimilarities to construct a weighted  
81 graph of spots, and then integrates the graph with gene expression using a graph  
82 convolutional network (GCN) to cluster the spots. stLearn<sup>25</sup> utilizes a deep learning model on  
83 the spot images to extract morphological features, on which morphological distances are  
84 calculated. It then uses the morphological distance and spatial neighborhood information to  
85 normalize the gene expression of each spot based on its identified neighbors. The normalized  
86 gene expression is then used as input for linear principal component analysis (PCA), followed  
87 by uniform manifold approximation and projection (UMAP), and spatial clustering. Notably,  
88 these methods mainly rely on PCA to extract the highly variable features of gene expression  
89 data, which involves a linear transformation, so they are unable to model complex non-linear  
90 relationships. While stLearn does utilize deep learning, it is only applied to the image modality,  
91 and the model still relies on linear PCA to extract features from the spatially normalized gene  
92 expression data. Moreover, these methods do not produce low-dimensional representations  
93 of jointly embedded gene expression and spatial information. The joint embedding of gene

94 expression and spatial information is essential to effectively integrate both modalities for better  
95 visualization, clustering, trajectory inference, and batch integration.

96 In this work, we developed an unsupervised spatially embedded deep representation  
97 (SEDR) method for learning a low-dimensional latent representation of gene expression  
98 embedded with spatial information. Our SEDR model consists of two main components, a deep  
99 autoencoder network for learning a gene representation, and a variational graph autoencoder  
100 network for embedding the spatial information. These two components are optimized jointly to  
101 generate a latent representation for spatial transcriptomics data analysis. We applied SEDR  
102 on the 10x Genomics Visium spatial transcriptomics and Stereo-seq datasets and  
103 demonstrated its ability to achieve better representations for various follow-up analysis tasks,  
104 including clustering, visualization, trajectory inference, and batch effects correction.

## 105 Results

### 106 Overview of SEDR.

107 SEDR learns a gene representation in a low-dimensional latent space with jointly embedded  
108 spatial information (Figure 1). Given spatial transcriptomics data, SEDR first learns a nonlinear  
109 mapping from the gene expression space to a low-dimensional feature space using a deep  
110 autoencoder network. Simultaneously, a variational graph autoencoder is utilized to aggregate  
111 the gene representation with the corresponding spatial neighborhood relationships to produce  
112 a spatial embedding. Then, the gene representation and spatial embedding are concatenated  
113 to form the final latent representation used to reconstruct the gene expression. Thereafter, an  
114 unsupervised deep clustering method<sup>26</sup> is employed to enhance the compactness of the  
115 learned latent representation. This iterative deep clustering generates a form of soft clustering  
116 that assigns cluster-specific probabilities to each cell, leveraging on the inferences between  
117 cluster-specific and cell-specific representation learning. Finally, the learned latent  
118 representation can be applied towards various analysis tasks.

119 **Quantitative assessment of SEDR on human dorsolateral prefrontal cortex (DLPFC)**  
120 **dataset.**

121 To perform a quantitative comparison of SEDR with other methods, we downloaded the 10x  
122 Genomics Visium spatial transcriptomics data and the manually annotated layers for LIBD  
123 human dorsolateral prefrontal cortex (DLPFC) data.<sup>2</sup> The LIBD data includes 12 slices from  
124 the human DLPFC, which span six cortical layers plus white matter. We chose this dataset  
125 because the human DLPFC has clear and established morphological boundaries which can  
126 serve as the ground truth. We first applied the standard Seurat pipeline<sup>21</sup> to process and cluster  
127 cells using only expression profiles and set the result as the benchmarking baseline to  
128 investigate the extent to which spatial information improves cell clustering. As Giotto,<sup>23</sup>  
129 stLearn,<sup>25</sup> SpaGCN,<sup>24</sup> and BayesSpace<sup>22</sup> integrate spatial information and RNA-seq data for  
130 clustering, we also applied them with recommended default parameters to the same dataset  
131 for benchmarking against SEDR.

132 In brain slice 151673 (Figure 2A) with 3,639 spots and 33,538 genes, SEDR and  
133 BayesSpace achieved the best performance in terms of both layer borders and adjusted rand  
134 index (ARI). When comparing the results on all 12 DLPFC samples, SEDR had the second  
135 highest mean ARI (0.427) (Figure 2A, bottom right), but the difference between SEDR and the  
136 top performer BayesSpace (0.457) was not significant (Mann-Whitney U Test:<sup>27</sup> p-value=0.55).  
137 It should be noted that BayesSpace's clustering algorithm is optimized for spatial omics, while  
138 SEDR is a dimension reduction method with its objective being to find the best latent  
139 representation. SEDR followed by Leiden clustering, which was not specifically designed or  
140 optimized for clustering spatial omics, achieved comparable clustering performance to  
141 BayesSpace. This indicated that SEDR latent representations effectively integrate gene  
142 expressions and spatial information for capturing inter-cluster differences. Coupling SEDR with  
143 clustering algorithms that are better-suited for spatial omics can be expected to further improve  
144 the clustering accuracy. Furthermore, in contrast to BayesSpace, which does not produce  
145 latent representations, SEDR-derived embeddings can be used for not only clustering but also

146 various downstream analysis tasks such as UMAP visualization, trajectory inference, and  
147 batch effect correction, thus providing more flexibility and utility. Similar to SEDR, SpaGCN  
148 also uses a GCN to process spatial transcriptomics data. Moreover, it incorporates histological  
149 information, which is not included in SEDR. However, the clustering performance of SEDR is  
150 better than that of SpaGCN (Mann-Whitney U Test, p-value < 0.05). stLearn also integrates  
151 histological data, but the performance is likewise poorer. This may indicate that the current  
152 approaches utilized by SpaGCN and stLearn to incorporate histological data are not optimal.  
153 To make full use of the histological information, we may need to treat it as a separate data  
154 modality and use dedicated multi-view algorithms for integration.

155 SEDR generates a set of low-dimensional representation features which can be used  
156 in various downstream analyses, such as trajectory inference.<sup>28</sup> In our experiments, we used  
157 Monocle3<sup>29</sup> to perform trajectory inference on sample 151673 with the Seurat output (RNA-  
158 only) and the low-dimensional SEDR representation features. We found that SEDR showed  
159 significantly improved performance over Seurat (Figure 2B). In the UMAP plot of SEDR's  
160 output, cells belonging to different layers were well-organized, and when we selected white  
161 matter (WM) as the root, the pseudo-time reflected the correct "inside-out" developmental  
162 ordering of cortical layers (Figure 2B). This demonstrated that, compared to RNA-only  
163 analyses, incorporating spatial information enabled SEDR to generate a better latent  
164 representation summarizing the spatial transcriptomics data. We further confirmed our  
165 observations with another trajectory inference method named partition-based graph  
166 abstraction (PAGA),<sup>30</sup> using the SEDR-derived latent space embedding instead of UMAP  
167 coordinates (Figure 2C). The PAGA results showed that adjacent cortical layers tend to share  
168 greater similarity, suggesting that spatial adjacency is linked with transcriptomic and even  
169 functional similarity. Notably, the trajectory was concordant with the chronological order of  
170 cortex development.<sup>31,32,33</sup> We then compared the PAGA graphs generated using the Seurat-  
171 derived principal components and SEDR embeddings. For each of the 12 DLPC slices, we  
172 calculated the ratio of the edge weights between adjacent cortical layers to the total sum of

173 the weights of all edges. We found a significantly higher ratio for SEDR compared to Seurat  
174 (Mann-Whitney U test p-value < 0.05) (Figure 2C, bottom).

### 175 **SEDR corrects for batch effects.**

176 The proliferation of spatial omics applications is generating ever increasing volumes of spatially  
177 resolved omics data across different labs. However, differences in protocols and technologies  
178 complicate comparisons and data integration when trying to achieve consensus on spatially  
179 resolved tissue atlases. As with single-cell RNA-seq (scRNA-Seq), removing batch effects in  
180 spatial omics datasets is a significant challenge. To date, no methods are available for this.  
181 Here, we demonstrate that SEDR can learn joint embeddings across multiple batches and  
182 project them into a shared latent space. Furthermore, SEDR employs a deep embedded  
183 clustering (DEC) loss function that enables it to retain biological variations while reducing  
184 technical variations. We evaluated the batch correcting performance of SEDR on the DLPCF  
185 datasets. We first assessed the batch variations among the twelve datasets and selected three  
186 sets (151507, 151672, 151673) which exhibited substantial batch effects. The common cortical  
187 layers from different batches were separated, as shown in the UMAP plot (Figure 3A). We first  
188 applied Harmony to remove the batch effects due to its superior performance in scRNA-seq  
189 data integration.<sup>34</sup> Harmony was able to mix batches while keeping different layers apart.  
190 However, when we zoomed into the individual layers, we found distinct batch-specific  
191 subclusters, suggesting that the batch effects were not completely removed (Figure 3B). Next,  
192 we tested SEDR and found that the batch effects were substantially reduced (Figure 3C).  
193 Common layers across batches were brought very close and were well-aligned, while different  
194 layers were minimally mixed. Further application of Harmony on the SEDR embeddings evenly  
195 mixed the batches while maintaining separation between layers (Figure 3D). Notably, batch-  
196 specific clusters were no longer present within individual layers. This showed that the  
197 combination of SEDR with Harmony effectively removed the batch effects. Among the other  
198 spatial omics analysis methods, only stLearn is able to produce a latent space embedding  
199 which can be fed to Harmony for batch correction. Therefore, we benchmarked SEDR against

200 stLearn. As stLearn is unable to jointly project different batches to a shared latent space due  
201 to its requirement of histological images as input, we generated latent space embeddings from  
202 each dataset and then concatenated them for Harmony integration. The results showed that  
203 batches were not well mixed and the layers were poorly separated (Figure 3E). In conclusion,  
204 SEDR combined with Harmony outperforms both Harmony alone and stLearn with Harmony,  
205 and can serve as an effective method for batch correction of spatial omics data.

### 206 **Dissecting tumor heterogeneity and immune microenvironments using SEDR.**

207 Intratumoral heterogeneity in cancer complicates effective treatment formulations and is  
208 associated with poor survival prospects.<sup>35</sup> Spatial transcriptomics is an effective tool for  
209 dissecting and characterizing intratumoral heterogeneity and tumor-immune crosstalk. We  
210 tested SEDR on the 10x Visium spatial transcriptomics data for human breast cancer, which  
211 is known for its high intratumoral and intertumoral differences. To aid in the interpretation of  
212 SEDR results, we performed manual pathology labeling based on H&E staining. It should be  
213 noted that, unlike the cerebral cortex which has clear and established morphological  
214 boundaries, tumor tissues are highly heterogeneous and encompass complex  
215 microenvironments. Manual labeling solely based on tumor morphology is inadequate for  
216 characterizing such complexity. Based on pathological features, we manually segmented the  
217 histological image into twenty regions, which we then grouped into four main morphotypes:  
218 ductal carcinoma *in situ*/lobular carcinoma *in situ* (DCIS/LCIS), healthy tissue (Healthy),  
219 invasive ductal carcinoma (IDC), and tumor surrounding regions with low features of  
220 malignancy (Tumor edge) (Figure 4A top left, Supplementary Figure 1A). Visually, all five  
221 methods agreed with the manual annotations at the macroscopic level (Figure 4A).  
222 Nevertheless, the SEDR clusters presented a smoother segmentation compared to other  
223 methods, while those derived by Seurat, stLearn, and SpaGCN appeared fragmented with  
224 irregular boundaries. Notably, SEDR found more sub-clusters within the tumor regions, while  
225 other methods were prone to dividing the healthy regions into subclusters, given that all  
226 methods were set to generate the same number of clusters. Specifically, within the seemingly

227 homogenous tumor region DCIS/LCIS\_3, SEDR separated an outer “ring” (Figure 4A, SEDR  
228 cluster 7) from the tumor core (Figure 4A, SEDR cluster 3). These SEDR clusters indicated  
229 transcriptionally and spatially distinct compartments within the visually homogenous tumor  
230 regions. In addition to clustering analysis, we also employed the Seurat3 ‘anchor’-based  
231 integration workflow to perform probabilistic transfer of annotations from scRNA-seq reference  
232 data for human breast cancer<sup>36</sup> to the spatial data. For each spot, we obtained a probabilistic  
233 classification for each of the scRNA-seq derived classes (Figure 4B, Supplementary Figure  
234 1B). The transferred class probabilities were able to delineate the tumor regions and regions  
235 where immune cells or fibroblasts were present, which were useful for further dissecting the  
236 tumor microenvironment.

237 To further characterize the transcriptional differences between SEDR cluster 3 (tumor  
238 core) and cluster 7 (tumor edge) of DCIS/LCIS\_3 region, we performed differential expression  
239 analysis followed by pathway enrichment analysis (Figure 4C). In cluster 3, we observed the  
240 upregulation of interferon signaling pathways (IFIT1, IFITM1, IFITM3 and TAP1) and NK or  
241 neutrophil activities (FCGR3B and TNFSF10) (Figure 4C, Supplementary Figure 2E). In  
242 addition, RHOB was upregulated in this region, pointing towards reduced metastatic  
243 potential.<sup>40</sup> Cluster 3 represented a region where cancer growth was limited by pro-  
244 inflammatory immune responses. On the other hand, in cluster 7, we observed the presence  
245 of TAMs (Figure 4B, Supplementary Figure 2D), memory B cells (IGHG1, IGHG3, IGHG4,  
246 IGLC2 and IGLC3) and fibroblasts (COL1A1, COL1A2, COL3A1, COL5A1, COL6A1, COL6A2  
247 and FN1) (Figure 4C, Supplementary Figure 2E). Upregulated cathepsin activity (CTSB,  
248 CTSD and CTSZ) and complement pathway (C1QA, C1S) indicated pro-tumor activity by the  
249 TAMs in this region.<sup>41,42,43</sup> Upregulation of actin cytoskeleton signalling also suggested higher  
250 metastasis potential of cluster 7 (Figure 4C). Moreover, upregulated cathepsin activity and  
251 metalloproteinase inhibitors (TIMP1 and TIMP3) also indicated disturbance in the extracellular  
252 matrix integrity (Supplementary Figure 2E). Overall, cluster 7 represented a region with an  
253 immune-suppressed pro-tumor microenvironment and high potential for cancer metastasis.

254 A number of driving forces have been hypothesized as responsible for the metastatic

255 transition of tumor cells from a pre-invasive state to invasive carcinoma, including pro-tumor  
256 immune microenvironments and reduced cell-cell interactions within the tumor.<sup>37</sup> In this study,  
257 we employed PAGA to infer the inter-relatedness between the manually annotated tumor  
258 regions to trace the metastatic transition process. The PAGA graph generated using the SEDR  
259 embeddings suggested that DCIS\_LCIS\_3 was related to the neighboring IDC\_6 region  
260 (Figure 4D). The differentially expressed genes (DEGs) and enriched pathways of  
261 DCIS\_LCIS\_3 compared to all other DCIS\_LCIS regions showed that DCIS\_LCIS\_3 had more  
262 immune infiltrates (Supplementary Figure 2A, 2B, 2C), particularly tumor associated  
263 macrophages (TAMs) (Figure 4B, top), while the other DCIS\_LCIS regions were mainly  
264 comprised of actively dividing/cycling epithelial cells (Figure 4B, bottom) with upregulated  
265 glycolytic and metabolic processes (Supplementary Figure 2C). TAM infiltration is known to  
266 be strongly associated with poor survival rate in solid tumor patients due to its promotion of  
267 tumor angiogenesis and induction of tumor migration, invasion and metastasis.<sup>38,39</sup> We thus  
268 performed Monocle3 analysis to infer the pseudo-time of the transition from DCIS\_LCIS\_3 to  
269 IDC\_6. As DCIS\_LCIS\_3 and IDC\_6 coincided with SEDR clusters 3, 7, and 11 (Figure 4A,  
270 4D), we applied Monocle3 on these three clusters and set cluster 3 (tumor core) as the starting  
271 point (Figure 4D bottom). Monocle3 analysis showed that pseudo-time derived from SEDR  
272 embeddings better traced the inside-out tumor progression compared to that from Seurat PCA  
273 embeddings. We subsequently identified genes that changed expression along the Monocle3  
274 pseudo-time and revealed sequential waves of gene regulation along the trajectory (Figure  
275 4E).

276 In summary, SEDR analysis revealed intratumoral heterogeneity within visually  
277 homogenous tumor regions and revealed the tumor outer ring (cluster 7) with TAM infiltration  
278 and cancer associated fibroblasts (CAFs), both of which have been reported to facilitate tumor  
279 spread.<sup>44,45</sup> SEDR also enabled the mapping of a molecular trajectory from the tumor core to  
280 its neighboring invasive region, demonstrating the transition from a pro-inflammatory to an  
281 immune-suppressive microenvironment, which may contribute to tumor metastasis.

## 282 **SEDR can handle spatial transcriptomics of high resolution.**

283 Currently available spatial omics technologies, including 10x Visium Spatial Omics,  
284 Nanostring GeoMX DSP, SLIDE-seq<sup>4</sup>, and DBIT-seq<sup>46</sup>, do not provide single-cell resolution,  
285 with each capture spot containing 1 to 10 cells. However, newly emerging methods such as  
286 Stereo-seq<sup>5</sup>, PIXEL-Seq<sup>6</sup>, and Seq-Scope<sup>7</sup> can achieve submicrometer and thus subcellular  
287 resolution. With continued technology advancement, the spatial resolution and number of cells  
288 detected per tissue will significantly improve, producing large datasets with high throughput.  
289 As such, we evaluated SEDR's performance on one type of such data produced by Stereo-  
290 seq from mouse olfactory bulb tissues (Figure 5). The coronal section of a mouse olfactory  
291 bulb contains the olfactory nerve layer (ONL), glomerular layer (GL), external plexiform layer  
292 (EPL), mitral cell layer (MCL), internal plexiform layer (IPL), granule cell layer (GCL), and  
293 rostral migratory stream (RMS) (Figure 5A). We performed unsupervised clustering using the  
294 Seurat-derived principal components and SEDR-derived embeddings to computationally  
295 reconstruct the spatial identity of the olfactory bulb tissues. Compared to Seurat clusters, those  
296 produced by SEDR better reflected tissue organization and were more consistent with known  
297 anatomical layers (Figure 5B, 5C). We also performed quantitative assessment using local  
298 inverse Simpson's index (LISI) and found that SEDR produced significantly lower LISI than  
299 Seurat, implying better spatial separation by SEDR (Figure 5D).

## 300 **Discussion**

301  
302 Cell type heterogeneity is a feature of both healthy and diseased tissue. Capturing this  
303 heterogeneity, coupled with its spatial arrangement in the tissue, is crucial when studying the  
304 roles of cells and their cross-talk. Spatial omics technologies represent the state-of-the-art  
305 approaches for capturing omics data with corresponding spatial information from tissue  
306 samples. In this paper, we have introduced SEDR, which leverages on cutting edge graph  
307 neural network techniques to achieve a better representation of spatial omics data that can be  
308 used for clustering and further downstream analyses. SEDR first learns a low-dimensional

309 latent space representation of the transcriptome information with a deep autoencoder network,  
310 which is then aggregated with spatial neighborhood information by a variational graph  
311 autoencoder to create a spatial embedding. This spatial embedding is then concatenated with  
312 the gene expression to be decoded to reconstruct the final gene expression for further  
313 analyses. We first demonstrated the efficacy of SEDR in delineating the different cerebral  
314 cortex layers with higher clarity than competing methods, and recapitulated the associated  
315 development order by using the joint latent representation with Monocle3.

316         To enhance the analytical power and resolution of spatial omics, we need to integrate  
317 multiple datasets from the same tissue. Similar to single-cell transcriptomic data, spatial omics  
318 datasets generated in different batches also contain batch-specific systematic variations that  
319 present a challenge to batch-effect removal and data integration. In our study, we  
320 demonstrated that by combining SEDR and Harmony, we were able to effectively remove  
321 batch effects present. In the future, we will integrate Harmony into the SEDR workflow.

322         Spatial omics technologies such as Stereo-seq are able to measure a large number of  
323 cells in a single experiment through high spatial resolutions and large tissue sizes. In the near  
324 future, we expect to see ever-increasing throughput from spatial omics experiments, which will  
325 result in spatial omics big data that will pose significant challenges to data analysis and  
326 integration. Computational methods that employ GCNs require the entire graph to be loaded  
327 into GPU memory, which inhibits their application to very large datasets. We will improve the  
328 memory efficiency of SEDR using a GCN mini-batch or parallel techniques to construct large-  
329 scale graphs for spatial omics data of high throughput and resolution. Furthermore,  
330 technologies with a capture spot size smaller than the diameter of a cell will also require new  
331 computational methods that can accurately delineate cells based on capture spots. In the  
332 future, we will integrate cell segmentation based on H&E or DAPI staining into the SEDR  
333 workflow.

334         The current SEDR methodology focuses on gene expression and spatial information,  
335 and does not make use of histological images. Contemporary methods such as SpaGCN and

336 stLearn use histological images as input, but in a suboptimal fashion, as demonstrated in our  
337 study. Specifically, SpaGCN utilizes histological image pixels as features by calculating the  
338 mean color values from the RGB channels directly. However, the pixel values are easily  
339 affected by noise and cannot provide semantic features for cell analysis. A more effective  
340 approach can be to adopt a deep CNN model which can learn high-level representations for  
341 histological images. stLearn introduces a deep learning model to extract image features of the  
342 spots and integrates them with the spatial location and gene expression. However, stLearn  
343 employs a model pre-trained on natural images, and does not fine-tune the network for  
344 histological images. In the future, we will incorporate histological images as an additional  
345 modality into the SEDR model. We will add an image autoencoder network to learn image  
346 features, and jointly learn the latent representation by integrating gene expression, image  
347 morphology, and spatial information.

348 In summary, SEDR is a promising new approach that builds an integrated  
349 representation of cells using both transcriptomic data and spatial coordinates. SEDR-derived  
350 low-dimensional embedding enables more accurate clustering, trajectory inference and batch  
351 effect correction. Our model is also able to handle spatial transcriptomics with capture spot  
352 sizes ranging from 50 $\mu$ m to less than 1 $\mu$ m. Furthermore, we applied SEDR on human breast  
353 cancer to reveal heterogeneous sub-regions within the seemingly homogenous tumor region and  
354 shed light on the role of immune microenvironments on tumor invasiveness.

## 355 Methods

### 356 **Dataset preprocessing.**

357 Our SEDR method takes spatial transcriptomic gene expressions and spatial coordinates as  
358 inputs. The raw gene expression counts are first normalized using the respective library sizes  
359 (by `normalize_total` in Scanpy (v.1.5.0)), with very highly expressed genes excluded when  
360 computing the normalization factor (size factor) for each cell<sup>47</sup>. PCA is then applied to extract  
361 the first 200 principal components to generate the initial gene expression matrix.

362

### 363 **Graph construction for spatial transcriptomics data.**

364 To create a graph representing the cell–cell spatial relationships in spatial transcriptomics data,  
365 we calculate the Euclidean distances between cells using the image coordinates, and select  
366 the top 10 nearest neighbors of each cell to construct an adjacency matrix. The adjacency  
367 matrix, denoted by  $A$ , is a symmetric matrix, where  $A_{ij} = A_{ji} = 1$  if  $i$  and  $j$  are neighbors, and  
368 0 otherwise.

### 369 **Deep autoencoder for latent representation learning.**

370 The latent representation of gene expression is learned using a deep autoencoder. The  
371 encoder part, consisting of two fully connected stacked layers, generates a low-dimensional  
372 representation  $Z_f \in \mathbb{R}^{N \times D_f}$  from the input gene expression matrix  $X \in \mathbb{R}^{N \times M}$ . Meanwhile, the  
373 decoder part with one fully connected layer reconstructs the expression matrix  $X' \in \mathbb{R}^{N \times M}$   
374 from the latent representation  $Z \in \mathbb{R}^{N \times D}$ , which is obtained by concatenating the low-  
375 dimensional representation  $Z_f$  and spatial embedding  $Z_g \in \mathbb{R}^{M \times D_g}$ , where  $N$  is the number of  
376 cells,  $M$  is the number of input genes, and  $D_f, D_g, D$  are the dimensions of the low-dimensional  
377 expression representation learned by the encoder, the spatial embedding learned by the GCN,  
378 and the final latent representation of SEDR, respectively with  $D = D_f + D_g$ . The objective  
379 function of the deep autoencoder maximizes the similarity between the input gene and

380 reconstructed expressions measured by the mean squared error (MSE) loss function  
 381  $\sum(X - X')^2$ .

382 **Variational graph autoencoder for spatial embedding.**

383 SEDR utilizes a variational graph autoencoder<sup>48</sup> (VGAE) to embed the spatial information of  
 384 neighboring cells. With the adjacency matrix  $A$  and its degree matrix  $D$ , the VGAE learns a  
 385 graph embedding  $Z_g$  with the following format:  $g: (A, Z_f) \rightarrow Z_g$ , where  $Z_f$  is the node/gene  
 386 representation from the deep autoencoder. The inference part of the VGAE is parameterized  
 387 by a two-layer GCN<sup>49</sup> :

388 
$$g(Z_g|A, Z_f) = \prod g(z_i|A, Z_f), \text{ with } g(z_i|A, Z_f) = \mathcal{N}(z_i|\mu_i, \text{diag}(\sigma_i^2)),$$

389 where  $\mu = GCN_\mu(A, Z_f)$  is the matrix of mean vectors, and  $\log\sigma = GCN_\sigma(A, Z_f)$ . The two-layer  
 390 GCN is defined as:

391 
$$GCN(A, Z_f) = \tilde{A} \text{ReLU}(\tilde{A}Z_fW_0)W_1,$$

392 with a weight  $W_i$  and symmetrically normalized adjacency matrix  $\tilde{A} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ . The spatial  
 393 embedding  $Z_g$  and reconstructed adjacency matrix  $A'$  are generated as:

394 
$$A' = \sigma(Z_g \cdot Z_g^T),$$

395 with  $Z_g = GCN(A, Z_g)$ . The objective of the VGAE is to minimize the cross-entropy (CE) loss  
 396 between the input adjacency matrix  $A$  and reconstructed adjacency matrix  $A'$ , while  
 397 simultaneously minimizing the Kullback-Leibler (KL) divergence between  $g(Z_g|A, Z_f)$  and the  
 398 Gaussian prior:

399 
$$p(Z_g) = \prod_i \mathcal{N}(z_i|0, I).$$

400 **Batch effect correction for spatial transcriptomics.**

401 Spatial relationships only exist within single spatial transcriptomic measurement; cells/spots  
 402 from different transcriptomic measurements have no direct spatial relation. Let  $A^k$  and  $Z_f^k$   
 403 denote the adjacency matrix and deep gene representation of spatial omics  $k$ , we then create

404 a block-diagonal adjacency matrix  $A^k$  and concatenate the deep gene representation in the  
 405 cell dimension, as:

$$406 \quad A = \begin{bmatrix} A^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A^K \end{bmatrix}, \quad Z_f = \begin{bmatrix} Z_f^1 \\ \vdots \\ Z_f^K \end{bmatrix},$$

407 where  $K$  is the number of spatial omics. Based on this formulation, we transform different  
 408 spatial omics datasets (of potentially different sizes) into multiple graph instances in the form  
 409 of one block-diagonal adjacency matrices as inputs to SEDR.

410 To remove batch effects and enhance the compactness of its latent representation, SEDR  
 411 employs an unsupervised deep embedded clustering (DEC) method<sup>26</sup> to iteratively group the  
 412 cells into different clusters. To initialize the cluster centers, we employ the KMeans of scikit-  
 413 learn on the learned latent representations. The number of clusters is pre-defined as a  
 414 hyperparameter. With the initialization, DEC improves the clustering using an unsupervised  
 415 iterative method of two steps. In the first step, a soft assignment  $q_{ij}$  of latent point  $z_i$  to cluster  
 416 center  $\mu_j$  is calculated using the Student's t-distribution:

$$417 \quad q_{ij} = \frac{(1 + \|z_i - \mu_j\|^2)^{-1}}{\sum_{j'} (1 + \|z_i - \mu_{j'}\|^2)^{-1}}.$$

418 In the second step, we iteratively refine the clusters by learning from their high confidence  
 419 assignments with the help of an auxiliary target distribution  $p$  based on  $q_{ij}$ :

$$420 \quad p_{ij} = \frac{q_{ij}^2 / \sum_i q_{ij}}{\sum_{j'} (q_{ij'}^2 / \sum_i q_{ij'})}.$$

421 Based on the soft assignment  $q_{ij}$  and auxiliary target distribution  $p_{ij}$ , an objective function is  
 422 defined using the KL divergence:

$$423 \quad KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}.$$

424 The SEDR parameters and cluster centers are then simultaneously optimized by using  
 425 stochastic gradient descent (SGD) with momentum.

426

427 **Seurat.**

428 Raw mRNA counts were preprocessed to remove low-quality genes and sctransformed to  
429 remove technical artifacts and normalize the data.<sup>50</sup> We then ran PCA to extract the top 30  
430 principal components (PCs) and used them to calculate the shared nearest neighbors (SNNs).  
431 Then, the Louvain clustering algorithm was used to identify clusters with the SNN networks.  
432 We tried clustering at different resolutions to obtain the same number of clusters as the ground  
433 truth layers.

434 **SpaGCN, stLearn, BayesSpace, Giotto.**

435 We ran these methods with the recommended parameters and set each one to generate the  
436 same number of clusters as the ground truth layers. The stLearn-derived low-dimensional  
437 embedding was used for downstream UMAP visualization and Harmony batch correction.

438 **Evaluation metrics for clustering.**

439 For datasets with cell-type labels (e.g., DLPFC), we employed ARI to compare the  
440 performance of different clustering algorithms. ARI calculates the similarity between the  
441 clustering labels predicted by the algorithm and reference cluster labels as:

442 
$$ARI = \frac{RI - E[RI]}{\max(RI) - E[RI]},$$

443 where the unadjusted rand index (RI) is defined as  $RI = (a + b)/C_n^2$ , with  $a$  being the number  
444 of pairs correctly labeled as coming from the same set,  $b$  being the number of pairs correctly  
445 labeled as not in the same set, and  $C_n^2$  being the total number of possible pairs.  $E[RI]$  is the  
446 expected  $RI$  of random labeling. A higher ARI score indicates better performance.

447 **Monocle3.**

448 On the DLPFC #151673 slice and breast cancer data, we ran Monocle3 using both the Seurat  
449 and SEDR outputs. For Seurat, we ran the standard pipeline to obtain the UMAP. For SEDR,  
450 we first extracted the low-dimensional embedding and then used the uwot package to calculate

451 the UMAP. We then ran Monocle3 on both UMAPs using the recommended parameters and  
452 set white matter (WM) as the starting point to generate the pseudo-time. Finally, we used the  
453 Moran\_I test to detect significant genes that showed correlations with the pseudo-time.

#### 454 **Leiden clustering, PAGA trajectory, and UMAPs.**

455 We used the Scanpy (v.1.5.0) package to compute the Leiden clustering, partition-based graph  
456 abstraction (PAGA), and uniform manifold approximation and projection (UMAP) from SEDR-  
457 derived joint embeddings of gene expression and spatial information. Briefly, we used SEDR  
458 embeddings to compute neighborhood graphs with 15 as the number of neighbors and  
459 Euclidean distance as the distance measure. To obtain the same number of unique Leiden  
460 clusters, grid-searching on the Leiden clustering resolutions between 0.2 and 2.5 was  
461 performed at intervals of 0.05/0.01. Subsequently, PAGA was applied to quantify the  
462 connectivity between Leiden clusters. Finally, the cluster positions suggested by PAGA were  
463 used to compute the UMAP for visualization.

#### 464 **Harmony.**

465 Harmony was used to correct batch effects on low-dimensional embeddings. For SEDR, we  
466 used latent space embeddings as input. For the raw data and stLearn, we used the PCA  
467 embeddings. We treated different samples as different batches, and set all other parameters  
468 to their default values. For each method, the uncorrected embeddings and batch-corrected  
469 Harmony embeddings were used for UMAP visualization and analysis.

#### 470 **Prediction of cell type composition of 10x Visium spatial spot.**

471 We downloaded a published scRNA-seq dataset of human breast cancer<sup>36</sup> as reference, and  
472 ran Seurat to find transfer anchors between the reference and our Visium spatial data. Cell  
473 types in the reference were then assigned to the spatial spots by label transferring. We  
474 removed cell types with probabilities equal to 0 for all spots.

#### 475 **Differential expression analysis and pathway analyses.**

476 We used Seurat to identify DEGs. Genes with adjusted p-values < 0.05 were used as the

477 input for QIAGEN Ingenuity Pathway Analysis (IPA). For IPA results, pathways with  
478 positive or negative z-scores were plotted.

#### 479 **Raw data processing of Stereo-seq data.**

480 Fastq files were generated using the MGI DNBSEQ-Tx sequencer. Coordinate identities  
481 (CIDs) and unique molecular identifiers (UMIs) were encoded in the forward reads (CID: 1-  
482 25bp, UMI: 26-35bp), while the reverse reads consisted of the cDNA sequences. CID  
483 sequences in the forward reads were first mapped to the designed coordinates of the *in situ*  
484 captured chip, allowing one base mismatch to correct for sequencing and PCR errors. Reads  
485 with UMIs containing either N bases or more than two bases with quality scores lower than 10  
486 were filtered out. The CIDs and UMIs associated with each read were appended to each read  
487 header. Retained reads were then aligned to the reference genome (mm10) using STAR<sup>51</sup>,  
488 and mapped reads with MAPQ  $\geq 10$  were counted and annotated using an *in-house* script  
489 (available at <https://github.com/BGIResearch/handleBam>). UMIs with the same CIDs and gene  
490 loci were collapsed together, allowing for one mismatch to correct for sequencing and PCR  
491 errors, to give the final gene expression matrix.

#### 492 **Local inverse Simpson's index (LISI).**

493 We first used Seurat and SEDR to generate cell clusters for the stereo-seq data, and then the  
494 R "lisi" package to calculate the LISIs using spatial coordinates as X and the clustering results  
495 of Seurat and SEDR as meta data.

496

#### 497 **Data availability.**

498 (1) LIBD human dorsolateral prefrontal cortex (DLPFC) Data  
499 (<http://spatial.libd.org/spatialLIBD/>); (2) 10x visium spatial transcriptomics data of human  
500 breast cancer and Stereo-seq of mouse olfactory bulb are at  
501 <https://github.com/JinmiaoChenLab/SEDR/> (3) Analysis results and scripts to reproduce the  
502 results are at <https://github.com/JinmiaoChenLab/SEDR/>

503

504 **Software availability.**

505 SEDR was written in Python using the PyTorch library. An open-source implementation of

506 SEDR has been released on <https://github.com/HzFu/SEDR>

507

508

509

510

511 **References**

- 512 1. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by  
513 spatial transcriptomics. *Science* (2016) doi:10.1126/science.aaf2403.
- 514 2. Maynard, K. R. *et al.* Transcriptome-scale spatial gene expression in the human  
515 dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).
- 516 3. Ji, A. L. *et al.* Multimodal Analysis of Composition and Spatial Architecture in Human  
517 Squamous Cell Carcinoma. *Cell* (2020) doi:10.1016/j.cell.2020.08.043.
- 518 4. Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide  
519 expression at high spatial resolution. *Science* (80-. ). (2019)  
520 doi:10.1126/science.aaw1219.
- 521 5. Chen, A. *et al.* Large field of view-spatially resolved transcriptomics at nanoscale  
522 resolution Short title: DNA nanoball stereo-sequencing. *bioRxiv* 2021.01.17.427004  
523 (2021).
- 524 6. Fu, X. *et al.* Continuous Polony Gels for Tissue Mapping with High Resolution and  
525 RNA Capture Efficiency. *bioRxiv* 2021.03.17.435795 (2021).
- 526 7. Cho, C.-S. *et al.* Seq-Scope: Submicrometer-resolution spatial transcriptomics for  
527 single cell and subcellular studies. *bioRxiv* (2021).
- 528 8. Stickels, R. R. *et al.* Highly sensitive spatial transcriptomics at near-cellular resolution  
529 with Slide-seqV2. *Nat. Biotechnol.* (2021) doi:10.1038/s41587-020-0739-1.
- 530 9. Liu, Y. *et al.* High-Spatial-Resolution Multi-Omics Sequencing via Deterministic  
531 Barcoding in Tissue. *Cell* (2020) doi:10.1016/j.cell.2020.10.026.
- 532 10. Lein, E., Borm, L. E. & Linnarsson, S. The promise of spatial transcriptomics for  
533 neuroscience in the era of molecular cell typing. *Science* (2017)  
534 doi:10.1126/science.aan6827.
- 535 11. Yoosuf, N., Navarro, J. F., Salmén, F., Ståhl, P. L. & Daub, C. O. Identification and  
536 transfer of spatial transcriptomics signatures for cancer diagnosis. *Breast Cancer Res.*  
537 (2020) doi:10.1186/s13058-019-1242-9.

- 538 12. van den Brink, S. C. *et al.* Single-cell and spatial transcriptomics reveal somitogenesis  
539 in gastruloids. *Nature* (2020) doi:10.1038/s41586-020-2024-3.
- 540 13. Dong, R. & Yuan, G. C. SpatialDWLS: accurate deconvolution of spatial  
541 transcriptomic data. *Genome Biol.* **22**, 1–10 (2021).
- 542 14. Song, Q. & Su, J. DSTG: deconvoluting spatial transcriptomics data through graph-  
543 based artificial intelligence. *Brief. Bioinform.* (2021) doi:10.1093/bib/bbaa414.
- 544 15. Andersson, A. *et al.* Single-cell and spatial transcriptomics enables probabilistic  
545 inference of cell type topography. *Commun. Biol.* **3**, 1–8 (2020).
- 546 16. Biancalani, T. *et al.* Deep learning and alignment of spatially-resolved whole  
547 transcriptomes of single cells in the mouse brain with Tangram. *bioRxiv* (2020).
- 548 17. Gayoso, A. *et al.* scvi-tools: a library for deep probabilistic analysis of single-cell omics  
549 data. *bioRxiv* (2021).
- 550 18. Lopez, R. *et al.* Multi-resolution deconvolution of spatial transcriptomics data reveals  
551 continuous patterns of inflammation. *bioRxiv* (2021).
- 552 19. Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I. & Heyn, H. SPOTlight: seeded NMF  
553 regression to deconvolute spatial transcriptomics spots with single-cell  
554 transcriptomes. *Nucleic Acids Res.* (2021) doi:10.1093/nar/gkab043.
- 555 20. Danaher, P. *et al.* Advances in mixed cell deconvolution enable quantification of cell  
556 types in spatially-resolved gene expression data. *bioRxiv* (2020).
- 557 21. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Bioarxiv* (2020).
- 558 22. Zhao, E. *et al.* Bayesspace enables the robust characterization of spatial gene  
559 expression architecture in tissue sections at increased resolution. *bioRxiv* (2020)  
560 doi:10.1101/2020.09.04.283812.
- 561 23. Dries, R. *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial  
562 expression data. *Genome Biol.* **22**, 78 (2021).
- 563 24. Hu, J. *et al.* Integrating gene expression, spatial location and histology to identify  
564 spatial 1 domains and spatially variable genes by graph convolutional network 2 3.  
565 *bioRxiv* 2020.11.30.405118 (2020).

- 566 25. Pham, D. *et al.* stLearn: Integrating spatial location, tissue morphology and gene  
567 expression to find cell types, cell-cell interactions and spatial trajectories within  
568 undissociated tissues. *bioRxiv* (2020) doi:10.1101/2020.05.31.125658.
- 569 26. Xie, J., Girshick, R. & Farhadi, A. Unsupervised Deep Embedding for Clustering  
570 Analysis. in *ICML* (eds. Balcan, M. F. & Weinberger, K. Q.) vol. 48 478–487 (PMLR,  
571 2016).
- 572 27. Mann, H. B. & Whitney, D. R. On a Test of Whether one of Two Random Variables is  
573 Stochastically Larger than the Other. *Ann. Math. Stat.* **18**, 50–60 (1947).
- 574 28. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by  
575 pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
- 576 29. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis.  
577 *Nature* **566**, 496–502 (2019).
- 578 30. Wolf, F. A. *et al.* PAGA: graph abstraction reconciles clustering with trajectory  
579 inference through a topology preserving map of single cells. *Genome Biol.* **20**, 1–9  
580 (2019).
- 581 31. Gilmore, E. G. & Herrup, K. Cortical development: Layers of complexity. *Current*  
582 *Biology* (1997) doi:10.1016/s0960-9822(06)00108-4.
- 583 32. Chini, M. & Hanganu-Opatz, I. L. Prefrontal Cortex Development in Health and  
584 Disease: Lessons from Rodents and Humans. *Trends in Neurosciences* (2021)  
585 doi:10.1016/j.tins.2020.10.017.
- 586 33. Nadarajah, B. & Parnavelas, J. G. Modes of neuronal migration in the developing  
587 cerebral cortex. *Nature Reviews Neuroscience* (2002) doi:10.1038/nrn845.
- 588 34. Tran, H. T. N. *et al.* A benchmark of batch-effect correction methods for single-cell  
589 RNA sequencing data. *Genome Biol.* (2020) doi:10.1186/s13059-019-1850-9.
- 590 35. Nguyen, P. H. D. *et al.* Intratumoural immune heterogeneity as a hallmark of tumour  
591 evolution and progression in hepatocellular carcinoma. *Nat. Commun.* **12**, 1–13  
592 (2021).
- 593 36. *et al.* A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic

- 594 states in the human breast. *EMBO J.* **40**, 1–23 (2021).
- 595 37. Friedl, P. & Alexander, S. Cancer invasion and the microenvironment: Plasticity and  
596 reciprocity. *Cell* (2011) doi:10.1016/j.cell.2011.11.016.
- 597 38. Kuroda, H. *et al.* Tumor microenvironment in triple-negative breast cancer: the  
598 correlation of tumor-associated macrophages and tumor-infiltrating lymphocytes. *Clin.*  
599 *Transl. Oncol.* (2021) doi:10.1007/s12094-021-02652-3.
- 600 39. Asiry, S. *et al.* The Cancer Cell Dissemination Machinery as an Immunosuppressive  
601 Niche: A New Obstacle Towards the Era of Cancer Immunotherapy. *Front. Immunol.*  
602 **12**, 1–19 (2021).
- 603 40. Ju, J. A., Godet, I., DiGiacomo, J. W. & Gilkes, D. M. RhoB is regulated by hypoxia  
604 and modulates metastasis in breast cancer. *Cancer Rep.* (2020)  
605 doi:10.1002/cnr2.1164.
- 606 41. Olson, O. C. & Joyce, J. A. Cysteine cathepsin proteases: Regulators of cancer  
607 progression and therapeutic response. *Nature Reviews Cancer* (2015)  
608 doi:10.1038/nrc4027.
- 609 42. Roumenina, L. T. *et al.* Tumor cells hijack macrophage-produced complement C1q to  
610 promote tumor growth. *Cancer Immunol. Res.* (2019) doi:10.1158/2326-6066.CIR-18-  
611 0891.
- 612 43. Fraser, D., Melzer, E., Camacho, A. & Gomez, M. Macrophage production of innate  
613 immune protein C1q is associated with M2 polarization (INM1P.434). *J. Immunol.*  
614 (2015).
- 615 44. Monteran, L. & Erez, N. The dark side of fibroblasts: Cancer-associated fibroblasts as  
616 mediators of immunosuppression in the tumor microenvironment. *Frontiers in*  
617 *Immunology* (2019) doi:10.3389/fimmu.2019.01835.
- 618 45. Lin, Y., Xu, J. & Lan, H. Tumor-associated macrophages in tumor metastasis:  
619 Biological roles and clinical therapeutic applications. *Journal of Hematology and*  
620 *Oncology* (2019) doi:10.1186/s13045-019-0760-3.
- 621 46. Deng, Y. *et al.* Resource High-Spatial-Resolution Multi-Omics Sequencing via

622 Deterministic Barcoding in Tissue II II Resource High-Spatial-Resolution Multi-Omics  
623 Sequencing via Deterministic Barcoding in Tissue. *Cell* **183**, 1665-1681.e18 (2020).

624 47. Weinreb, C., Wolock, S. & Klein, A. M. SPRING: a kinetic interface for visualizing high  
625 dimensional single-cell expression data. *Bioinformatics* **34**, 1246–1248 (2018).

626 48. Kipf, T. N. & Welling, M. Variational Graph Auto-Encoders. in *NIPS Workshop on*  
627 *Bayesian Deep Learning* (2016).

628 49. Kipf, T. N. & Welling, M. Semi-Supervised Classification with Graph Convolutional  
629 Networks. in *International Conference on Learning Representations (ICLR)* (2017).

630 50. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell  
631 RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296  
632 (2019).

633 51. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* (2013)  
634 doi:10.1093/bioinformatics/bts635.

635

636

637 **Acknowledgements:**

638 This research was supported by funding from Singapore Immunology Network (SIgN), A \*  
639 STAR, Singapore.

640

641 **Author contributions:**

642 Huazhu Fu designed and implemented SEDR. Hang Xu, Huazhu Fu, Kelvin Chong, Mengwei  
643 Li, Hong Kai Lee and Jingjing Ling performed data analysis. Hang Xu, Huazhu Fu, Mengwei  
644 Li generated figures. Jinmiao Chen, Huazhu Fu, Hang Xu, Kok Siong Ang, Kelvin Chong,  
645 Jingjing Ling and Ling Shao drafted the manuscript. Ao Chen and Longqi Liu provided Stereo-  
646 seq data. Jinmiao Chen conceptualized and supervised the study.

647

648 **Competing interests:**

649 The authors declare no competing interests.

650

651 **Figure legends:**

652 **Figure 1. Overview of SEDR.** SEDR learns a low-dimensional latent representation of gene  
653 expression embedded with spatial information by jointly training a deep autoencoder and a  
654 variational graph autoencoder. The low-dimensional embedding produced by SEDR can be  
655 used for downstream visualization, cell clustering, trajectory inference, and batch effect  
656 correction.

657 **Figure 2. Quantitative assessment of SEDR on the human dorsolateral prefrontal cortex**

658 **(DLPFC) dataset.** A) Ground-truth segmentation of cortical layers; clustering results of Seurat,  
659 Giotto, stLearn, SpaGCN, BayesSpace, and SEDR on DLPFC slice #151673; and adjusted  
660 rand index (ARI) of various cluster sets on the 12 DLPFC slices. B) UMAP visualization and  
661 Monocle3 trajectory generated using the Seurat-derived PCA embedding (left) and SEDR  
662 embedding (right). Monocle pseudotimes visualized on UMAP plots (middle) and spatial co-  
663 ordinates (bottom). C) PAGA graphs generated using the Seurat-derived PCA embedding (top)

664 and SEDR embedding (middle). The ratios of the sum of weights of correctly inferred PAGA  
665 edges to the sum of weights of all edges produced by SEDR and Seurat (bottom).

666 **Figure 3. Batch effects present in DLPFC dataset and assessment of SEDR's**  
667 **performance on batch correction.** A) Slices #151507, #151672 and #151673 showed  
668 substantial inter-slice variations before batch effect correction. UMAP plots colored by ground-  
669 truth cortical layers (left), slices (right), split by layers and colored by slices (bottom). B)  
670 Harmony alone was unable to remove the batch effects present. C) SEDR alone substantially  
671 reduced the batch effects. D) SEDR combined with Harmony effectively corrected the batch  
672 effects. E) stLearn combined with Harmony was unable to correct the batch effects.

673 **Figure 4. Application of SEDR on 10x Visium spatial transcriptomics data of human**  
674 **breast cancer.** A) Manual pathology labeling based on H&E staining (annotation); clustering  
675 results of SEDR, Seurat, stLearn, SpaGCN, and BayesSpace. B) Seurat3 'anchor'-based  
676 integration workflow was used to perform probabilistic transfer of annotations from a reference  
677 scRNA-seq data of human breast cancer to the spatial data. This gives a probabilistic  
678 classification of the scRNA-seq derived classes for each spot. The probabilities of tumor  
679 associated macrophage (TAMs) and cycling epithelials (C-EPI) were visualized. C) Pathways  
680 enriched by genes differentially expressed between SEDR clusters 3 (core) and 7 (outer ring).  
681 Red bars represent pathways upregulated in cluster 3. D) Trajectory analysis results using  
682 PAGA (top) and Monocle3 (bottom). The PAGA graph predictions of the inter-relatedness  
683 between the manually annotated DCIS/LCIS and IDC regions. The edge width denotes  
684 connectivity strength, thus indicating the likelihood of an actual connection being present.  
685 Monocle3 inferred the pseudo-times of spots in SEDR clusters 3, 7, and 11 using the Seurat-  
686 derived PCA embedding (termed "rna\_pseudotime") and SEDR embedding (termed  
687 "SEDR\_pseudotime"). E) Heatmap of genes with expression changes along the Monocle-  
688 derived pseudo-time.

689 **Figure 5. Application of SEDR on Stereo-seq spatial transcriptomics data of mouse**  
690 **olfactory bulb tissue sections.** A) Laminar organization of DAPI-stained mouse olfactory

691 bulb. B) Unsupervised clustering of the spatial voxels analyzed by Seurat and SEDR. C) The  
692 four clusters with the highest numbers of voxels were selected and visualized. D) Quantitative  
693 comparison of Seurat and SEDR clusters using local inverse Simpson's index (LISI).

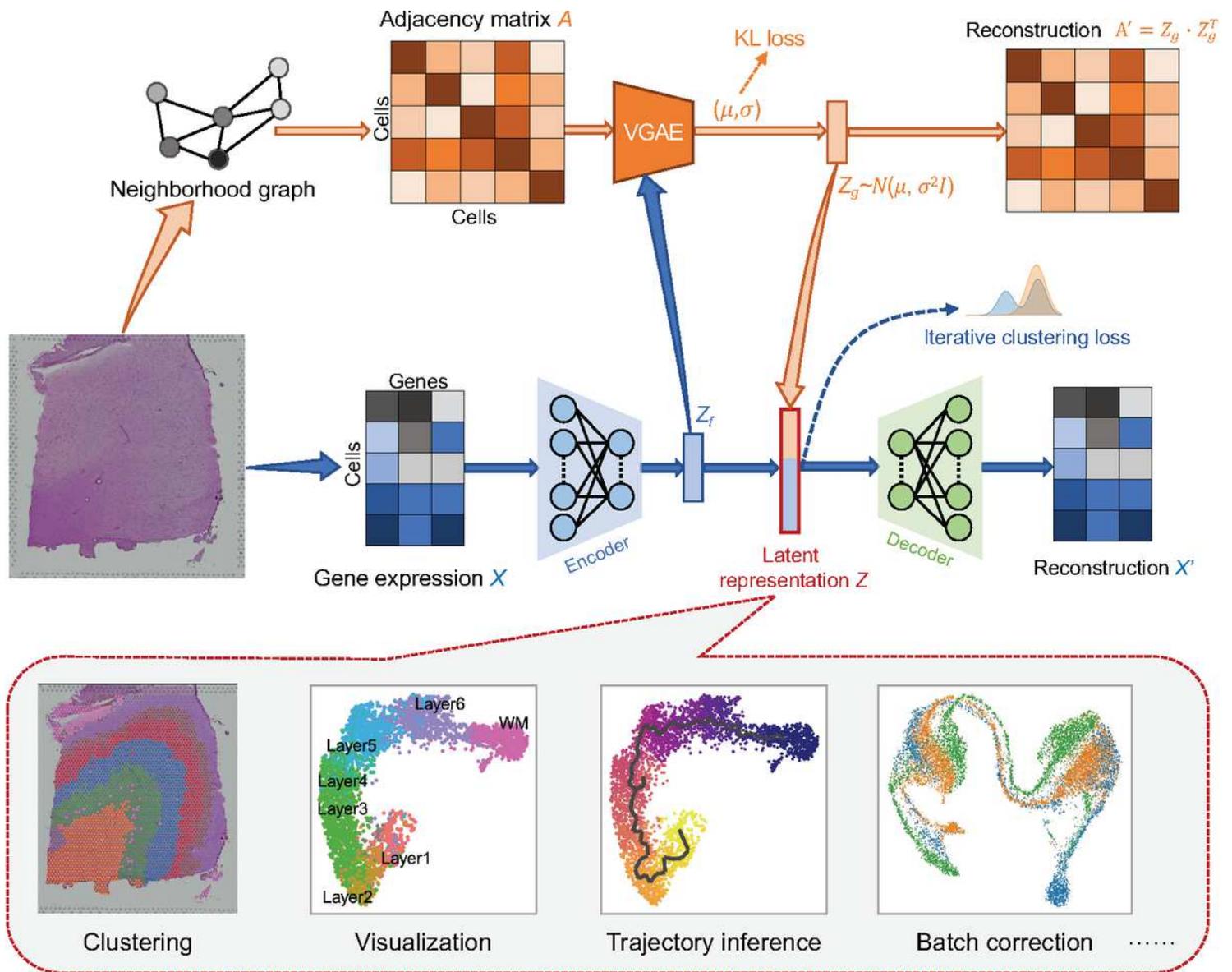
694 **Supplementary:**

695 Figure 1. **Human breast cancer histology and cell type mixtures of spatial spots.** A) H&E  
696 staining. B) Seurat3 predicted probabilities of scRNA-seq derived cell types.

697 Figure 2. **Differentially expressed genes (DEGs) and enriched pathways.** A) Locations of  
698 DCIS\_LCIS\_3 and other DCIS\_LCIS regions. B) Top DEGs between DCIS\_LCIS\_3 and other  
699 DCIS\_LCIS regions. C) Enriched pathways of DEGs between DCIS\_LCIS\_3 and other  
700 DCIS\_LCIS regions. Red bars represent pathways up-regulated in DCIS\_LCIS\_3 D)  
701 Percentages of tumor associated macrophages (TAMs) in SEDR cluster 3 (tumor core) and  
702 cluster 7 (tumor edge). E) Violin plots of selected DEGs between SEDR clusters 3 and 7.

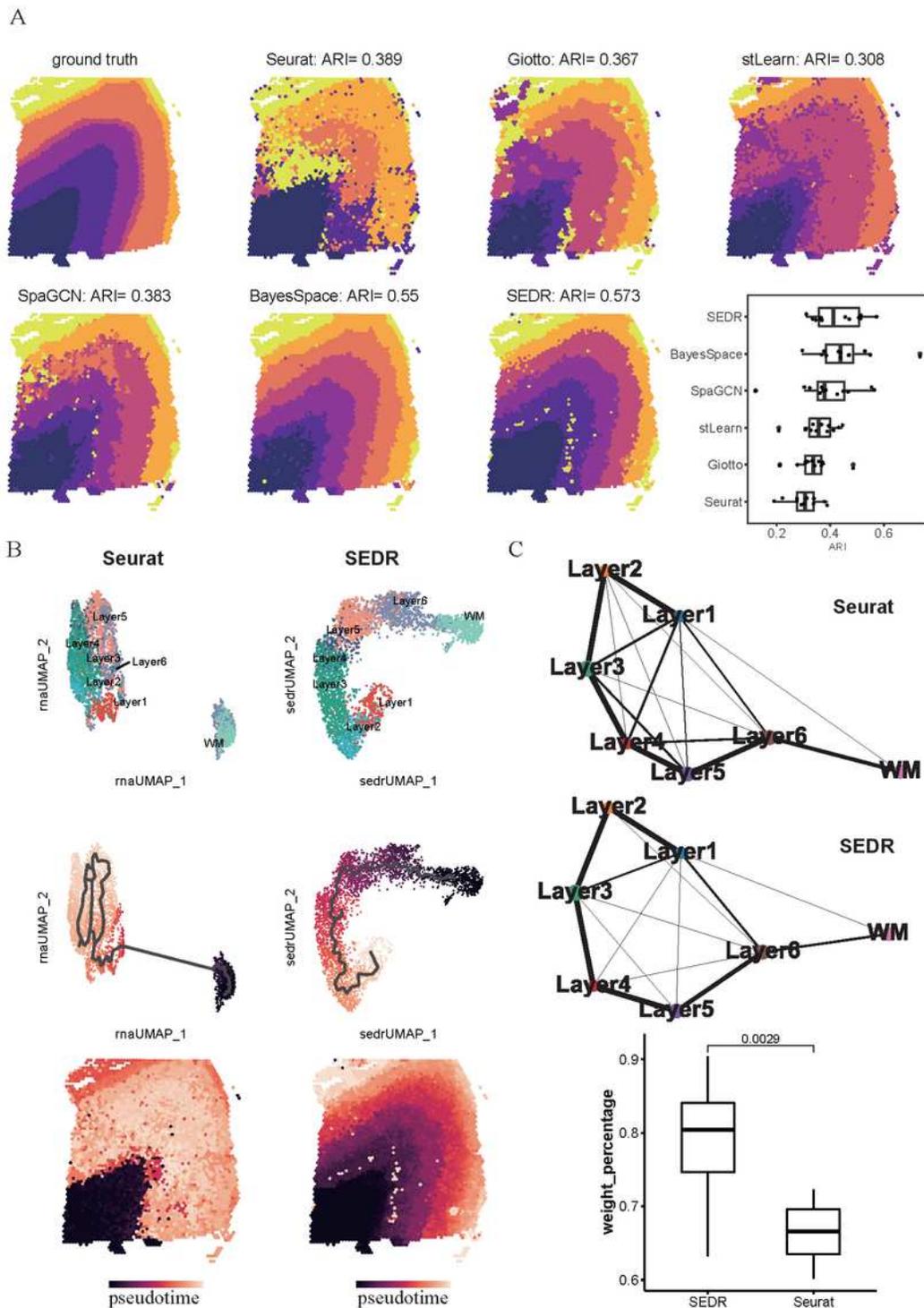
703

# Figures



**Figure 1**

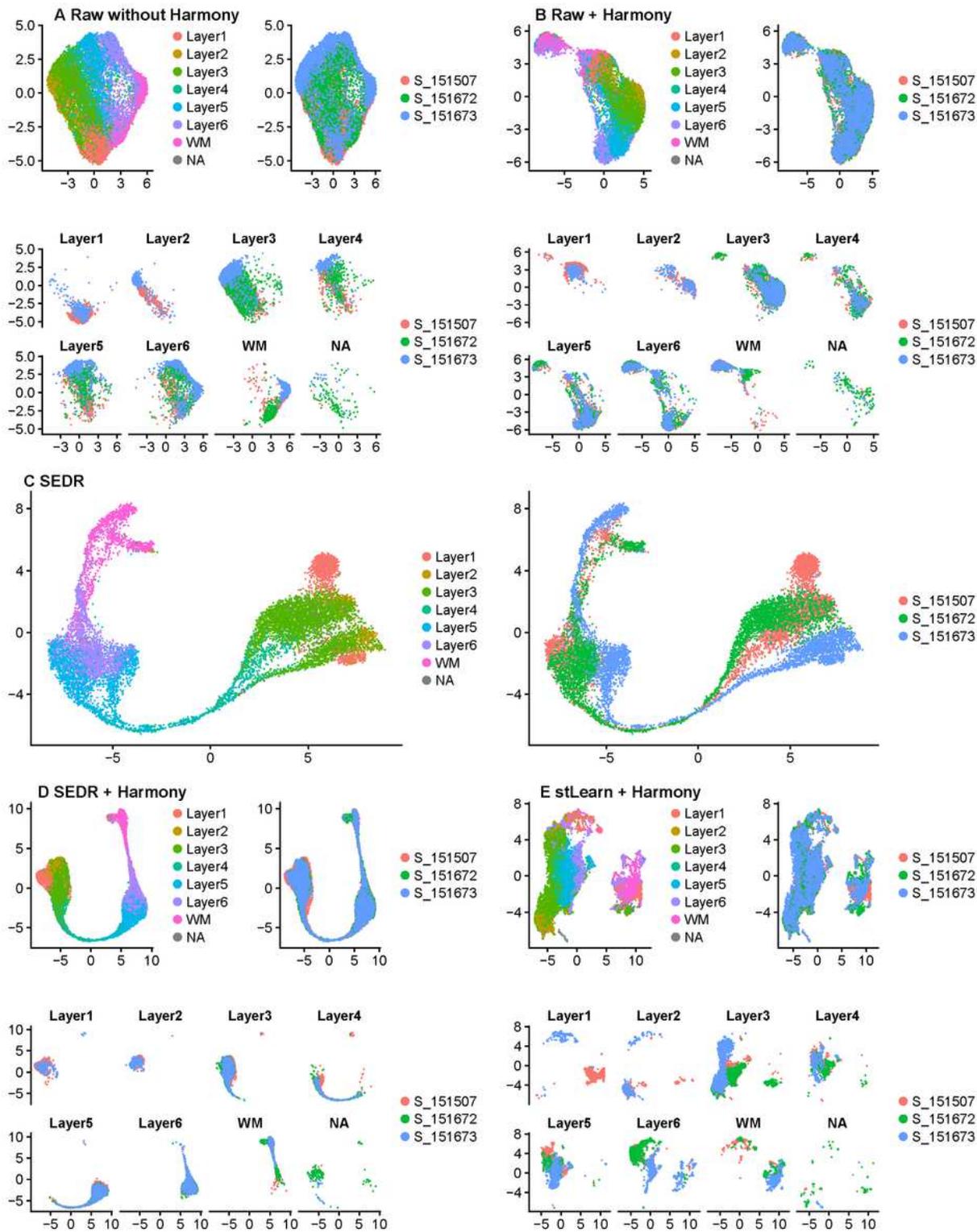
Overview of SEDR. SEDR learns a low-dimensional latent representation of gene expression embedded with spatial information by jointly training a deep autoencoder and a variational graph autoencoder. The low-dimensional embedding produced by SEDR can be used for downstream visualization, cell clustering, trajectory inference, and batch effect correction.



**Figure 2**

Quantitative assessment of SEDR on the human dorsolateral prefrontal cortex (DLPFC) dataset. A) Ground-truth segmentation of cortical layers; clustering results of Seurat, Giotto, stLearn, SpaGCN, BayesSpace, and SEDR on DLPFC slice #151673; and adjusted rand index (ARI) of various cluster sets on the 12 DLPFC slices. B) UMAP visualization and Monocle3 trajectory generated using the Seurat-derived PCA embedding (left) and SEDR embedding (right). Monocle pseudotimes visualized on UMAP plots

(middle) and spatial co-ordinates (bottom). C) PAGA graphs generated using the Seurat-derived PCA embedding (top) and SEDR embedding (middle). The ratios of the sum of weights of correctly inferred PAGA edges to the sum of weights of all edges produced by SEDR and Seurat (bottom).



**Figure 3**

Batch effects present in DLPFC dataset and assessment of SEDR's performance on batch correction. A) Slices #151507, #151672 and #151673 showed substantial inter-slice variations before batch effect

correction. UMAP plots colored by ground-truth cortical layers (left), slices (right), split by layers and colored by slices (bottom). B) Harmony alone was unable to remove the batch effects present. C) SEDR alone substantially reduced the batch effects. D) SEDR combined with Harmony effectively corrected the batch effects. E) stLearn combined with Harmony was unable to correct the batch effects.

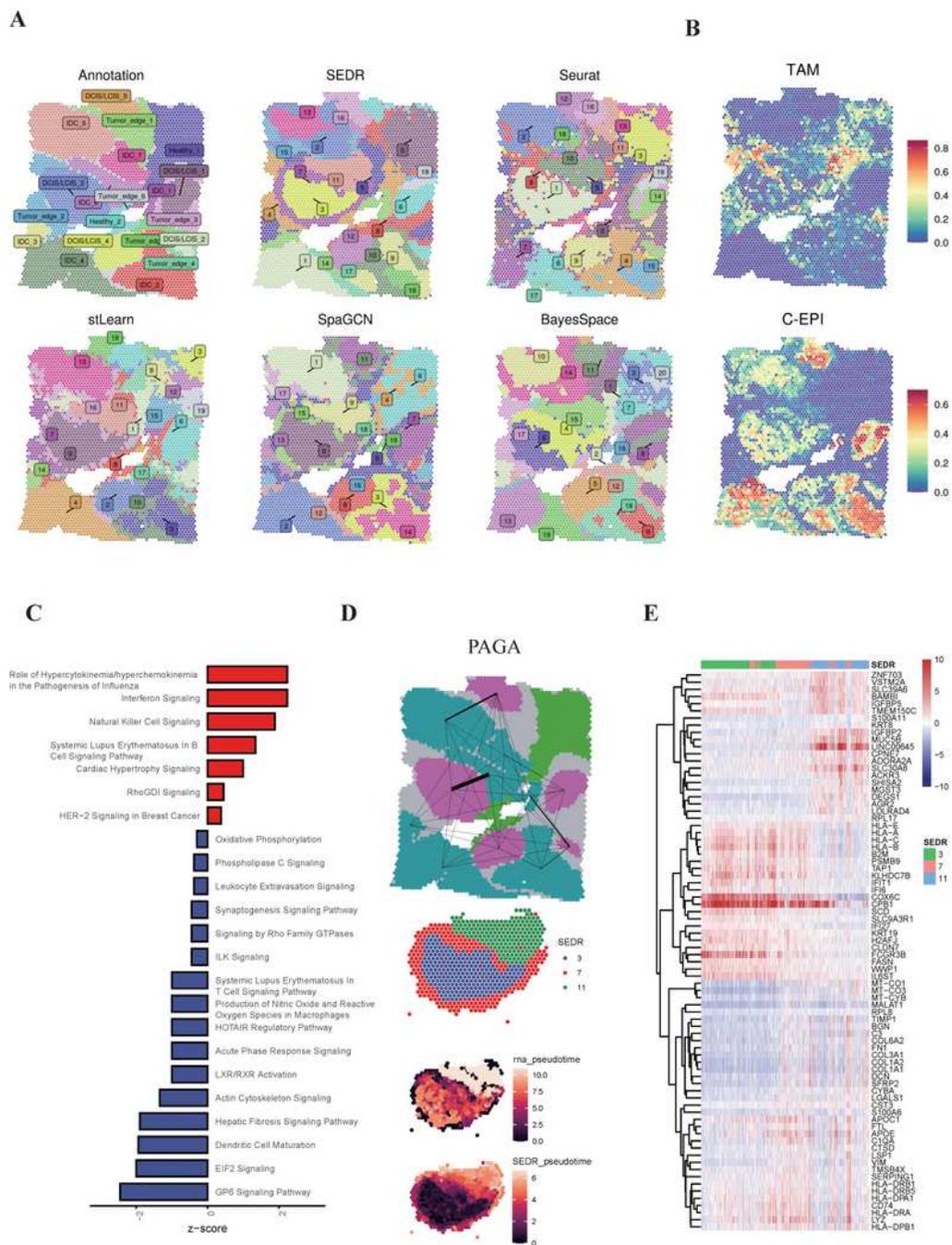
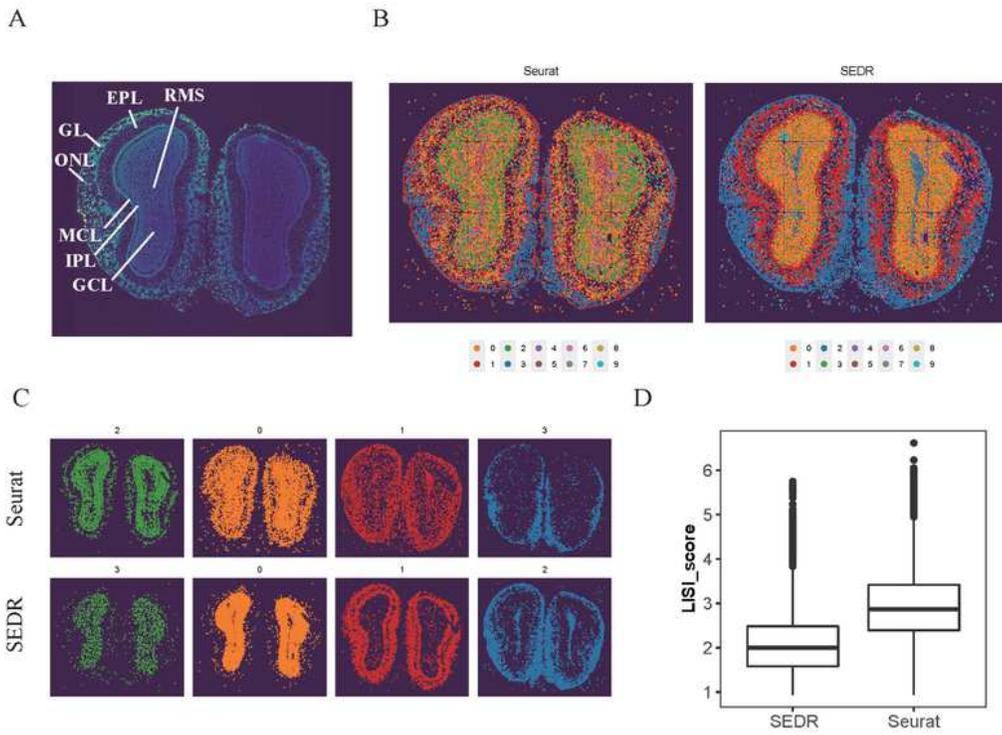


Figure 4

Application of SEDR on 10x Visium spatial transcriptomics data of human breast cancer. A) Manual pathology labeling based on H&E staining (annotation); clustering results of SEDR, Seurat, stLearn, SpaGCN, and BayesSpace. B) Seurat3 'anchor'-based integration workflow was used to perform probabilistic transfer of annotations from a reference scRNA-seq data of human breast cancer to the spatial data. This gives a probabilistic classification of the scRNA-seq derived classes for each spot. The probabilities of tumor associated macrophage (TAMs) and cycling epithelials (C-EPI) were visualized. C) Pathways enriched by genes differentially expressed between SEDR clusters 3 (core) and 7 (outer ring). Red bars represent pathways upregulated in cluster 3. D) Trajectory analysis results using PAGA (top) and Monocle3 (bottom). The PAGA graph predictions of the inter-relatedness between the manually annotated DCIS/LCIS and IDC regions. The edge width denotes connectivity strength, thus indicating the likelihood of an actual connection being present. Monocle3 inferred the pseudo-times of spots in SEDR clusters 3, 7, and 11 using the Seurat-derived PCA embedding (termed "rna\_pseudotime") and SEDR embedding (termed "SEDR\_pseudotime"). E) Heatmap of genes with expression changes along the Monocle-derived pseudo-time.



**Figure 5**

Application of SEDR on Stereo-seq spatial transcriptomics data of mouse olfactory bulb tissue sections. A) Laminar organization of DAPI-stained mouse olfactory bulb. B) Unsupervised clustering of the spatial voxels analyzed by Seurat and SEDR. C) The four clusters with the highest numbers of voxels were selected and visualized. D) Quantitative comparison of Seurat and SEDR clusters using local inverse Simpson's index (LISI).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supp1.pdf](#)
- [supp2.pdf](#)