

1 **Feasibility of Inferring Spatial Transcriptomics from Single-Cell Histological Patterns for Studying**  
2 **Colon Cancer Tumor Heterogeneity**

3 Michael Y. Fatemi<sup>1,\*</sup>, Yunrui Lu<sup>2,\*</sup>, Cyril Sharma<sup>3</sup>, Eric Feng<sup>4</sup>, Zarif L. Azher<sup>4</sup>, Alos B. Diallo<sup>2,5,6</sup>, Gokul  
4 Srinivasan<sup>2</sup>, Grace M. Rosner<sup>7,8</sup>, Kelli B. Pointer<sup>7,8</sup>, Brock C. Christensen<sup>5</sup>, Lucas A. Salas<sup>5</sup>, Gregory J.  
5 Tsongalis<sup>2</sup>, Scott M. Palisoul<sup>2</sup>, Laurent Perreard<sup>9</sup>, Fred W. Kolling IV<sup>9</sup>, Louis J. Vaickus<sup>2</sup>, Joshua J.  
6 Levy<sup>2,6,10,11,12,\*\*</sup>

- 7 1. Department of Computer Science, University of Virginia, Charlottesville, VA
- 8 2. Emerging Diagnostic and Investigative Technologies, Department of Pathology and Laboratory  
9 Medicine, Dartmouth Health, Lebanon, NH
- 10 3. Department of Computer Science, Purdue University, West Lafayette, IN
- 11 4. Thomas Jefferson High School for Science and Technology, Alexandria, VA
- 12 5. Department of Epidemiology, Dartmouth College Geisel School of Medicine, Hanover, NH
- 13 6. Program in Quantitative Biomedical Sciences, Dartmouth College Geisel School of Medicine,  
14 Hanover, NH
- 15 7. Department of Medicine, Section of Radiation Oncology, Dartmouth Health, Lebanon, NH
- 16 8. Department of Molecular and Cell Biology, Dartmouth College Geisel School of Medicine,  
17 Hanover, NH
- 18 9. Genomics Shared Resource, Dartmouth Cancer Center, Lebanon, NH
- 19 10. Department of Dermatology, Dartmouth Health, Lebanon, NH
- 20 11. Department of Pathology and Laboratory Medicine, Cedars Sinai Medical Center, Los Angeles,  
21 CA
- 22 12. Department of Computational Biomedicine, Cedars Sinai Medical Center, Los Angeles, CA

23 \* Denotes equal contribution as first authors

24 \*\* To whom correspondence should be addressed

25  
26 **Corresponding Author Information:**

27 Joshua Levy  
28 Director of Digital Pathology Research  
29 Department of Pathology and Computational Medicine  
30 Cedars Sinai Medical Center  
31 (925) 457-5752  
32 [joshua.levy@cshs.org](mailto:joshua.levy@cshs.org)

33

34

## 35 Abstract

36 **Background.** Spatial transcriptomics involves studying the spatial organization of gene expression within  
37 tissues, offering insights into the molecular diversity of tumors. While spatial gene expression is  
38 commonly amalgamated from 1-10 cells across 50-micron spots, recent methods have demonstrated the  
39 capability to disaggregate this information at subspot resolution by leveraging both expression and  
40 histological patterns. However, elucidating such information from histology alone presents a significant  
41 challenge but if solved can better permit spatial molecular analysis at cellular resolution for instances  
42 where Visium data is not available, reducing study costs. This study explores integrating single-cell  
43 histological and transcriptomic data to infer spatial mRNA expression patterns in whole slide images  
44 collected from a cohort of stage pT3 colorectal cancer patients. A cell graph neural network algorithm was  
45 developed to align histological information extracted from detected cells with single cell RNA patterns  
46 through optimal transport methods, facilitating the analysis of cellular groupings and gene relationships.  
47 This approach leveraged spot-level expression as an intermediary to co-map histological and  
48 transcriptomic information at the single-cell level.

49 **Results.** Our study demonstrated that single-cell transcriptional heterogeneity within a spot could be  
50 predicted from histological markers extracted from cells detected within a spot. Furthermore, our model  
51 exhibited proficiency in delineating overarching gene expression patterns across whole-slide images. This  
52 approach compared favorably to traditional patch-based computer vision methods as well as other  
53 methods which did not incorporate single cell expression during the model fitting procedures. Topological  
54 nuances of single-cell expression within a Visium spot were preserved using the developed methodology.

55 **Conclusion.** This innovative approach augments the resolution of spatial molecular assays utilizing  
56 histology as a sole input through synergistic co-mapping of histological and transcriptomic datasets at the  
57 single-cell level, anchored by spatial transcriptomics. While initial results are promising, they warrant  
58 rigorous validation. This includes collaborating with pathologists for precise spatial identification of distinct  
59 cell types and utilizing sophisticated assays, such as Xenium, to attain deeper subcellular insights.

60 **Keywords.** Spatial transcriptomics, Tumor-immune microenvironment, Whole slide images, Cell graph  
61 neural network, Gene expression patterns, Pathway analysis, Molecular profiling

## 62 Background

63 Cancer warrants considerable attention given its profound impact on individuals and their families, the  
64 absence of a clear cure in many cases, and the necessity to advance technologies for prevention, early  
65 detection, and treatment selection. By the end of 2023, nearly 2 million new cancer cases and more than  
66 600,000 cancer deaths will occur in the United States [1,2]. Colorectal cancer (CRC) persists as one of  
67 the most formidable solid tumors, with an annual incidence of approximately 150,000 new cases in the  
68 United States and a 63% 5-year survival rate [1,2]. With the shift in CRC to younger demographics and  
69 tumor metastasis being responsible for most cancer deaths, there is a pressing need for a high-fidelity  
70 screening and prognostication program [3]. The treasure trove of imaging and genomics information  
71 provided by nascent molecular assays and informatics techniques has the potential to inform more  
72 effective, targeted treatment options by revealing novel prognostic biomarkers.

73  
74 The pathological TNM-staging system (pTNM) is the most predictive factor for the risk of recurrence and  
75 prognosis [4–7]. Cancer staging is broadly characterized by assessing local invasiveness (T-stage),  
76 followed by the extent of lymph node involvement (N-stage), then the presence of metastasis to distant  
77 sites. Successful extraction of lymph nodes at the time of resection is crucial for establishing the N-stage,  
78 which primarily informs the risk of recurrence [6]. A higher yield of lymph nodes during surgery is often  
79 associated with favorable outcomes at the expense of potential morbidity; as such, the recommended  
80 lymph node yield is set at 12 [6,8]. Yet, one recent population study found that only 37% of such  
81 assessments extract this number, underscoring significant implications for accurate risk assessment and  
82 selection of optimal therapies [4]. For example, the usage of neoadjuvant chemotherapy is typically  
83 considered for stage III patients with proficient mismatch repair (MMR) status. As there are instances  
84 where the presence of metastasis is uncertain, thereby impacting the fidelity of prognostic assessments,  
85 informatics tools that can study the spatial biology of the tumor at the primary site can infer missing  
86 staging information (e.g., lymph node stage) or identify risk factors independent of pTNM staging relevant  
87 which provide additional predictive value for the recurrence risk assessment.

88

89 Tumor Infiltrating Lymphocytes (TIL) play a crucial role in understanding and modulating the Tumor  
90 Microenvironment (TME) and Tumor Immune Microenvironment (TIME) [9]. The TME consists of  
91 malignant and benign cells, blood vessels, and extracellular matrix, interconnected through complex  
92 communication via cytokine recruitment factors [9]. Recent studies highlight the importance of immune  
93 infiltrates, such as T cells, B cells, NK cells, and monocyte/lymphocyte cells, and their distribution,  
94 density, and relationships in mounting an effective anti-tumor response. Microsatellite Instability (MSI)  
95 status also influences this response [10]. For example, high levels of cytotoxic T cells within the tumor  
96 may indicate immune exhaustion, while increased immune cell density can suggest a favorable prognosis  
97 [11]. Understanding the molecular changes and spatial arrangements associated with colon cancer  
98 metastasis is still incomplete. Nonetheless, several digital pathology assays have incorporated existing  
99 findings to complement pTNM staging and serve as independent risk factors for recurrence. These  
100 assays include: 1) Immunoscore, which measures the density of cytotoxic T-cells at the tumor's invasive  
101 margin and inside the tumor [12], 2) CDX2, an epithelial marker of pluripotency indicating the tumor's  
102 ability to bypass immune response and growth inhibition checkpoints [13–15], and 3) circulating tumor  
103 DNA, such as mutations in the Vascular Endothelial Growth Factor (VEGF) pathway [16–18]. While these  
104 assays are predictive of recurrence risk, they provide only a limited perspective on tumor metastasis  
105 phenomenology.

106  
107 Spatial omics technologies, like 10x Genomics Spatial Transcriptomics (ST) or GeoMX Digital Spatial  
108 Profiling (DSP), have facilitated the simultaneous analysis of multiple biomarkers, including the whole  
109 transcriptome, with remarkable spatial resolution [19–22]. These technologies have been applied to  
110 further characterize TIL subpopulations in TME. However, their clinical utility is limited due to high costs,  
111 low throughput, and limited reproducibility. In previous work, we demonstrated the feasibility of utilizing  
112 machine learning algorithms to extract TIL and spatial biology information from Hematoxylin and Eosin  
113 (H&E) stains. This can be a cost-effective and high-throughput digital biomarker that could be employed  
114 prospectively as an adjunct test similar to Immunoscore for recurrence risk assessment [23–25]. We  
115 found that careful selection of algorithms is crucial to capture molecular alterations and pathways  
116 reflective of histomorphological changes or large-scale tissue architecture changes [26,27].

117  
118 Nevertheless, the resolution of these findings is currently restricted to the available resolution of Visium  
119 spots, typically around 50 microns, which aggregates expression data across a small number of cells (1-  
120 10 cells). Incorporating single-cell information, captured through the new Chromium Flex technology, can  
121 enable further characterization of spatial cellular heterogeneity to enhance the resolution of the Visium  
122 data. Recent advancements in 10x profiling technologies, including Chromium Flex and CytAssist assays,  
123 enable the profiling of single-cell transcriptomics (scRNASeq) on serial sections of formalin-fixed paraffin-  
124 embedded (FFPE) tissue. By utilizing automated staining devices and advanced imaging techniques prior  
125 to conducting Visium spatial transcriptomics (ST) assays, the CytAssist technology significantly improves  
126 workflow that allow for higher resolution whole slide images (WSI) to be collected. Pairing this high-  
127 resolution slide imaging information with ST and serial section scRNASeq has the potential to enhance  
128 the capacity to perform spatial assessments at single-cell resolution on external study cohorts.

129  
130 While several technologies have been developed to increase the resolution of Visium data, such  
131 algorithms require both ST and histological information and do not operate on tissue images alone.  
132 Previous studies have made attempts to infer single-cell RNA sequencing (scRNA-seq) data from breast  
133 cancer tissue slide sections, improving the resolution of the data and enabling the identification of  
134 different cell types within the tissue [28]. Other studies have made attempts to infer Visium ST expression  
135 patterns aggregated across several cells per spot using image classification techniques with some  
136 domain-specific adaptations. For example, recent studies have trained DenseNet-121 and InceptionV3  
137 models to predict gene expression [29,30], and another work used a custom convolutional layer along  
138 with a graph attention network and transformer model to share information between Visium spots [25].  
139 While the Visium platform primarily provides low-resolution, aggregated expression measurements across  
140 cells contained within a 50-micron spot [31,32], single-cell analyses offer a more comprehensive view of  
141 cellular heterogeneity. One study attempts to reconstruct false zero-counts in cell-level omics data by  
142 training a masked autoencoder over known cell data [33], and other recent efforts have been made to  
143 directly infer single-cell transcriptomes from whole slide images by collecting information from single cells

144 in serial sections following H&E staining, yet lacked ST data to guide the co-registration of the single cell  
145 data to the slide [34].

146

147 The primary goal of this study is to enhance the predictive capability of algorithms that infer spatial  
148 transcriptomics (ST) data solely from histology images, capturing single-cell heterogeneity within a spot  
149 and their aggregate spot-level expression. To achieve this, we combine the precise locations of individual  
150 cells, as identified in whole cell images, with the granular data from single-cell RNA sequencing (scRNA-  
151 seq). This approach integrates histological details from localized nuclei within and around Visium spots  
152 with corresponding scRNA-seq profiles mapped to the same spots. By seamlessly merging these  
153 datasets, our framework stands poised to extract richer molecular insights from cells, facilitating a more  
154 accurate prediction of both Visium ST and individual cell information.

155

156 Through feature engineering methods, we develop attribution methods to examine the structural  
157 organizations of cells that are most correlated with the expression of specific genes. This can contribute  
158 to a better understanding of the dynamics of the tumor-immune microenvironment and potentially aid in  
159 developing prognostic tools for the stage and aggressiveness of colorectal tumors. In this paper, we  
160 compare the accuracy of methods that use cells as features with conventional computer vision methods  
161 featured in our previous work. It is important to emphasize that this study does not claim to infer  
162 scRNASeq data at specific locations of individual cells. Rather, we demonstrate the ability to leverage  
163 single-cell information to enhance the expression prediction at Visium spots on held-out tissue slides.  
164 This research establishes a foundational workflow and conceptual framework for the future inference of  
165 such information.

166

## 167 Results

### 168 Overview of Cells2RNA Framework: Bridging Histological Patterns with 169 Single-Cell Expression

170 Cells2RNA was crafted to infer single-cell expression insights from discernible histological patterns in  
171 instances where spatial transcriptomics and single-cell data might be lacking (**Figure 1**). The crux of the  
172 challenge lies in deducing single-cell nuances solely from histological patterns surrounding pinpointed  
173 cells (**Figure 2A**). Historically, prior research has been limited to interpreting aggregated spot-level data.  
174 Yet, when disaggregated to the individual cell level, a richer tapestry of heterogeneity emerges, which  
175 becomes our focal point for inference. This innovation holds the potential to augment findings across  
176 expansive cohorts without the constraints of traditional spatial analyses that heavily rely on Visium data.  
177 The goal of this study is to derive molecular insights paralleling the depth of Visium-based investigations,  
178 but strictly from histological imaging.

179  
180 Central to our approach is a co-mapping methodology. Here, histological patterns detected at the cellular  
181 level are intricately aligned with single-cell expression data (**Figure 1A**). Spatial transcriptomics serves as  
182 a vital intermediary in this process: during training, it not only maps single-cell RNASeq data to  
183 corresponding Visium spots (**Figure 1B**) where cells are located but also acts as a crucial inference target  
184 for the expression-centric histological attributes derived from these located cells.

185  
186 To accomplish this, Cell Graph Neural Networks (CGNNs) are employed (**Figure 2B,E,F**). They efficiently  
187 extract the essence of histological features from located cells within Visium spots (**Figure 1C**). These  
188 features are updated during model training through simultaneous harmonization with scRNASeq data via  
189 optimal transport methodologies (**Figures 1D,2G**), forging a cohesive link between cellular histological  
190 characteristics and their intricate gene expressions.

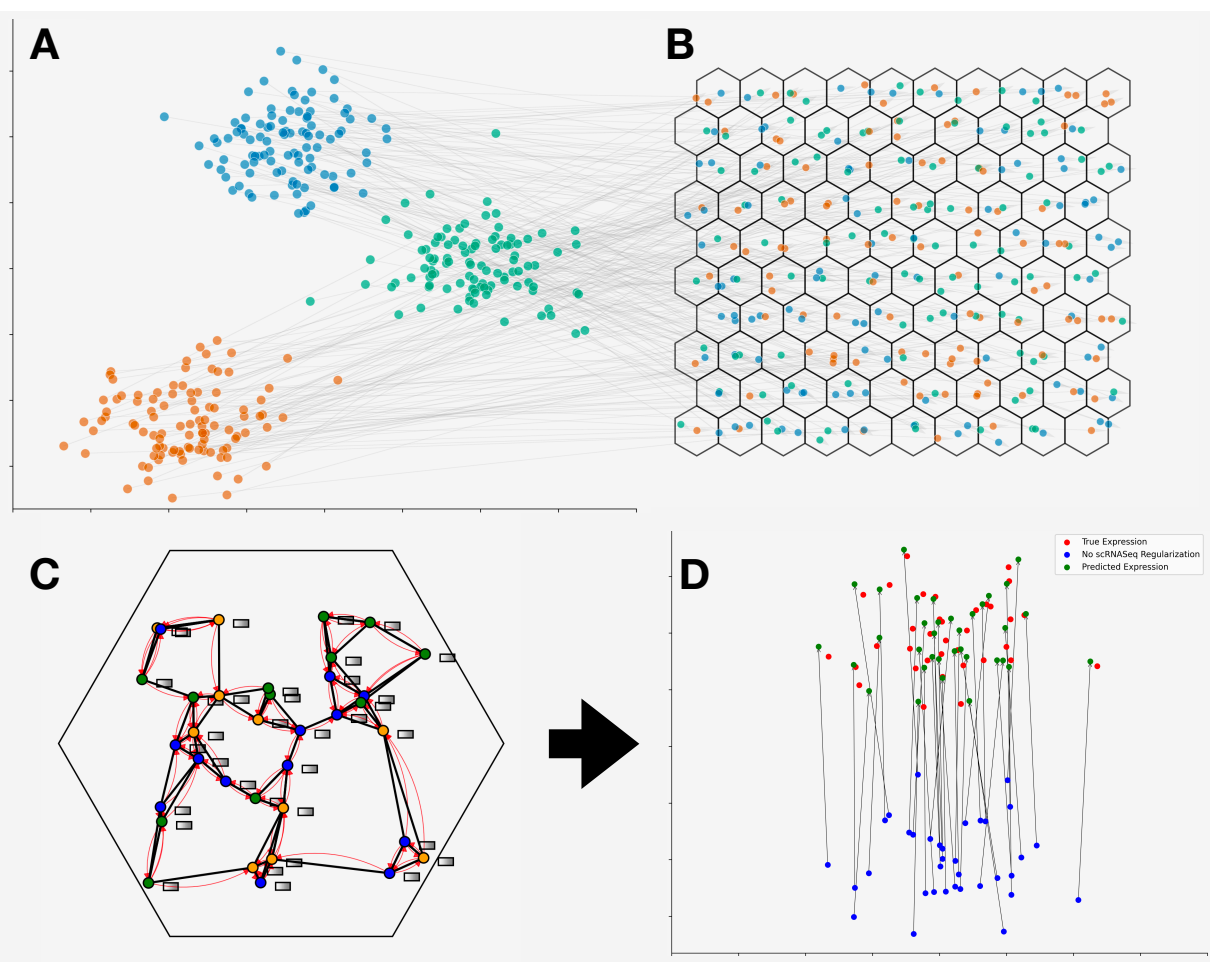
191

192 The culmination of this methodological synergy is a framework where the expression profiles predicted  
193 from detected cells align with the broader single-cell expression landscape. Although this alignment might  
194 not be perfect, it closely mirrors genuine single-cell expression dynamics within each Visium spot. Thus,  
195 our anticipation is twofold: the accurate convergence of aggregated inferred single-cell expression to  
196 spot-level benchmarks and a holistic reflection of single-cell expression diversity within the spot when  
197 applied to external, held-out slides (**Figure 2C,D**).

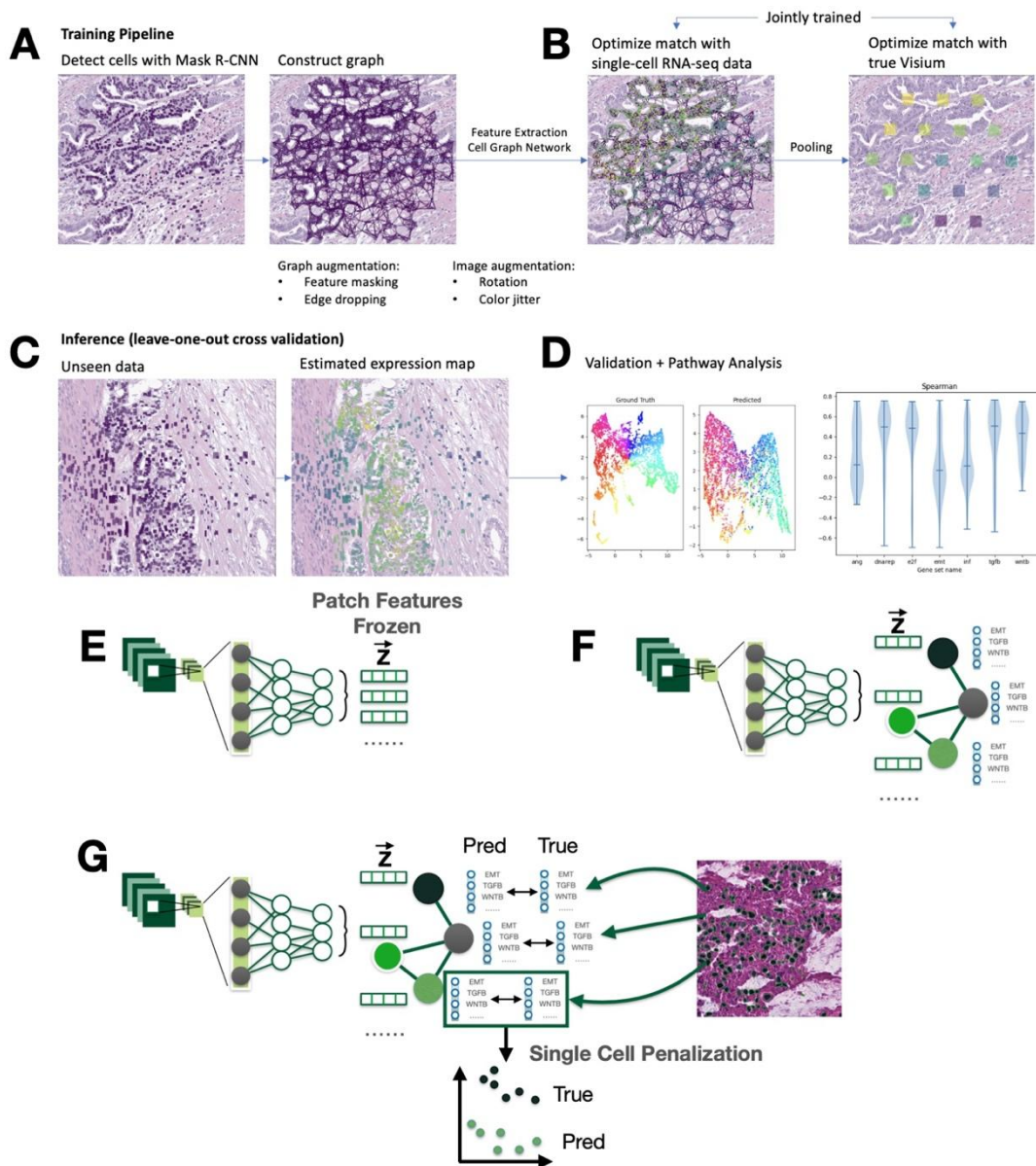
198  
199 Using Visium and paired 40X resolution whole slide imaging from a cohort of nine stage pT3 colorectal  
200 patients (see section “Data Collection and Preprocessing”), the co-mapping technique was benchmarked  
201 against patch-level models (Inceptionv3) and other CGNNs that utilize alternative information extraction  
202 methods. We assessed their performance based on predicting spot-level expression, capturing cellular  
203 heterogeneity within spots (using Wasserstein distance), maintaining tissue architectural relationships,  
204 and conducting a pathway analysis. A detailed examination of these methodologies and comparisons is  
205 provided in the Methods section.

206





207  
208 **Figure 1: Overview of Cell2RNA's Co-Mapping Approach:** (A) Low-dimensional visualization of single-  
209 cell RNA profiles, clusters indicating cell-type. (B) Spatial layout of identified cells across the tissue slide  
210 (assignment to spots represented by hexagons), color-coded by distinct gene expression patterns  
211 mapped from single cell profiles featured in (A). (C) In-depth view of cells located within a specific Visium  
212 spot, illustrating connectivity and cell relationships. Expression-related histological features, represented  
213 by grey rectangles, are shared among neighboring cells through red curves via a graph neural network.  
214 (D) A side-by-side low-dimensional comparison of scRNASeq profiles for a representative Visium spot:  
215 actual expression (red), model-predicted expression using the co-mapping training approach (green), and  
216 expression prediction without co-mapping training (blue).



217

218 **Figure 2: Schematic Representation of the Neural Network Workflow for Single-Cell Analysis.**

219 During the training phase, (A) a pre-trained Mask R-CNN model is applied to raw histology images to  
 220 detect individual cells, after which a 6-nearest neighbors graph is constructed for the detected cells. (B)  
 221 Features for each cell are extracted using a ResNet-50 neural network, and the aggregation of  
 222 neighboring cell information is modeled using a Graph Attention Network (GAT). For each Visium spot,  
 223 the node features are aggregated using sum pooling. (C) Pre-pooled node values are jointly optimized  
 224 against single-cell RNA-sequencing (scRNA-seq) data, and (D) pooled Visium spot predictions are  
 225 optimized against the corresponding ground truth data, using a mean-squared-error loss computed  
 226 across log-transformed counts. (E)-(G) Visual description of neural network architectures and  
 227 penalizations employed: (E) a two-stage neural network comprising a feature extraction stage and a  
 228 prediction stage, this was not used in this work, (F) an end-to-end neural network encompassing the  
 229 entire process from cell detection to feature extraction, graph convolutions and prediction, utilized in this  
 230 study, and (G) the incorporation of single-cell-level penalties into the loss function to enforce consistent  
 231 predictions with scRNA-seq data.  
 232

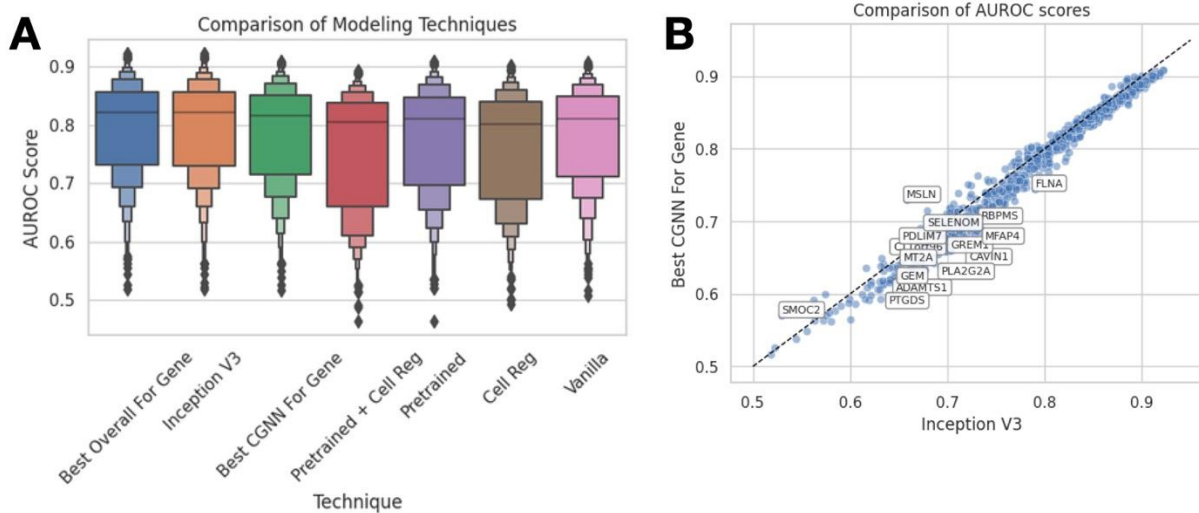
## 233 Model Comparison

234 Overall, models have strong performance— selecting the top CGNN model per gene resulted in an  
235 AUROC of  $0.8138 \pm 0.0069$  and Spearman's statistic of  $0.5724 \pm 0.0133$  (**Table 1**). However, across all  
236 experiments, model performances did not appear significantly different from each other, though we  
237 noticed several important trends (**Figure 3,4**). CGNN models performed on par with the Inception model  
238 (AUC= $0.8204 \pm 0.0073$ ). The most predictive cell-based model had an AUROC of  $0.8093 \pm 0.0083$ ,  
239 similar to the InceptionV3 model's AUROC interval of  $0.8204 \pm 0.0073$ , which leveraged additional  
240 information beyond the cell's immediate neighborhood and may have also benefited from the built-in  
241 structural feature extraction of CNNs. There was high agreement in top-performing genes between CGNN  
242 methods using graph contrastive learning or single-cell penalization as compared to a CGNN with no  
243 penalization/pretraining (**Supplementary Figure 1, Supplementary Table 1**). Notable genes with high  
244 predictive performance include TMSB4X, which encodes an actin-sequestering protein vital for actin  
245 polymerization, cell proliferation, migration, differentiation, and bypasses X inactivation with a homolog on  
246 chromosome Y; and ELF3, which regulates the inflammatory response (**Supplementary Table 1**).  
247

248 **Table 1: Comparison of model performance.** Aggregate AUROC is calculated as the median AUROC  
249 across genes. Gene-level AUROC is calculated as the mean across cross-validation folds.

Modeling Approach	Spearman	AUROC	Optimal Transport (EMD)
Vanilla CGNN	$0.5591 \pm 0.0146$	$0.8093 \pm 0.0083$	$0.2113 \pm 0.0018$
CGNN: Graph Contrastive Learning	$0.5356 \pm 0.0177$	$0.8049 \pm 0.0083$	$0.1900 \pm 0.0020$
CGNN: Single-Cell Penalization	$0.5381 \pm 0.0158$	$0.8012 \pm 0.0074$	$0.1473 \pm 0.0018$
CGNN: GCL and Single-cell penalization	$0.5464 \pm 0.0156$	$0.8084 \pm 0.0093$	$0.1415 \pm 0.0018$
Top CGNN per Gene	$0.5637 \pm 0.0135$	$0.8138 \pm 0.0069$	N/A
Top Model per Gene	$0.5766 \pm 0.0122$	$0.8206 \pm 0.0076$	N/A
InceptionV3 (256x256)	$0.5724 \pm 0.0133$	$0.8204 \pm 0.0073$	N/A

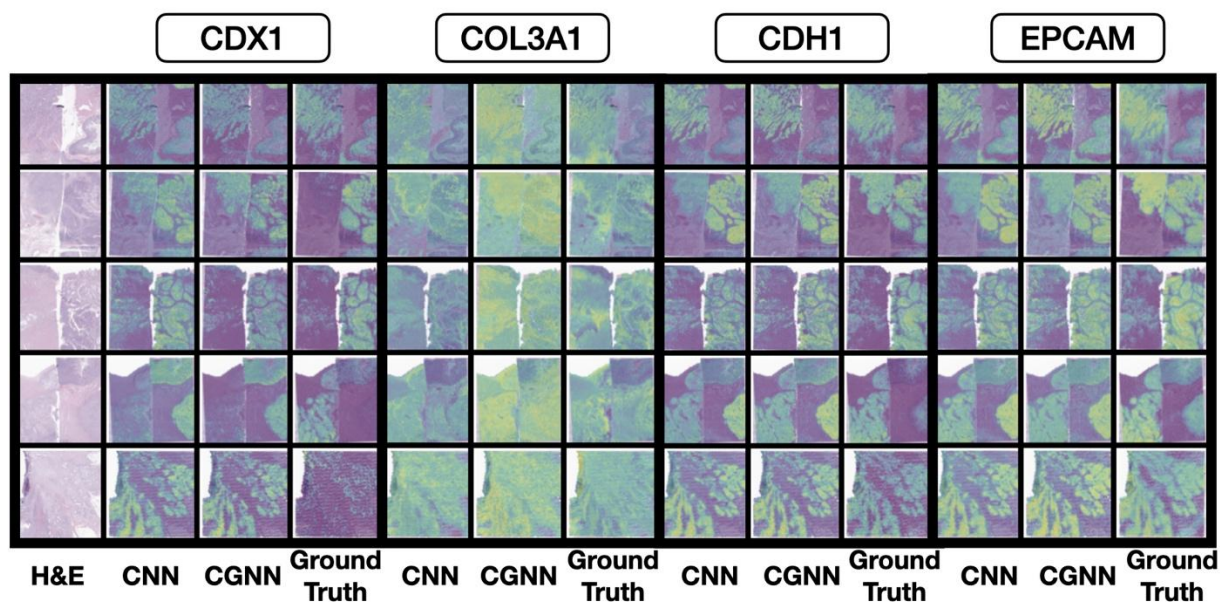
250



251

252 **Figure 3: Performance comparison between methods. A)** Boxplot of AUROC scores from each  
 253 method; **B)** comparison of AUROC for best CGNN and CNN for each gene

254



255

256 **Figure 4: Predicted expression for various genes: CNN, CGNN, compared to ground truth for genes**  
 257 *CDX1, COL3A1, CDH1* and *EPCAM* across sections from all nine patients

## 258 Single-Cell Attribution Maps Point to Spatial Cellular Heterogeneity

259 Single-cell regularization was able to significantly improve the alignment of cellular information extracted  
 260 from located cells within held-out slides with their corresponding single-cell profiles, as measured by the  
 261 Earth Mover's (Wasserstein) distance between cells assigned to spots using Tangram and their closest

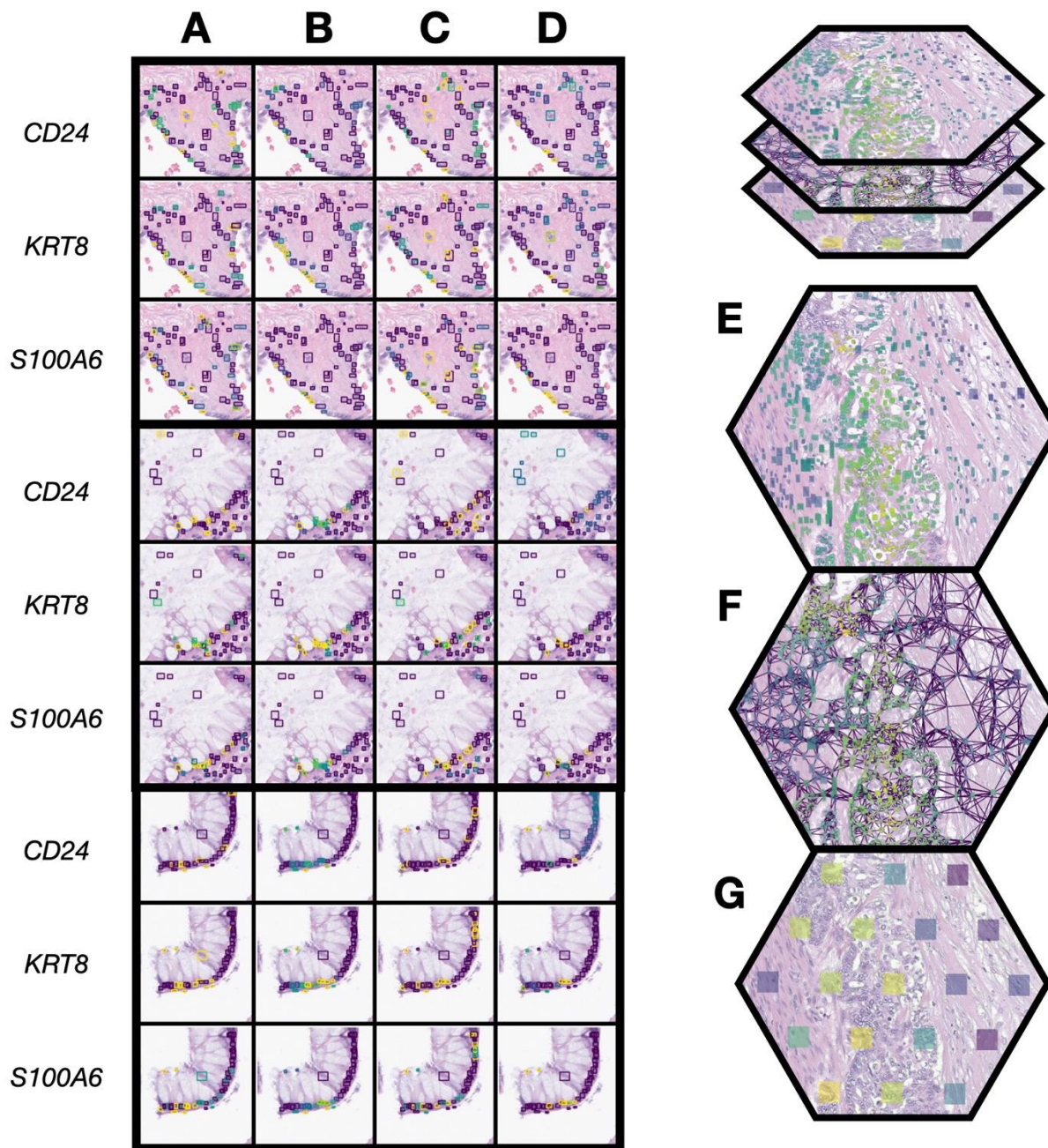


262 detected matches ( $EMD=0.1415 \pm 0.0018$  with penalization,  $0.2113 \pm 0.0018$  without penalization). This  
263 improvement does not negatively impact AUROC. Cells were embedded using UMAP based on the  
264 ground truth and predicted expression, with and without penalization with scRNASeq. Visual inspection of  
265 these UMAP embeddings confirmed the quantitative results of differences in EMD (**Supplementary**  
266 **Figure 2**), that single-cell penalization causes node-level predicted expression from cellular  
267 histomorphology for genes to more closely resemble the distribution of single-cell data assigned to the  
268 Visium spot with Tangram.

269  
270 An optimal transport approach matched single-cell profiles between the predicted and ground truth. We  
271 then assessed the correlation between individual cells for specific genes based on their performance on  
272 the Visium spot level. Overall, more than 80% of the genes exhibited a positive correlation between  
273 ground truth and predicted single-cell expression when single-cell regularization was employed,  
274 compared to around 20-30% of the genes without such regularization was not used (**Supplementary**  
275 **Figure 3**). Examining inferred cell-level expression before the final pooling layer makes the source of our  
276 network's predicted gene expression evident. As illustrated in **Figure 5E-G**, we juxtapose the predicted  
277 level of EPCAM expression for each cell against ground truth data from a Visium assay. Notably, the  
278 Visium assay offers aggregate expression measurements spanning a broader region. Our model's  
279 predictions and the ground truth at cellular resolution are visually consistent (**Figure 5A-D**), further  
280 corroborating with both the high accuracy results reported from the previous section as well as the lower  
281 EMD reported through single-cell penalization.

282  
283 Cell-level attribution (i.e., inferred expression) maps mirror tissue architecture. A closer inspection of a  
284 tissue section reveals patterns, such as cells exhibiting high expression juxtaposed with a cluster of low-  
285 expression cells. Such patterns potentially highlight activities conforming to the underlying tissue  
286 structure. The structural coherence of granular cell-level information likely stems from our regularization  
287 strategies, compelling the neural networks to internalize a general representation of cell graphs and the  
288 sheer volume of available data. In instances where Visium spots exhibit diverse expression levels yet  
289 have similar cell graphs, the graph neural network may pinpoint certain cells that act as primary

290 influencers or representatives of the expression within that spot (i.e., have markedly similar/different  
291 expression for specific genes). These specific cells could either epitomize the overall expression profile or  
292 significantly sway the expression heterogeneity at that location. Further validation is necessary to  
293 corroborate these findings, and this will be a central focus of our future research.

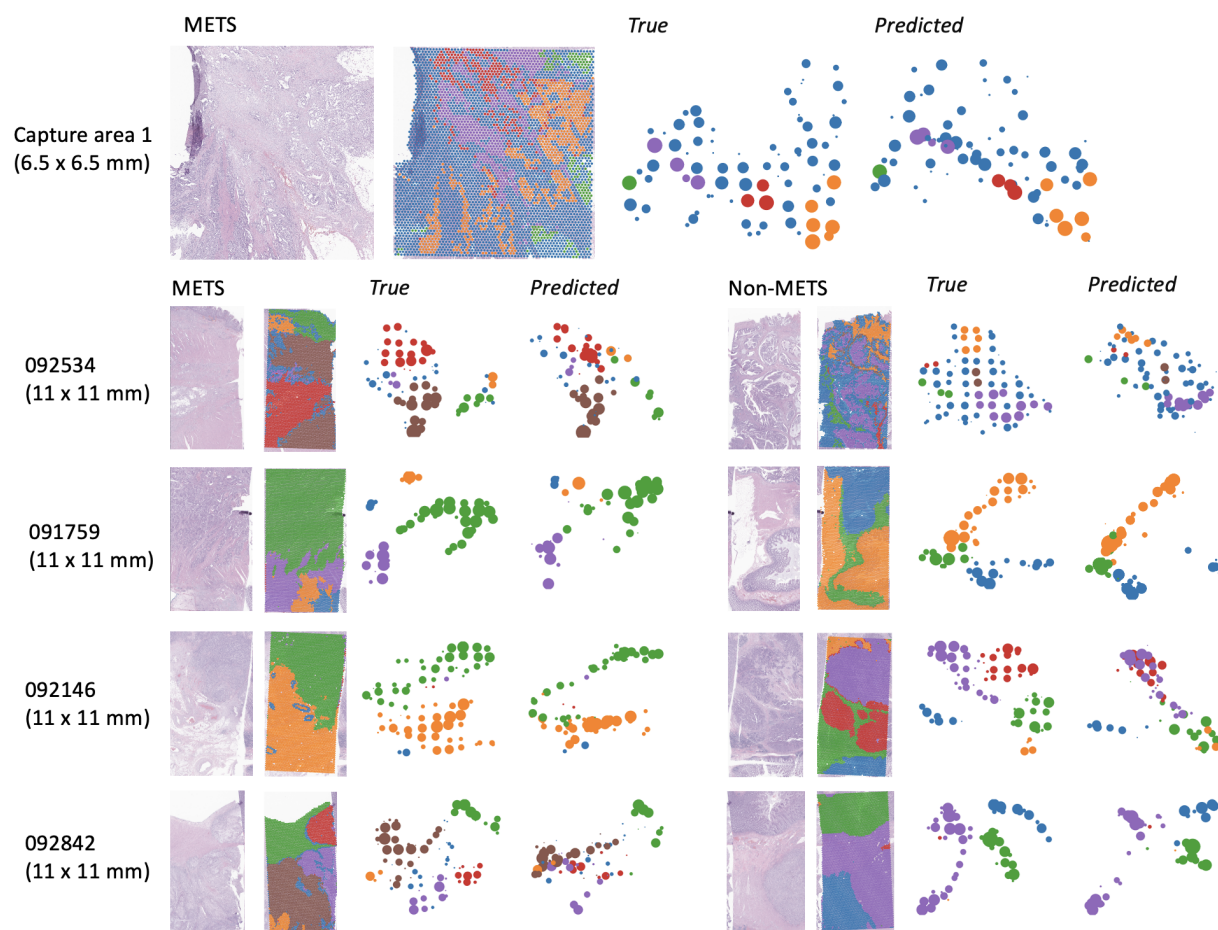


294  
295 **Figure 5: Alignment of True and Predicted Single-Cell and Visium-Spot Level Expression on a**  
296 **Histological Section.** The figure illustrates the relationship between true and predicted single-cell  
297 expression on a histological section for genes CD24, KRT8, and S100A6. Ground truth expression was

298 assigned to each cell using optimal transport based on comparing predicted and actual expressions  
299 rather than their spatial arrangement. A) and C) display the ground truth of single-cell expression with and  
300 without single-cell regularization, respectively. B) and D) visualize the respective predicted single-cell  
301 expressions. Progressing from individual cellular predictions to a broader view, D)-G) detail the transition  
302 through EPCAM expression: from predicted cell-level expression in D) to an overarching cell graph across  
303 multiple Visium spots in E) and concluding with spot-level Visium expression in G).

## 304 Topological Consistency of Inferred Expression Patterns

305 Across all capture areas, predicted spot level expression clustered similarly to the true expression as the  
306 relative positioning between the true and predicted clusters in the UMAP plots was preserved (**Figure 6**).  
307 However, there were notable differences. Overlaying the clusters assigned to ground truth embeddings  
308 over the predicted expression embeddings, we found that clusters were less separated, more connected,  
309 clumped, and fuzzier than the ground truth. Nonetheless, overlaying cluster assignments across the  
310 whole slide image demonstrates the ability of these models to derive expression signatures that can  
311 delineate key histological architectures that will be the subject of inquiry in future work.



312



313 **Figure 6: UMAP embeddings of tissue slides from selected capture areas, color-coded by**  
314 **HDBSCAN clusters.** Comparisons include CGNN, CGNN with single-cell penalization, and patch-based  
315 methods against the ground truth. Clusters derived from the ground truth are overlaid on the slides for  
316 context.

## 317 Pathway Analysis

318 To compare performance across several potential prediction targets, we selected pathways from  
319 MSigDB's Hallmark Gene Sets [35,36] and reported the average AUC for genes from these sets. Across  
320 all modeling approaches, genes involved in DNA repair and E2F targets were predicted with higher  
321 performance as compared to other molecular pathways (**Supplementary Figure 4**). Dysregulation of  
322 DNA repair can impact the acceleration of tumor progression [37], and therefore accurately detecting the  
323 presence of relevant genes may be useful in the prognostication of the speed of progression. We did  
324 notice that for some pathways, e.g., *Epithelial to Mesenchymal Transition*, penalizing by single-cell  
325 expression did lead to some loss of performance in distinguishing these molecular signatures  
326 (**Supplementary Figure 4**).

327  
328 We performed a pathway analysis by subsetting the top 10% of genes (100 out of 1000) per modeling  
329 approach for further analysis using the Enrichr software/database [38–41]. The top ten statistically-  
330 significant pathways (from the Elsevier pathway collection), divided by each modeling approach, are  
331 shown in **Supplementary Table 2**. We also found that the *WNT in Epithelial to Mesenchymal Transition*  
332 *in Cancer* pathway, a chief contributor to the migration and metastasis of cancer cells, and several  
333 pathways associated with *desmosome assembly* (which regulate intercellular adhesion between  
334 metastasizing cells) were among the top ten most statistically significant gene sets detected in all four  
335 techniques, and *EPCAM in Cancer Cell Motility and Proliferation* is a statistically significant gene set in all  
336 four techniques. The *WNT in Epithelial to Mesenchymal Transition in Cancer* pathway has an AUROC of  
337  $0.8686 \pm 0.0273$  for the Inception model and  $0.8638 \pm 0.0238$  for the "vanilla" cell graph model.

338



## 339 Discussion

340 In this research, our primary objective was to draw inferences about spatial mRNA expression patterns  
341 from whole slide images (WSI), specifically by fusing single-cell histological and transcriptomic data. One  
342 of the key advantages of deriving such spatial information from routine stains is the potential to  
343 substantially reduce the costs associated with understanding the tumor immune microenvironment  
344 (TIME). Instead of relying on expensive spatial molecular staining for multiple targets, this technique  
345 offers an economical avenue for spatial molecular assessment, which can subsequently aid in the risk  
346 evaluation of recurrence. It is becoming clear that the spatial positioning, functional status, and mere  
347 presence of immune cells within TIME are crucial determinants of the tumor's immune response. We  
348 pioneered a cell graph neural network algorithm in response to this understanding. This was designed to  
349 meticulously analyze cell structures within WSI and associate them with specific gene expressions. The  
350 development of this algorithm highlights the viability of utilizing spatial transcriptomics as a rich pretraining  
351 source, using scRNASeq to guide single-cell level interpretations that could benefit from graph-based  
352 representations. As we merged scRNASeq data with detailed histological imaging, we also introduced  
353 attribution methods that aspire to clarify the expression patterns evident in Visium spots. Even though  
354 these methods are still evolving and need further optimization and comprehensive validation, they usher  
355 in a promising avenue to understand the nuanced relationship between spatial positioning and gene  
356 expression. Importantly, our work underscores the potential of harnessing neural networks to derive  
357 molecular insights directly from tissue histology without necessitating additional assays, even though  
358 such assays played a critical role in training our models.

359  
360 Our study revealed that by considering cells' histomorphology and spatial relationships, we could  
361 effectively predict gene expression patterns across whole slide images. In some instances, these  
362 approaches outperformed traditional patch-based computer vision methods that rely on regression  
363 models using cropped images around each Visium spot. However, the predictive capacity of these  
364 approaches was found to be similar to patch-based methods, which is not surprising considering that the  
365 cells are contained within these patches and should present some loss of information. By explicitly

366 incorporating cells as nested observations, attribution methods enabled the identification of structural cell  
367 organizations that exhibited the strongest correlation with the expression of specific genes. This finding  
368 has the potential to enhance our understanding of the tumor-immune microenvironment dynamics in  
369 future work.

370

371 *Comparison of Cell-Level Approach to Patch-Based Methods:* The performance of the CNN model does  
372 not surpass that of the cell-based approaches. Interestingly, our basic cell model, devoid of any  
373 augmentations or pretraining, demonstrates a bootstrapped AUROC confidence interval overlapping with  
374 that of the Inception model. This indicates that, even when operating with potentially less diverse  
375 information like the extracellular matrix and connective tissue, the cell-based model remains competitive  
376 against its CNN counterpart. We posit that patch-based CNNs have an inherent advantage due to their  
377 richer input dataset and innate spatial reasoning granted by 2D convolutional kernels. In contrast, graph  
378 neural networks (GNNs) rely exclusively on adjacency matrices for spatial information. The crux here is  
379 the trade-off between robustness and interpretability. Although CNN may show a slight performance  
380 advantage, its insights are limited to single-pixel attributions, neglecting the broader scope of cell-cell  
381 interactions. Conversely, the GNN model offers superior explainability, permitting direct visualization of  
382 pivotal cell-cell interactions for particular genes, with tools like Captum, GNNExplainer, and topological  
383 methods that can be used to decipher important structural motifs.

384

385 *Impact of Single-Cell Penalization:* Several of our experiments, including single-cell penalization and  
386 contrastive pretraining, showed minimal influence on the final outcome. This lack of influence indicates  
387 that employing single-cell penalization can shed light on the spatial nuances of cellular disparities without  
388 compromising performance, emphasizing the significance of cellular phenomena and connectivity. We  
389 believe this is due to the large dataset size (more than 60,000 Visium spots), which may mitigate the  
390 need or potential benefit of pretraining. Additionally, although we hoped that single-cell penalization would  
391 improve the model's robustness (by grounding predictions in real single-cell RNA quantification), the  
392 penalization provided modest performance gains over methods that did not employ this penalization. This  
393 modest gain suggests that models may produce the same optimum regardless of the intermediate feature

394 values (i.e., cell-level predictions). However, we do note that single-cell penalization causes cell-level  
395 values to be more aligned with true single-cell data, which may indicate potential future applications in  
396 deconvolution at single-cell resolution pending further validation. Cell regularization seems to have a  
397 slight negative impact across all pathways, but this might be expected since it more directly impacts node  
398 predictions rather than graph predictions. Without this regularization, the model is more directly trained  
399 towards Visium alignment, on which it is validated. However, with the regularization, we notice that  
400 agreement between node-level predictions and single-cell profiles improves (as measured by a decrease  
401 in mean-squared error). Therefore, we propose that this method would be useful in exploring the potential  
402 for graph neural networks in cell deconvolution.

403  
404 *Revisiting Topological Consistency and Intermediate Histologically-Associated Molecular States: We*  
405 discovered that although the predicted expression patterns mirrored the essential topological  
406 relationships tied to specific histological structures, they were more intertwined compared to the true  
407 expression, resulting in less pronounced clustering. Such mixed clustering might suggest that these  
408 clusters signify different degrees of cellular activity for various phenomena. It seems easier for machine  
409 learning models to distinguish between low and high activity levels, but interpolating intermediate levels of  
410 activity poses a challenge from a visual standpoint. In some instances, while the ground truth UMAP  
411 expression plot showed nuanced expression, the predicted expression indicated centralized expression  
412 around specific profiles. This implies a potential loss of intricate genetic interrelations, which is expected  
413 when deducing genomic data mainly from visual information in histology. This nuanced loss in the  
414 distinctiveness of the models' topological embeddings could indicate the limitation of extracting molecular  
415 data from histology alone, suggesting richer molecular details exist beyond just histology. A potential  
416 enhancement might involve using single-cell expression profiles to better guide the model toward a more  
417 accurate gene distribution representation. Nevertheless, overall, the model's predictions are topologically  
418 in line with the ground truth. Areas of tissue with similar ground truth measurements also exhibit similar  
419 predicted expressions.

420

421 *Reflections on Pathway Analysis: The WNT in Epithelial to Mesenchymal Transition in Cancer and*  
422 *EPCAM in Cancer Cell Motility and Proliferation* were notable pathways from the results section. Wnt/ $\beta$ -  
423 catenin signaling is implicated in cell differentiation and proliferation and has been implicated in  
424 increasing the number of "stem-like" cells in a tumor [42]. EPCAM is responsible for modulating epithelial  
425 cell adhesion, and - while having conflicting trends in recent research - can result in adhesive and  
426 migratory cell activity, potentially impacting the potential for metastasis [43].

427  
428 *Immunological Considerations: Our approach to unveil single-cell heterogeneity from whole slide images*  
429 through alignment with single-cell expression bears several important immunological implications. First,  
430 the spatial arrangement of immune cells not only influences processes governing the anti-tumoral  
431 response but may offer insights as to the efficacy of immunotherapies including checkpoint inhibitors  
432 which has been a timely subject of inquiry [44,45]. Deciphering the spatial make-up may also further  
433 reveal how tumors can establish immunosuppressive environments or contribute to an immune exhaustion  
434 phenotype [46–48]. These topics underscore work being done to study how tumors can alter their  
435 immunogenicity and immune evasion tactics, potentially informing CAR T-cell therapies or selection of  
436 specific antibodies which can be applied in a personalized manner [49–51]. Revealing additional  
437 heterogeneity may refine the selection of adjuvant therapy choices outside of existing prognostic  
438 measures (e.g., pTNM staging) depending on the kind of immune reactivity.

439  
440 *Limitations and Future Directions: Our study, while promising, has several limitations that offer avenues*  
441 for future research. First, the generalizability of our detection neural network models may be constrained  
442 due to the limited sample size of the Lizard dataset used to train the cell detection models, which may not  
443 be representative of the broader colorectal tumor population. Inaccurate cell detection may have  
444 hampered the predictive performance of the cell-graph neural networks. Additionally, while our cohort of  
445 nine samples is large for a Visium study, given its cost, plans are underway to amass a larger, more  
446 diverse cohort to bolster the robustness of our findings by accounting for further tumor heterogeneity. As  
447 our cohort was restricted to pT3 patients, future work will examine the predictiveness of these algorithms  
448 at additional tumor sites and levels of invasiveness, particularly in relation to the tumor-immune

449 microenvironment. Another limitation is that our cell graphs solely relied on local connectivity for  
450 information flow, lacking positional embeddings or integration with patch-level information. The slight  
451 performance advantage of convolutional neural networks (CNNs) over graph neural networks (GNNs)  
452 could be attributed to CNNs' inherent structural benefits. Future work could explore the benefits of deeper  
453 or more extensive cell graphs and the incorporation of single-cell RNA priors for more accurate  
454 predictions. While cell regularization and similar strategies might not enhance the ultimate predictive  
455 capability for pooled Visium expression, they do bolster performance concerning optimal transport  
456 between single-cell profiles assigned to Visium spots using deconvolution methods (like Tangram). Direct  
457 model predictions can display notable correlations in gene expression levels across the cell distribution  
458 within each spot, even without an input single-cell profile. In the future, these methods could serve as  
459 adjustable enhancements or alternatives to Tangram. Inaccurate mapping of single-cell profiles to Visium  
460 spots may have also impacted the validity of the associations between single-cell inferred expression and  
461 could improve with the adoption of other single-cell spatial mapping methods. Potentially this approach  
462 could be used to map single cell profiles, where scRNASeq information has been collected, to their  
463 precise locations within slides without the use of Visium as an intermediary. Our approach could also  
464 extend to identify other spatially-resolved molecular features, such as protein interactions [52,53].  
465 Validation of our findings using pathological examination of underlying cellular phenomena, through  
466 immunohistochemistry and other single-cell level spatial analysis platforms like Xenium/merScope is  
467 essential [54,55], as this study serves as a proof of concept. We anticipate applying our models to a  
468 larger cohort to identify metastasis and recurrence predictors at a population scale. Furthermore, the  
469 trade-off in model accuracy is mitigated by the potential for scalability, allowing for the assessment of a  
470 larger number of genes with reasonable predictive performance and statistical power. Overall, our study  
471 signifies a crucial step towards improving cancer diagnostics and prognosis by incorporating spatial  
472 transcriptomics into histological images, and future efforts will focus on refining these techniques.  
473  
474 Ultimately our findings could shed light on the molecular alterations occurring in these immune cell  
475 subsets, which may help identify potential immune evasion mechanisms employed by tumor cells, such  
476 as the upregulation of immune checkpoint molecules or the recruitment of immunosuppressive cell

477 populations. By integrating this spatial and molecular information, we can potentially uncover the  
478 functional "hotspots" within the tumor microenvironment, where effective immune responses are either  
479 promoted or hindered. These functional "hotspots" could ultimately aid in designing immunotherapeutic  
480 strategies that aim to reinvigorate the anti-tumor immune response, either by modulating the activity of  
481 immune cells in these hotspots or by altering the spatial organization of immune cells within the tumor  
482 microenvironment.

## 483 Conclusion

484 Our investigation into the spatial patterns of T cells, NK cells, and B cells near the primary tumor site and  
485 their molecular alterations has potential implications for our understanding of tumor microenvironment  
486 dynamics and the broader clinical and immunological context. These patterns can offer valuable insights  
487 into the potential for concurrent nodal and/or distant metastasis and serve as indicators of the risk of  
488 recurrence and mortality. They provide critical complementary information to traditional lymph node  
489 assessment screening programs. Understanding these patterns can significantly influence disease  
490 management strategies for patients with colorectal cancer, as the immune response is intrinsically tied to  
491 the disease's pathogenesis and progression. The promise of spatially inferring gene expression patterns  
492 from routine histological staining within larger cohorts could vastly enhance our understanding of specific  
493 transcriptomic changes occurring within distinct spatial architectures associated with metastasis.  
494 Traditionally, conducting highly multiplexed spatial profiling in such cohorts has been cost-prohibitive.  
495 However, with the advent of new methods, it is now feasible to explore and analyze spatial transcriptomic  
496 data at a previously challenging scale. This advancement opens up unprecedented possibilities for more  
497 cost-effectively investigating spatial cell-type specific changes associated with metastasis.

498  
499 Our study revealed that by considering cells' histomorphology and spatial relationships, we could  
500 effectively predict gene expression patterns across whole slide images and recover local patterns of  
501 cellular heterogeneity. Identifying structural cell organizations that exhibited the strongest correlation with  
502 the expression of specific genes has the potential to drastically improve our understanding of the tumor-  
503 immune microenvironment dynamics and potentially guide personalized treatment plans. Future

504 applications of this method could include predicting response to immunotherapy based on the spatial  
 505 distribution and expression patterns of immune cells in the tumor microenvironment. In conclusion, our  
 506 study underscores the potential of integrating spatial transcriptomic information into histological images  
 507 using feature engineering approaches. It offers a promising direction for enhancing not only the diagnosis  
 508 and prognosis of cancer but also our broader understanding of the clinical and immunological intricacies  
 509 of the tumor microenvironments. As the field of spatial transcriptomics advances, we anticipate that this  
 510 methodology will play an increasingly pivotal role in shaping personalized and precision medicine  
 511 strategies for cancer treatment.

512

## 513 Methods

514 **Table 2: Patient characteristics:** This table details nine unique patients' profiles across five capture  
 515 areas. Four of these areas are divided equally between two patients, while one is dedicated to a single  
 516 patient. Critical patient data, including microsatellite stability status, metastasis occurrence, and tumor  
 517 site, are provided for each individual.

Capture Area	Dimensions (mm)	Section Placed on Left Side / Center of Capture Area					Section Placed on Right Side of Capture Area				
		Age	Sex	Tumor Site	MSI Status	Metastasis	Age	Sex	Tumor Site	MSI Status	Metastasis
1	6.5 x 6.5	80-85	M	Left Colon	MSI	yes	–	–	–	–	–
2	11 x 11	90-95	F	Hepatic Flexure	MSI	yes	80-85	M	Left Colon	MSI	no
3	11 x 11	85-90	F	Splenic Flexure	MSS	yes	75-80	F	Hepatic Flexure	MSS	no
4	11 x 11	80-85	F	Cecum	MSI	yes	70-75	M	Cecum	MSI	no
5	11 x 11	55-60	F	Left Colon	MSS	yes	65-70	M	Sigmoid	MSS	no

518

## 519 Data Collection and Preprocessing

520 The dataset used in this study comprised nine patients with pathologic T Stage-III (pT3) colorectal cancer.  
 521 Following IRB approval, these patients were selected through a retrospective review of pathology reports  
 522 from 2016 to 2019. Notably, this cohort is distinct from our initial cohort of four slides, profiled using the  
 523 original Visium assay without using CytAssist.

524

525 Patients were matched based on various criteria such as age, sex, tumor grade, tissue size, mismatch  
 526 repair/microsatellite instability (MMR/MSI/MSS) status, and tumor site. MSI status was determined

527 through the loss of expression of MLH1 and PMS2 as assessed through immunohistochemistry (there  
528 was no loss of expression for MSH2 and MSH6).

529

530 The dataset included both patients with concurrent tumor metastasis and patients without metastasis,  
531 with equal representation from each group. Patients were carefully selected to ensure a balanced  
532 representation across various factors, including tumor site, grade, node status, MSI status, and sex.  
533 Tissue blocks were sectioned into 5-micron thick layers. Specific regions of interest within these sections,  
534 including epithelium, tumor-invasive front, intratumoral areas, and lymphatics, were annotated by a board-  
535 certified GI pathologist. Following annotation, these regions were dissected from the tissue, and  
536 subjected to H&E staining, imaging, and Visium profiling at the Pathology Shared Resource at Dartmouth  
537 Cancer Center and Single Cell Genomics Core in the Center for Quantitative Biology.

538

539 *Enhanced Pathology Workflow for Visium Profiling:* In this study, we utilized the enhanced workflow  
540 described in(10)., to profile specimens utilizing the Visium assay. Forv four slides, which corresponded to  
541 eight patients, we macrodissected tissue sections from Formalin-Fixed Paraffin-Embedded (FFPE) blocks  
542 into 5.5mm by 11mm rectangular segments—precisely delineated by the pathologist in serial Whole Slide  
543 Images (WSI)—to isolate specific tissue architectures. To optimize study expenses, rectangular tissue  
544 segments from two different patients were juxtaposed at the center of a standard histology slide to form  
545 11mm<sup>2</sup> capture areas (each corresponding to approximately 14,300, 50-micron Visium spots) and then  
546 secured with a coverslip. Within each of these dual-patient capture areas, we maintained an equal  
547 representation of metastasis and MSI status and ensured each capture area comprised tissue from  
548 similar anatomic sites. For one of these slides, we profiled approximately 5,000 50-micron Visium spots  
549 within a 6.5mm by 6.5mm capture area. Our analysis involved a total of five tissue capture areas,  
550 representing nine distinct patients. The characteristics of the patients, such as age, sex, and tumor site  
551 for these batches, are provided in **Table 2**.

552

553 To achieve uniform staining and enhance image quality, we incorporated the CytAssist workflow, which  
554 allows Visum profiling of tissues on standard histology slides, enabling the use of pathology department



555 automated staining (Sakura Tissue-Tek Prisma Stainer– Sakura Finetek USA, Inc. 1750 West 214th  
556 Street, Torrance, CA 90501) and advanced WSI at 40x resolution (0.25 micron per pixel) via Aperio  
557 GT450s to obtain reproducible, high-quality images. Following the preparation of the tissue slides, we  
558 employed the Visium assay using the CytAssist technology according to the manufacturer's protocol  
559 (CG000495) [56,57]. Here, transcriptomic probes are hybridized to the tissue section and ligated to form  
560 an amplifiable product. CytAssist facilitates the precise transfer of the ligated probes from the tissue slide  
561 to the Visium CytAssist Spatial Gene Expression slide (GEX), which harbors spatially-barcoded capture  
562 oligonucleotides, bearing a poly-A sequence complementary to the probes. During transfer, CytAssist  
563 images the tissue within the fiducial frame of the GEX slide, allowing for the co-registration of 40x H&E  
564 tissue images with Visium ST spots. transferred probes are then extended and amplified prior to  
565 sequencing on an Illumina NovaSeq instrument, targeting 50,000 reads/spot. This comprehensive  
566 procedure allows for unbiased and gridded profiling of up to 5,000 spots, each containing 1-10 cells,  
567 within a 6.5 mm<sup>2</sup> capture area, and up to 14,300 spots within an 11 mm<sup>2</sup> capture area. For data  
568 processing, we utilized Spac Ranger V to align the CytAssist images with the corresponding 40X H&E  
569 stains, conduct quality control, and convert the Visium Spatial Transcriptomics (ST) data into a spots x  
570 genes expression matrices used for downstream analysis [58].

571  
572 *Single Cell Profiling:* We utilized the Chromium Flex assay to acquire single-cell RNA-Seq data,  
573 specifically from serial sections of patients identified in Capture Areas 2 (left section) and 5 (right section),  
574 as detailed in **Table 2**. This method allows for single cell profiling of disaggregated FFPE tissue sections  
575 using the same transcriptomic probe set as the Visium assay, thereby revealing the diverse cell types  
576 within the tissue. This method was performed according to the manufacturer's Demonstrated Protocol  
577 (CG000606). Data were processed using Cell Ranger v7.1.0 to generate quality control metrics and a  
578 cells by genes expression matrices for downstream processing.

579

## 580 Preprocessing and Augmentation

581 *Selection of Gene Prediction Targets:* We curated a list of 1,000 target genes by initially filtering out those  
582 not appearing in at least 100 spots per patient. These genes were subsequently ranked based on the  
583 fraction of their spatial variance, as determined through SpatialDE analysis. To rectify aberrant gene  
584 expression levels, we applied a transformation to both prediction and target gene counts using the  
585 expression  $\log(1 + \text{counts})$ .

586

587 *Extraction of Single Cell Imaging Information:* Cell detection was performed using the Mask-RCNN  
588 framework, which was trained on both the Lizard dataset and our internal dataset [59–61]. The nuclei  
589 detection model, available through the public Detectron2 Model Zoo, served as our pre-trained base. This  
590 model was further trained on our dataset for up to 5,000 epochs, with training halted upon the observation  
591 of overfitting (indicated by a peak in mAP on the validation set). The architecture employed was a Mask  
592 RCNN with a Residual Network + Feature Pyramid Network (ResNet+FPN) backbone based on the  
593 ResNet-101 model. After training, this cell detection model was systematically applied across each Whole  
594 Slide Image (WSI).

595

596 *Image Normalization and Augmentation:* The associated image was normalized for each detected cell  
597 through standard scaling applied over the image channels. We implemented data augmentation  
598 techniques to enhance our dataset, including random rotations (up to 90°) and color jitter adjustments.  
599 These augmentations were specifically applied to the images and cell detections cropped around the  
600 Visium spots during the training phase.

601

## 602 Deep Learning to Integrate Information from Localized Cells to Predict 603 Spatial Gene Expression

604 Cell graph neural networks (CGNN) facilitate the exchange of messages between adjacent cells, enabling  
605 the exchange/incorporation of contextual information [62–67]. This approach effectively captures the  
606 relationships between different cell populations within the tissue, including tumor cells and surrounding  
607 immune and other cell subpopulations. The important difference between this approach and a more  
608 general convolutional neural network model is that cells and the relationships between them are modeled  
609 explicitly. Leveraging these relationships can enhance the predictive performance of our spatial RNA  
610 inference algorithms while providing additional information as to relevant cells for these predictions.

611

612 We implemented an end-to-end training strategy that integrates the simultaneous training of a  
613 Convolutional Neural Network (CNN) and a Graph Neural Network (GNN). The CNN is designed to  
614 extract cell-level features from histological images, while the GNN contextualizes these features by  
615 incorporating information from neighboring cells during the model-fitting process. This stands in contrast  
616 to previous cell-level neural network models, which typically employ a two-stage approach. In these two-  
617 stage models, features are initially extracted using a CNN, and these extracted features serve as input to  
618 train the GNN in a separate, subsequent step. Such a two-stage paradigm can potentially lead to  
619 suboptimal results, as the CNN's parameters are fixed before the GNN training begins, thereby limiting  
620 the CNN's ability to adapt based on the contextual information that the GNN could provide. Our end-to-  
621 end approach aims to harmonize the feature extraction and contextualization processes, enabling the  
622 CNN to learn cell-level features that are more effectively contextualized through iterative, integrated  
623 training with the GNN (**Figure 1,2**).

624

625 The backbone of the model is a four-layer graph attention network (GAT) [68,69], which uses self-  
626 attention mechanisms to update the representation of each cell with the information of its neighbors. We

627 extract nodal attributes from detected cells using a ResNet-50 model, which is trained jointly with the  
628 graph attention layers. The Euclidean distances between the spatial locations of detected cells are used  
629 to form k-nearest-neighbor cell graphs (k=6, determined through a sensitivity analysis), connecting cells  
630 based on their spatial proximity. Patients within the same Visium capture area were divided according to  
631 pathologist-annotated segmentation masks— we ensured that cell graphs did not overlap across patients.  
632 The model updates cell-level embeddings through information sharing up to four neighbors away,  
633 facilitated by message passing between nodes across four graph attention layers, each mapping cells to  
634 512-dimensional numerical vectors. Final node embeddings pass through a linear layer, producing a  
635 vector representing each gene's relative pseudocount-transformed expression for each cell. Cells  
636 corresponding to the same Visium spot are then aggregated through global sum pooling to predict  
637 expression at the Visium spot. This is then compared to the pseudocount-transformed ground-truth  
638 Visium data with mean squared error.

639

## 640 Comparison of Cell-Graph Neural Network Regularization Strategies

641 In addition to evaluating the congruence between ground truth and predicted expression at the spot level,  
642 we explored the following methodological variations:

- 643 1. **Vanilla Supervised Learning Objective:** This baseline approach focuses solely on the  
644 supervised learning objective, serving as a reference for evaluating the potential gains from  
645 additional regularization strategies.
- 646 2. **Incorporating Graph Contrastive Learning:** This approach introduces a self-supervised  
647 regularization term that encourages the model to learn embeddings through the comparison of  
648 augmented viewpoints of the same cell graph / Visium spot to different cell-graphs / Visium spot.  
649 This can enhance the model's sensitivity to spatial patterns in the data, potentially improving its  
650 predictive accuracy for spatial transcriptomics patterns.
- 651 3. **Incorporating Single-Cell RNA-Seq Penalization through Optimal Transport:** This strategy  
652 introduces a penalty term that encourages the model to align cell-level histological features more  
653 closely with corresponding single-cell RNA-Seq data. By leveraging optimal transport theory, this

654 term effectively "guides" the model towards a solution where the spatial patterns inferred from  
655 histology are maximally consistent with independent single-cell RNA-Seq measurements, thereby  
656 enhancing the biological validity of the model's predictions.

657 4. **Combining Graph Contrastive Learning and Single-Cell Penalization:** This approach  
658 synergistically combines both the graph contrastive learning and the single-cell RNA-Seq  
659 penalization strategies, aiming to leverage the benefits of both spatial context awareness and  
660 alignment with single-cell RNA-Seq data. This dual-regularization strategy is designed to promote  
661 a model that is both sensitive to spatial patterns and tightly aligned with independent molecular  
662 measurements, potentially offering a balance between spatial sensitivity and biological validity.

663

## 664 Graph Contrastive Learning

665 Using the PyGCL package, graph contrastive learning was implemented through augmentations to  
666 random cell positions in the nearest neighbor graph construction, dropping edges with a probability of 0.1,  
667 and masking out features with a probability of 0.3. Graph contrastive learning is a form of self-supervised  
668 learning that can improve the generalizability and robustness of graphs [70–72]. By intentionally adding  
669 noise to the training cell graphs and comparing these representations at different Visium spots, we aimed  
670 to improve the model's generalizability when tested on held-out data by forcing it to learn a stable, robust,  
671 foundational representation.

672

## 673 Incorporating Single-Cell Expression

674 For two of our patients in the study, we obtained corresponding single-cell RNA-Seq data, which were  
675 selected to be representative across sex, microsatellite instability, and metastasis status. Due to budget  
676 constraints, we were unable to acquire matching single-cell data for all the slides, a limitation we aim to

677 address in future work. Our approach involves the integration of single-cell RNA-Seq data as a form of  
678 model regularization. By encouraging the predictions derived from histological images of individual cells  
679 to align closely with the corresponding true single-cell expression profiles, we aim to enhance the  
680 interpretability of our models through more consistent and biologically meaningful cellular information.  
681 This, in turn, is expected to improve the model's generalizability. The objective is to develop a nuanced  
682 understanding of cell type populations at specific spatial locations while maintaining fidelity to the ground-  
683 truth expression data sampled through the Visium platform. By fine-tuning the alignment between these  
684 profiles, we strive to increase the likelihood that our predictions accurately reflect the true cellular  
685 composition at each spatial location. In contrast to previous attempts to integrate single-cell information  
686 with histology, where spatial transcriptomics data were not employed to guide the spatial mapping of  
687 single-cell information, our approach uniquely capitalizes on the spatial context offered by the spatial  
688 transcriptomics data. Specifically, we align single-cell profiles with individual Visium spots, thereby  
689 avoiding spurious cell assignments that can arise when attempting to map single-cell information across  
690 entire whole-slide images without the benefit of such spatial guidance.

691  
692 We initiated our analysis by mapping scRNA profiles to Visium spots using the Tangram tool [73], and we  
693 selected the top  $k$  most likely cells to be assigned to each spot, where  $k$  represents the number of  
694 detected cells in that spot. In the context of our CGNN model, which aims to predict log-expression levels  
695 for individual cells, it was critical to assess the alignment between these predictions and the assigned  
696 single-cell profiles. Originally, we employed a dynamic matching approach wherein the pool of cells  
697 assigned to each spot was paired with our detected cells based on the Euclidean distance between their  
698 predicted and observed expression profiles. This matching was formulated as a variant of the linear sum  
699 assignment problem, solved using the Hungarian algorithm. This algorithm established the optimal one-  
700 to-one correspondence between detected cells and their closest-matching assigned scRNA cells, thereby  
701 minimizing the Euclidean distance between them. We applied a mean squared error (MSE) loss across  
702 the log-transformed expression values to penalize discrepancies between these matched profiles.  
703 However, following internal comparisons, we shifted our approach to leverage the Wasserstein loss as a  
704 more effective metric for aligning our predictive single-cell expression profiles with the true expression

705 profiles derived from scRNA data. The Wasserstein distance, a loss metric formulated on principles from  
706 the theory of optimal transport [74–77], quantifies the minimum cost required to transform one distribution  
707 into another, effectively capturing the overall distributional differences between predicted and ground truth  
708 single-cell expression profiles. This approach offers more flexible comparisons, as it can accommodate  
709 shifts in the distribution and discrepancies in the shape of the distributions without necessitating a direct  
710 one-to-one matching between predicted and observed cells.

711  
712 Our internal comparisons revealed that the Wasserstein loss provided more precise and interpretable  
713 mappings between detected and observed cellular profiles, solidifying it as our chosen approach. The  
714 Wasserstein loss was computed for each Visium spot and used to compare the effectiveness of different  
715 regularization methods, measuring alignment between predicted and ground truth single-cell expression.

716

## 717 Comparison to Convolutional Neural Network Approaches

718 The CGNN approaches were compared to patch-based convolutional neural network methodologies  
719 deemed highly predictive from previous works— namely the InceptionV3 neural network. The Inceptionv3  
720 network can capture features at various receptive fields. This model was trained on images of tissue  
721 patches encompassing multiple cells inclusive of surrounding matrixed tissue architecture. We initialize  
722 the model with ImageNet weights (with the final layer truncated) on the same dataset. We apply the same  
723 visual transformations as for the cell embeddings, (standard scaling, color jitter, random 90° rotations).

724

## 725 Training and Validation

726 CGNN models were implemented with the torch-geometric Python package [78], which extends the  
727 PyTorch machine-learning framework. We use PyGCL [72] to apply graph augmentations. CGNN were

728 trained using the Adam optimizer [79] with a learning rate of 0.0001 on one Nvidia V100 GPU with  
729 Dartmouth Research Computing, quickly converging after two epochs. Similarly, the CNN model was  
730 trained for around 100000 iterations on a Nvidia V100 GPU.

731  
732 The final performances of these models were compared using leave-one-patient-out cross-validation.  
733 Statistics are reported with the Spearman correlation coefficients. We also sought to assess the  
734 performance of predicting dichotomized gene expression (low/high), a binary classification problem— this  
735 was accomplished by dichotomizing expression according to [30], used to calculate the area under the  
736 receiver operating characteristic curve (AUROC) as another performance measure. Models were trained  
737 and then cross-validated using all capture areas except one, reserved for testing, and repeated for all  
738 cross-validation folds. Performance statistics were generated for each cross-validation fold, including  
739 Spearman's correlation coefficients and area under the receiver operating characteristic curves  
740 (AUROCs) by gene. The results were then averaged across all folds to assess the best-performing model  
741 on a gene-specific basis. We calculated 95% confidence intervals for all performance statistics, reported  
742 using 1000 sample non-parametric bootstrapping.

743

## 744 Model Interpretation through Single Cell Attributions, Gene Embeddings 745 and Pathway Analysis

746 *Single-cell attributions:* We visually inspected the CGNN's ability to reconstruct single-cell gene  
747 expression based on the aggregate neighborhood's gene expression. Since the CGNN calculates the  
748 overall expression of a Visium spot by summing predictions for individual cells, we hypothesized that the  
749 predictions for each cell would reflect their relative contribution to the spot's expression. We believed that  
750 the model would learn consistent and meaningful gene expression counts due to the large amount of data  
751 and our aggregation method.

752



753 Operating under this assumption, we generated *attribution maps* representing each cell's expression for  
754 the 1000 genes across the slide. These multidimensional maps provided insights into the expression  
755 patterns of individual cells, even though they were not explicitly labeled for that purpose. To assess the  
756 validity of these attribution maps, we collaborated with a pathologist for their expert evaluation. However,  
757 we acknowledge that further detailed analysis is needed in future work to delve deeper into their  
758 significance and accuracy. To complement these attribution maps, we also generated analogous  
759 *explanation maps*, which utilized the GNNExplainer algorithm to assign importance scores to specific cells  
760 based on their relevance for predictions of specific genes [80].

761  
762 *Gene embeddings:* Similar to our previous work, we sought to intuitively understand how well each  
763 approach could recapitulate the relationships between the Visium spots. This recapitulation was  
764 accomplished by applying Uniform Manifold Approximation and Projection (UMAP) to each predicted  
765 expression profile [81]. Each method's predicted and actual gene expressions were aligned and clustered  
766 using the AlignedUMAP method. Clusters determined by running HDBSCAN [82] on the ground truth  
767 expression data were overlaid on top of the UMAP plots for the other methods. To do this, we created  
768 HDBSCAN clusters with a minimum size of 3 across the log-transformed ground truth Visium data. Then,  
769 we annotated each of our prediction points with the corresponding HDBSCAN cluster of the ground truth  
770 and performed an aligned UMAP, jointly minimizing the distance between similar expressions in the  
771 embedding space and between paired ground truth and true locations. In addition, we annotated our  
772 histology images with the HDBSCAN clusters to interpret the tissue type of origin for each point. The  
773 visual similarity of clustering patterns would indicate an overall similarity of the predicted and ground truth  
774 expression.

775  
776 *Pathway analysis:* Pathway analyses were performed to assess the ability of the methods to capture  
777 broader biological phenomena. Pathway analyses were accomplished using two separate methods: 1)  
778 aggregating the Spearman correlation and AUROC statistics across genes associated with pathways  
779 identified from the MSigDB Hallmarks gene set, and 2) evaluating the enrichment of the highest genes as  
780 ranked using their performance statistics, utilizing enrichR, which employs a modified Fisher's exact test.

781 By examining the average performance across pathway analysis and overlap tests for the top-performing  
782 genes, we can gain insights into which biological phenomena each method effectively represents.

## 783 **Declarations**

784 **Ethics approval and consent to participate.** Human Research Protection Program IRB of Dartmouth  
785 Health gave ethical approval for this work.

786  
787 **Availability of data and materials.** Access to manuscript data is limited due to patient privacy concerns.  
788 All data produced in the present study are available upon reasonable request. Requests should be  
789 directed to senior author Dr. Joshua Levy (email: [joshua.j.levy@dartmouth.edu](mailto:joshua.j.levy@dartmouth.edu)).

## 791 **Competing interests**

792 None to disclose.

## 794 **Funding**

795 JL is funded under NIH subawards P20GM130454 and P20GM104416.

## 797 **Acknowledgements**

798 This study was carried out in the Genomics and Molecular Biology Shared Resource (GMBSR) at  
799 Dartmouth which is supported by NCI Cancer Center Support Grant 5P30CA023108 and NIH S10  
800 (1S10OD030242) awards. Additionally, single cell genomics projects should include the following text (or  
801 similar): "Single cell studies were conducted through the Dartmouth Center for Quantitative Biology in  
802 collaboration with the GMBSR with support from NIGMS (P20GM130454) and NIH S10 (S10OD025235)  
803 awards.

804

## 805 References

- 806 1. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer  
807 statistics, 2020. *CA: a cancer journal for clinicians*. 2020;70:145–64.
- 808 2. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*.  
809 2023;73:17–48.
- 810 3. Cheng E, Ou F-S, Ma C, Spiegelman D, Zhang S, Zhou X, et al. Diet-and Lifestyle-Based Prediction  
811 Models to Estimate Cancer Recurrence and Death in Patients With Stage III Colon Cancer (CALGB  
812 89803/Alliance). *Journal of Clinical Oncology*. 2022;JCO-21.
- 813 4. Baxter NN, Virnig DJ, Rothenberger DA, Morris AM, Jessurun J, Virnig BA. Lymph node evaluation in  
814 colorectal cancer patients: a population-based study. *J Natl Cancer Inst*. 2005;97:219–25.
- 815 5. Senthil M, Trisal V, Paz IB, Lai LL. Prediction of the Adequacy of Lymph Node Retrieval in Colon  
816 Cancer by Hospital Type. *Archives of Surgery*. 2010;145:840–3.
- 817 6. Norwood MGA, Sutton AJ, West K, Sharpe DP, Hemingway D, Kelly MJ. Lymph node retrieval in  
818 colorectal cancer resection specimens: national standards are achievable, and low numbers are  
819 associated with reduced survival. *Colorectal Disease*. 2010;12:304–9.
- 820 7. Liu Q, Luo D, Cai S, Li Q, Li X. P&ndash;TNM staging system for colon cancer: combination of P-stage  
821 and AJCC TNM staging system for improving prognostic prediction and clinical management. *CMAR*.  
822 2018;Volume 10:2303–14.
- 823 8. Nelson H, Petrelli N, Carlin A, Couture J, Fleshman J, Guillem J, et al. Guidelines 2000 for Colon and  
824 Rectal Cancer Surgery. *JNCI: Journal of the National Cancer Institute*. 2001;93:583–96.
- 825 9. de Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic  
826 outgrowth. *Cancer Cell*. 2023;41:374–403.
- 827 10. Amira AT, Mouna T, Ahlem B, Raoudha A, Majid BH, Amel H, et al. Immunohistochemical expression  
828 pattern of MMR protein can specifically identify patients with colorectal cancer microsatellite instability.  
829 *Tumour Biol*. 2014;35:6283–91.
- 830 11. Collier JL, Weiss SA, Pauken KE, Sen DR, Sharpe AH. Not-so-opposite ends of the spectrum: CD8+  
831 T cell dysfunction across chronic infection, cancer and autoimmunity. *Nat Immunol*. 2021;22:809–19.
- 832 12. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, et al. Towards the introduction of the  
833 'Immunoscore' in the classification of malignant tumours. *The Journal of Pathology*. 2014;232:199–209.
- 834 13. Dalerba P, Sahoo D, Paik S, Guo X, Yothers G, Song N, et al. CDX2 as a Prognostic Biomarker in  
835 Stage II and Stage III Colon Cancer. *N Engl J Med*. 2016;374:211–22.
- 836 14. Tarazona N, Gimeno-Valiente F, Gambardella V, Huerta M, Roselló S, Zuniga S, et al. Detection of  
837 postoperative plasma circulating tumour DNA and lack of CDX2 expression as markers of recurrence in  
838 patients with localised colon cancer. *ESMO Open*. 2020;5:e000847.
- 839 15. Saad RS, Ghorab Z, Khalifa MA, Xu M. CDX2 as a marker for intestinal differentiation: Its utility and  
840 limitations. *World J Gastrointest Surg*. 2011;3:159–66.

- 841 16. Chen G, Peng J, Xiao Q, Wu H-X, Wu X, Wang F, et al. Postoperative circulating tumor DNA as  
842 markers of recurrence risk in stages II to III colorectal cancer. *Journal of Hematology & Oncology*.  
843 2021;14:80.
- 844 17. Li H, Jing C, Wu J, Ni J, Sha H, Xu X, et al. Circulating tumor DNA detection: A potential tool for  
845 colorectal cancer management (Review). *Oncology Letters*. 2019;17:1409–16.
- 846 18. Todenhöfer T, Struss WJ, Seiler R, Wyatt AW, Black PC. Liquid biopsy-analysis of circulating tumor  
847 DNA (ctDNA) in bladder cancer. *Bladder Cancer*. 2018;4:19–29.
- 848 19. Moses L, Pachter L. Museum of spatial transcriptomics. *Nat Methods*. 2022;19:534–46.
- 849 20. Hu J, Schroeder A, Coleman K, Chen C, Auerbach BJ, Li M. Statistical and machine learning  
850 methods for spatially resolved transcriptomics with histology. *Comput Struct Biotechnol J*. 2021;19:3829–  
851 41.
- 852 21. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. Spatially resolved, highly multiplexed RNA  
853 profiling in single cells. *Science [Internet]*. 2015 [cited 2021 May 15];348. Available from:  
854 <https://science.sciencemag.org/content/348/6233/aaa6090>
- 855 22. Lewis SM, Asselin-Labat M-L, Nguyen Q, Berthelet J, Tan X, Wimmer VC, et al. Spatial omics and  
856 multiplexed imaging to explore cancer biology. *Nature methods*. 2021;18:997–1012.
- 857 23. Monjo T, Koido M, Nagasawa S, Suzuki Y, Kamatani Y. Efficient prediction of a spatial transcriptomics  
858 profile better characterizes breast cancer tissue sections without costly experimentation. *Sci Rep*.  
859 2022;12:4133.
- 860 24. Schmauch B, Romagnoni A, Pronier E, Saillard C, Maillé P, Calderaro J, et al. A deep learning model  
861 to predict RNA-Seq expression of tumours from whole slide images. *Nature communications*.  
862 2020;11:3877.
- 863 25. Zeng Y, Wei Z, Yu W, Yin R, Yuan Y, Li B, et al. Spatial transcriptomics prediction from histology  
864 jointly through Transformer and graph neural networks. *Briefings in Bioinformatics*. 2022;23:bbac297.
- 865 26. Fatemi M, Feng E, Sharma C, Azher Z, Goel T, Ramwala O, et al. Inferring spatial transcriptomics  
866 markers from whole slide images to characterize metastasis-related spatial heterogeneity of colorectal  
867 tumors: A pilot study. *J Pathol Inform*. 2023;14:100308.
- 868 27. Srinivasan G, Davis M, LeBoeuf M, Fatemi M, Azher Z, Lu Y, et al. Potential to Enhance Large Scale  
869 Molecular Assessments of Skin Photoaging through Virtual Inference of Spatial Transcriptomics from  
870 Routine Staining [Internet]. *bioRxiv*; 2023 [cited 2023 Oct 7]. p. 2023.07.30.551188. Available from:  
871 <https://www.biorxiv.org/content/10.1101/2023.07.30.551188v1>
- 872 28. Choi YH, Kim JK. Dissecting Cellular Heterogeneity Using Single-Cell RNA Sequencing. *Mol Cells*.  
873 2019;42:189–99.
- 874 29. He B, Bergenstråhle L, Stenbeck L, Abid A, Andersson A, Borg Å, et al. Integrating spatial gene  
875 expression and breast tumour morphology via deep learning. *Nat Biomed Eng*. 2020;4:827–34.
- 876 30. Levy-Jurgenson A, Tekpli X, Kristensen VN, Yakhini Z. Spatial transcriptomics inferred from pathology  
877 whole-slide images links tumor heterogeneity to survival in breast and lung cancer. *Sci Rep*.  
878 2020;10:18802.
- 879 31. Duan H, Cheng T, Cheng H. Spatially resolved transcriptomics: advances and applications. *Blood Sci*.  
880 2022;5:1–14.

- 881 32. Liu J, Tran V, Vemuri VNP, Byrne A, Borja M, Kim YJ, et al. Concordance of MERFISH spatial  
882 transcriptomics with bulk and single-cell RNA sequencing. *Life Sci Alliance*. 2022;6:e202201701.
- 883 33. Wen H, Tang W, Jin W, Ding J, Liu R, Shi F, et al. Single Cells Are Spatial Tokens: Transformers for  
884 Spatial Transcriptomic Data Imputation [Internet]. *arXiv*; 2023 [cited 2023 Oct 7]. Available from:  
885 <http://arxiv.org/abs/2302.03038>
- 886 34. Comiter C, Vaishnav ED, Ciampicotti M, Li B, Yang Y, Rodig SJ, et al. Inference of single cell profiles  
887 from histology stains with the Single-Cell omics from Histology Analysis Framework (SCHAF). *BioRxiv*.  
888 2023;2023-03.
- 889 35. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set  
890 enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.  
891 *Proceedings of the National Academy of Sciences*. 2005;102:15545–50.
- 892 36. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular  
893 signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27:1739–40.
- 894 37. Li L, Guan Y, Chen X, Yang J, Cheng Y. DNA Repair Pathways in Cancer Therapy and Resistance.  
895 *Frontiers in Pharmacology* [Internet]. 2021 [cited 2023 Oct 7];11. Available from:  
896 <https://www.frontiersin.org/articles/10.3389/fphar.2020.629266>
- 897 38. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative  
898 HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;14:128.
- 899 39. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative  
900 HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;14:128.
- 901 40. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a  
902 comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44:W90-  
903 97.
- 904 41. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene Set Knowledge  
905 Discovery with Enrichr. *Current Protocols*. 2021;1:e90.
- 906 42. Pai SG, Carneiro BA, Mota JM, Costa R, Leite CA, Barroso-Sousa R, et al. Wnt/beta-catenin  
907 pathway: modulating anticancer immune response. *Journal of Hematology & Oncology*. 2017;10:101.
- 908 43. Fagotto F, Aslemar A. EpCAM cellular functions in adhesion and migration, and potential impact on  
909 invasion: A critical review. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*. 2020;1874:188436.
- 910 44. Wang X, Barrera C, Bera K, Viswanathan VS, Azarianpour-Esfahani S, Koyuncu C, et al. Spatial  
911 interplay patterns of cancer nuclei and tumor-infiltrating lymphocytes (TILs) predict clinical benefit for  
912 immune checkpoint inhibitors. *Science Advances*. 2022;8:eabn3966.
- 913 45. Dermani FK, Samadi P, Rahmani G, Kohlan AK, Najafi R. PD-1/PD-L1 immune checkpoint: Potential  
914 target for cancer therapy. *Journal of Cellular Physiology*. 2019;234:1313–25.
- 915 46. Yang L, Li A, Lei Q, Zhang Y. Tumor-intrinsic signaling pathways: key roles in the regulation of the  
916 immunosuppressive tumor microenvironment. *Journal of Hematology & Oncology*. 2019;12:125.
- 917 47. Ando M, Ito M, Srirat T, Kondo T, Yoshimura A. Memory T cell, exhaustion, and tumor immunity.  
918 *Immunological Medicine*. 2020;43:1–9.

- 919 48. Crespo J, Sun H, Welling TH, Tian Z, Zou W. T cell anergy, exhaustion, senescence, and stemness in  
920 the tumor microenvironment. *Current opinion in immunology*. 2013;25:214–21.
- 921 49. Peng Z, Ye M, Ding H, Feng Z, Hu K. Spatial transcriptomics atlas reveals the crosstalk between  
922 cancer-associated fibroblasts and tumor microenvironment components in colorectal cancer. *J Transl  
923 Med*. 2022;20:302.
- 924 50. Wang F, Long J, Li L, Wu Z-X, Da T-T, Wang X-Q, et al. Single-cell and spatial transcriptome analysis  
925 reveals the cellular heterogeneity of liver metastatic colorectal cancer. *Sci Adv*. 2023;9:eadf5464.
- 926 51. Liu Z, Zhou Z, Dang Q, Xu H, Lv J, Li H, et al. Immunosuppression in tumor immune  
927 microenvironment and its optimization from CAR-T cell therapy. *Theranostics*. 2022;12:6273.
- 928 52. Levy JJ, Zavras JP, Veziroglu EM, Nasir-Moin M, Kolling FW, Christensen BC, et al. Identification of  
929 Spatial Proteomic Signatures of Colon Tumor Metastasis: A Digital Spatial Profiling Approach. *Am J  
930 Pathol*. 2023;S0002-9440(23)00123-2.
- 931 53. Zhang Q, Jiang S, Schroeder A, Hu J, Li K, Zhang B, et al. Leveraging spatial transcriptomics data to  
932 recover cell locations in single-cell RNA-seq with CeLEry. *Nature Communications*. 2023;14:4050.
- 933 54. Henley R, Rapicavoli N, Janesick A, Shelansky R, Kim A, Hensel J, et al. 95 Characterization of  
934 human breast cancer tissue with the Xenium In Situ platform reveals a novel marker for invasiveness  
935 [Internet]. *BMJ Specialist Journals*; 2022 [cited 2023 Oct 7]. Available from:  
936 [https://jitc.bmj.com/content/10/Suppl\\_2/A104.abstract](https://jitc.bmj.com/content/10/Suppl_2/A104.abstract)
- 937 55. Marco Salas S, Czarnewski P, Kuemmerle LB, Helgadottir S, Mattsson Langseth C, Tiesmeyer S, et  
938 al. Optimizing Xenium In Situ data utility by quality assessment and best practice analysis workflows.  
939 *bioRxiv*. 2023;2023–02.
- 940 56. Singh H, Sukovich D, Mohabbat S, Li D, Kim H, Tentori A. Visium CytAssist: A Novel Platform for  
941 Spatial Transcriptomic Analysis of FFPE Sections Mounted on Standard Glass Slides. *MOLECULAR  
942 THERAPY. CELL PRESS 50 HAMPSHIRE ST, FLOOR 5, CAMBRIDGE, MA 02139 USA*; 2022. p. 589–  
943 90.
- 944 57. Rosasco MG, Ho C-S, Luo T, Stein MM, Lonini L, Stumpe MC, et al. Abstract 4692: Comparison of  
945 interassay similarity and cellular deconvolution in spatial transcriptomics data using Visum CytAssist.  
946 *Cancer Research*. 2023;83:4692.
- 947 58. Sun S, Zhu J, Zhou X. Statistical analysis of spatial expression patterns for spatially resolved  
948 transcriptomic studies. *Nature Methods*. 2020;17:193–200.
- 949 59. Graham S, Jahanifar M, Azam A, Nimir M, Tsang Y-W, Dodd K, et al. Lizard: a large-scale dataset for  
950 colonic nuclear instance segmentation and classification. *Proceedings of the IEEE/CVF International  
951 Conference on Computer Vision [Internet]*. 2021 [cited 2023 Oct 7]. p. 684–93. Available from:  
952 [https://openaccess.thecvf.com/content/ICCV2021W/CDPath/html/Graham\\_Lizard\\_A\\_Large-  
953 Scale\\_Dataset\\_for\\_Colonic\\_Nuclear\\_Instance\\_Segmentation\\_and\\_ICCVW\\_2021\\_paper.html](https://openaccess.thecvf.com/content/ICCV2021W/CDPath/html/Graham_Lizard_A_Large-Scale_Dataset_for_Colonic_Nuclear_Instance_Segmentation_and_ICCVW_2021_paper.html)
- 954 60. Vuola AO, Akram SU, Kannala J. Mask-RCNN and U-Net Ensembled for Nuclei Segmentation. 2019  
955 *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. 2019. p. 208–12.
- 956 61. He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. *Proceedings of the IEEE international conference  
957 on computer vision [Internet]*. 2017 [cited 2023 Oct 7]. p. 2961–9. Available from:  
958 [http://openaccess.thecvf.com/content\\_iccv\\_2017/html/He\\_Mask\\_R-CNN\\_ICCV\\_2017\\_paper.html](http://openaccess.thecvf.com/content_iccv_2017/html/He_Mask_R-CNN_ICCV_2017_paper.html)



- 959 62. Reddy R, Reddy R, Sharma C, Jackson C, Palisoul S, Barney R, et al. Graph Neural Networks  
960 Ameliorate Potential Impacts of Imprecise Large-Scale Autonomous Immunofluorescence Labeling of  
961 Immune Cells on Whole Slide Images. *Proceedings of the First International Workshop on Geometric  
962 Deep Learning in Medical Image Analysis [Internet]*. PMLR; 2022 [cited 2023 Apr 1]. p. 15–33. Available  
963 from: <https://proceedings.mlr.press/v194/reddy22a.html>
- 964 63. Jaume G, Pati P, Anklin V, Foncubierta A, Gabrani M. Histocartography: A toolkit for graph analytics  
965 in digital pathology. *MICCAI Workshop on Computational Pathology*. PMLR; 2021. p. 117–28.
- 966 64. Levy J, Haudenschild C, Barwick C, Christensen B, Vaickus L. Topological Feature Extraction and  
967 Visualization of Whole Slide Images using Graph Neural Networks. *Pac Symp Biocomput*. 2021;26:285–  
968 96.
- 969 65. Adnan M, Kalra S, Tizhoosh HR. Representation Learning of Histopathology Images Using Graph  
970 Neural Networks. 2020 [cited 2020 Jul 23]. p. 988–9. Available from:  
971 [https://openaccess.thecvf.com/content\\_CVPRW\\_2020/html/w57/Adnan\\_Representation\\_Learning\\_of\\_His  
972 topathology\\_Images\\_Using\\_Graph\\_Neural\\_Networks\\_CVPRW\\_2020\\_paper.html](https://openaccess.thecvf.com/content_CVPRW_2020/html/w57/Adnan_Representation_Learning_of_His_topathology_Images_Using_Graph_Neural_Networks_CVPRW_2020_paper.html)
- 973 66. Li MM, Huang K, Zitnik M. Graph representation learning in biomedicine and healthcare. *Nature  
974 Biomedical Engineering*. 2022;1–17.
- 975 67. Zhou Y, Graham S, Alemi Koohbanani N, Shaban M, Heng P-A, Rajpoot N. CGC-Net: Cell Graph  
976 Convolutional Network for Grading of Colorectal Cancer Histology Images. 2019 [cited 2019 Nov 28]. p.  
977 0–0. Available from: [http://openaccess.thecvf.com/content\\_ICCVW\\_2019/html/VRMI/Zhou\\_CGC-  
978 Net\\_Cell\\_Graph\\_Convolutional\\_Network\\_for\\_Grading\\_of\\_Colorectal\\_Cancer\\_ICCVW\\_2019\\_paper.html](http://openaccess.thecvf.com/content_ICCVW_2019/html/VRMI/Zhou_CGC-Net_Cell_Graph_Convolutional_Network_for_Grading_of_Colorectal_Cancer_ICCVW_2019_paper.html)
- 979 68. Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph Attention Networks.  
980 arXiv:1710.10903 [cs, stat] [Internet]. 2018 [cited 2020 Oct 25]; Available from:  
981 <http://arxiv.org/abs/1710.10903>
- 982 69. Raju A, Yao J, Haq MM, Jonnagaddala J, Huang J. Graph Attention Multi-instance Learning for  
983 Accurate Colorectal Cancer Staging. *Medical Image Computing and Computer Assisted Intervention –  
984 MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V  
985 [Internet]*. Berlin, Heidelberg: Springer-Verlag; 2020 [cited 2022 Sep 13]. p. 529–39. Available from:  
986 [https://doi.org/10.1007/978-3-030-59722-1\\_51](https://doi.org/10.1007/978-3-030-59722-1_51)
- 987 70. Qiu J, Chen Q, Dong Y, Zhang J, Yang H, Ding M, et al. Gcc: Graph contrastive coding for graph  
988 neural network pre-training. *Proceedings of the 26th ACM SIGKDD International Conference on  
989 Knowledge Discovery & Data Mining*. 2020. p. 1150–60.
- 990 71. You Y, Chen T, Sui Y, Chen T, Wang Z, Shen Y. Graph Contrastive Learning with Augmentations  
991 [Internet]. arXiv; 2021 [cited 2023 Oct 7]. Available from: <http://arxiv.org/abs/2010.13902>
- 992 72. Zhu Y, Xu Y, Liu Q, Wu S. An empirical study of graph contrastive learning. arXiv preprint  
993 arXiv:2109.01116. 2021;
- 994 73. Biancalani T, Scalia G, Buffoni L, Avasthi R, Lu Z, Sanger A, et al. Deep learning and alignment of  
995 spatially resolved single-cell transcriptomes with Tangram. *Nat Methods*. 2021;18:1352–62.
- 996 74. Flamary R, Courty N, Gramfort A, Alaya MZ, Boisbunon A, Chambon S, et al. Pot: Python optimal  
997 transport. *Journal of Machine Learning Research*. 2021;22:1–8.
- 998 75. Peyré G, Cuturi M. Computational optimal transport: With applications to data science. *Foundations  
999 and Trends® in Machine Learning*. 2019;11:355–607.

- 1000 76. Villani C. Optimal Transport [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009 [cited  
1001 2023 Oct 7]. Available from: <http://link.springer.com/10.1007/978-3-540-71050-9>
- 1002 77. Schiebinger G, Shu J, Tabaka M, Cleary B, Subramanian V, Solomon A, et al. Optimal-Transport  
1003 Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell*.  
1004 2019;176:928-943.e22.
- 1005 78. Fey M, Lenssen JE. Fast Graph Representation Learning with PyTorch Geometric. arXiv:190302428  
1006 [cs, stat] [Internet]. 2019 [cited 2020 Jul 23]; Available from: <http://arxiv.org/abs/1903.02428>
- 1007 79. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization [Internet]. arXiv; 2017 [cited 2023 Oct  
1008 7]. Available from: <http://arxiv.org/abs/1412.6980>
- 1009 80. Ying R, Bourgeois D, You J, Zitnik M, Leskovec J. GNNExplainer: Generating Explanations for Graph  
1010 Neural Networks. arXiv:190303894 [cs, stat] [Internet]. 2019 [cited 2020 Oct 25]; Available from:  
1011 <http://arxiv.org/abs/1903.03894>
- 1012 81. McInnes L, Healy J, Saul N, Großberger L. UMAP: Uniform Manifold Approximation and Projection.  
1013 *Journal of Open Source Software*. 2018;3:861.
- 1014 82. McInnes L, Healy J, Astels S. hdbscan: Hierarchical density based clustering. *Journal of Open Source  
1015 Software*. 2017;2:205.
- 1016
- 1017

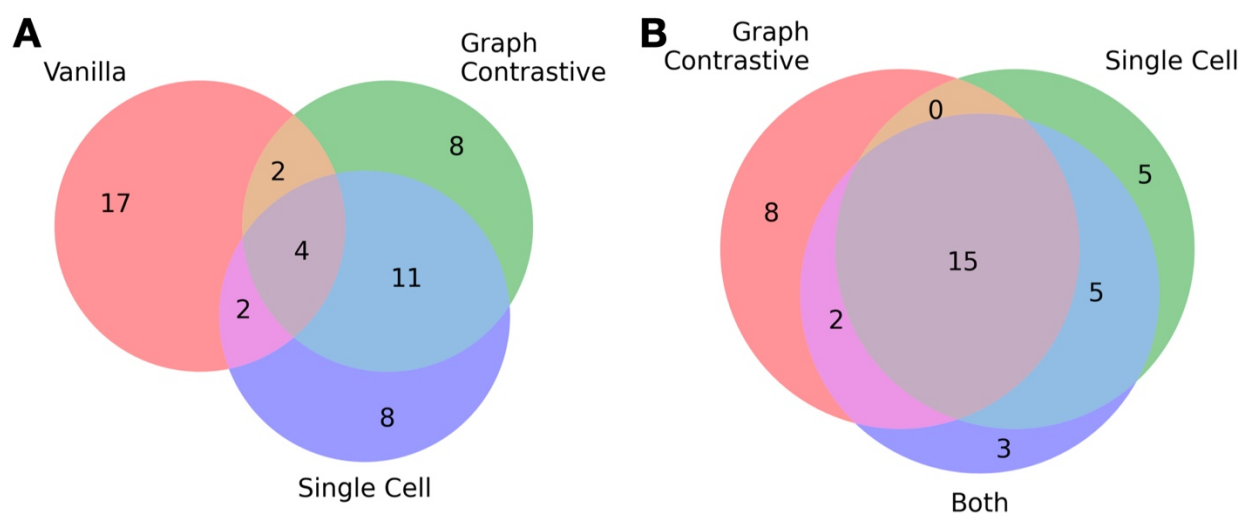


1018 **Supplementary Material**

1019 **Supplementary Table 1: Spearman correlation scores for top performing genes from each**  
 1020 **CGNN method**

	Vanilla		Graph Contrastive Pretraining		Single-Cell Regularization		Both	
	Gene	R	Gene	R	Gene	R	Gene	R
1	LGALS3BP	0.484	S100A6	0.575	FXYD3	0.624	TFF3	0.643
2	KRT8	0.443	FXYD3	0.568	LGALS4	0.590	KRT8	0.638
3	SPINT2	0.425	TMSB10	0.566	KRT8	0.590	S100A6	0.634
4	MYL6	0.412	TMSB4X	0.561	LGALS3	0.574	FXYD3	0.625
5	KLF5	0.404	TFF3	0.551	MYH14	0.567	PHGR1	0.602
6	S100A6	0.403	MT-ND4L	0.550	CLDN4	0.560	PIGR	0.595
7	JUN	0.392	PIGR	0.548	MISP	0.556	LGALS3	0.592
8	ETS2	0.390	KRT8	0.546	S100A6	0.541	ELF3	0.592
9	CSRP1	0.390	LGALS3	0.543	PHGR1	0.540	CLDN7	0.590
10	ACTB	0.389	MT-CO3	0.537	TFF3	0.535	KLF5	0.589
11	PHGR1	0.385	LGALS4	0.535	CLDN7	0.534	CD24	0.586
12	S100A10	0.384	ITM2C	0.534	CD24	0.530	EPCAM	0.585
13	ZFP36	0.377	MT-CO2	0.530	CD9	0.521	LGALS4	0.583
14	PERP	0.377	PHGR1	0.526	EPCAM	0.519	CLDN4	0.583
15	IGFBP4	0.373	MT-ND4	0.510	TMSB10	0.505	FCGBP	0.580
16	COL1A1	0.370	EPCAM	0.501	PIGR	0.502	MYH14	0.580
17	FLNB	0.367	CLDN4	0.501	SPINT2	0.501	TMSB10	0.579
18	SDC4	0.363	MT-ND6	0.494	REG4	0.499	TMSB4X	0.567
19	CLDN7	0.354	STARD10	0.490	KLF5	0.496	MISP	0.565
20	LMNA	0.348	LGALS3BP	0.490	ANXA2	0.491	STARD10	0.564
21	CDC42EP5	0.346	CLDN7	0.489	PPP1R1B	0.481	CD9	0.559
22	CANT1	0.344	MISP	0.479	IFI27	0.479	LGALS3BP	0.554
23	PYGB	0.337	CD24	0.477	STARD10	0.479	SPINT2	0.551
24	MYC	0.334	MT-CYB	0.467	MLEC	0.479	MUC13	0.549
25	GMDS	0.327	ACTB	0.457	ELF3	0.475	SDC4	0.542

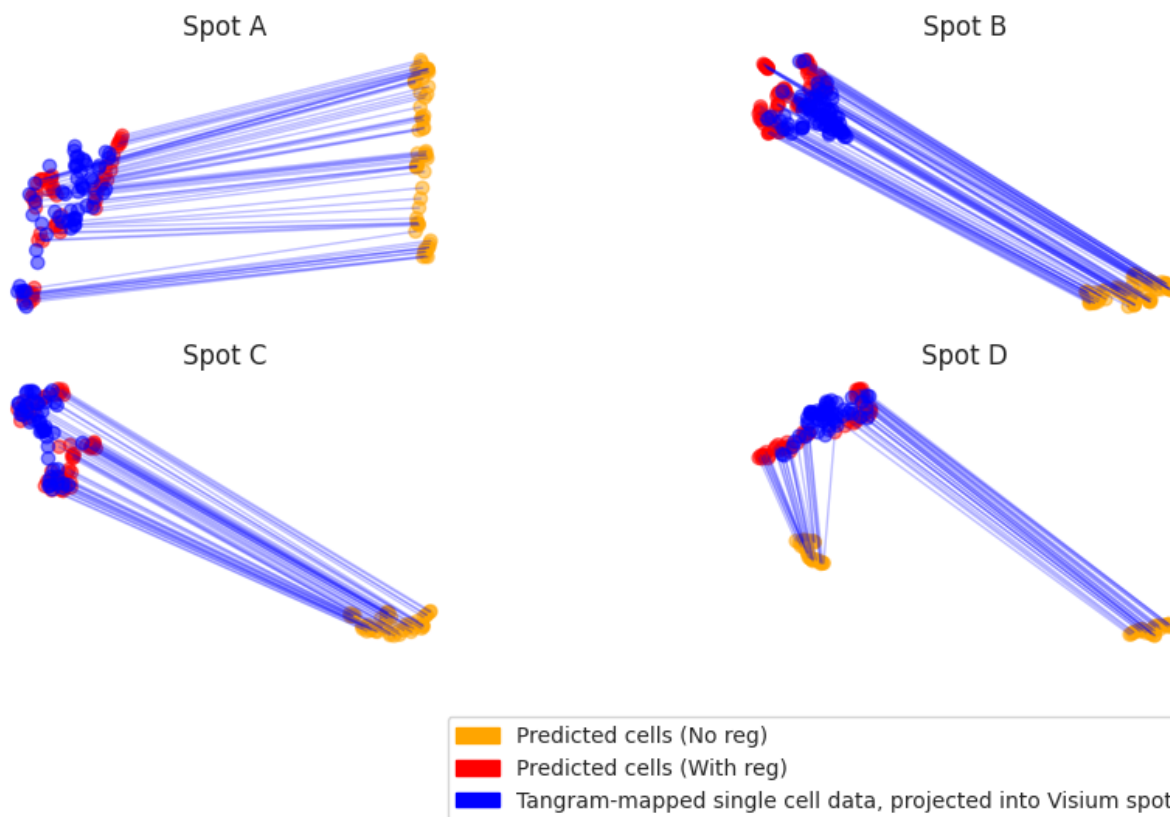
1021



1022

1023 **Supplementary Figure 1: Comparative analysis of top performing genes used in different**  
 1024 **penalization and pretraining methods for cell graph neural networks predicting spot-level gene**

1025 **expression. A)** Left Venn Diagram: Shows the overlaps between genes selected using the "Vanilla",  
1026 "Graph Contrastive", and "Single Cell" penalization methods. Each circle represents a distinct set of  
1027 genes, with overlapping regions indicating shared genes between methods. **B)** Right Venn Diagram:  
1028 Represents the overlaps between genes selected using the "Graph Contrastive", "Single Cell", and "Both"  
1029 (a combination of methods) penalization/pretraining approaches. Overlaps indicate genes that are  
1030 common across multiple methods, highlighting the consensus and divergence in gene selection across  
1031 these strategies. Numbers within each segment of the diagrams indicate the count of genes unique to or  
1032 shared by the corresponding methods.



1033

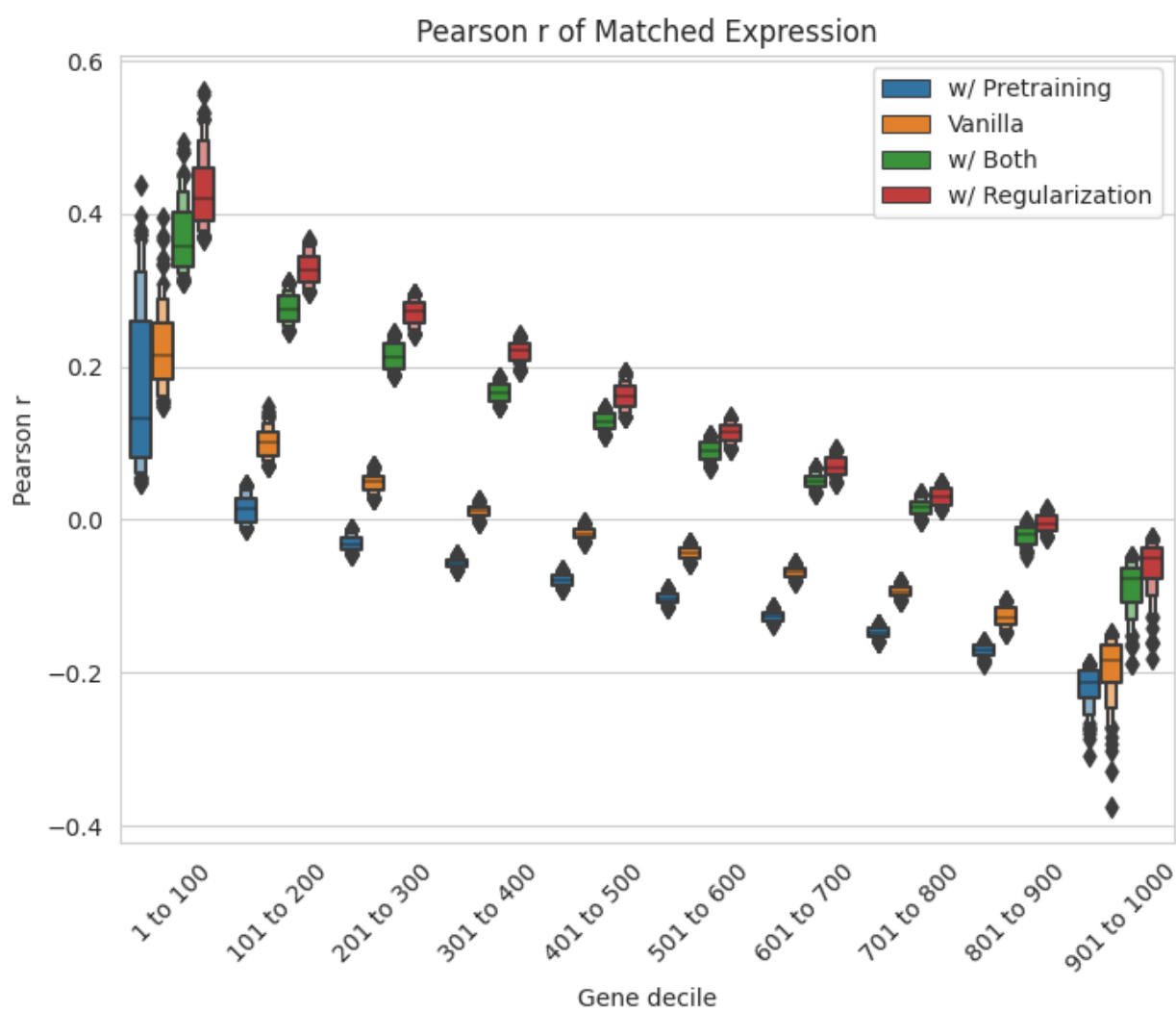
1034 **Supplementary Figure 2: UMAP Embedding of Predicted vs. Assigned Cell Expressions.** This figure  
1035 showcases UMAP embeddings contrasting predicted cell expressions against expressions of cells  
1036 designated to the Visium spot via the Tangram optimal matching algorithm. The red embeddings  
1037 showcase model outputs with regularization, blue embeddings correspond to the original ground truth  
1038 from single cell RNA-Seq, and yellow embeddings represent predicted single cell outputs without  
1039 regularization for the same spot. Evidently, cell regularization enhances the overlap of predicted and  
1040 assigned gene expressions. Performance assessment relies on the Earth Mover's Distance (EMD)  
1041 between a held-out single-cell profile designated via Tangram and predicted expressions on an isolated  
1042 slide. Distances between cells employ the Euclidean metric, while optimal matchings leverage the  
1043 Hungarian algorithm via the "scipy.optimize.linear\_sum\_assignment" function. Notably, post-  
1044 regularization, the EMD decreases from  $0.2113 \pm 0.0018$  to  $0.1473 \pm 0.0018$ .  
1045

1046

1047

1048

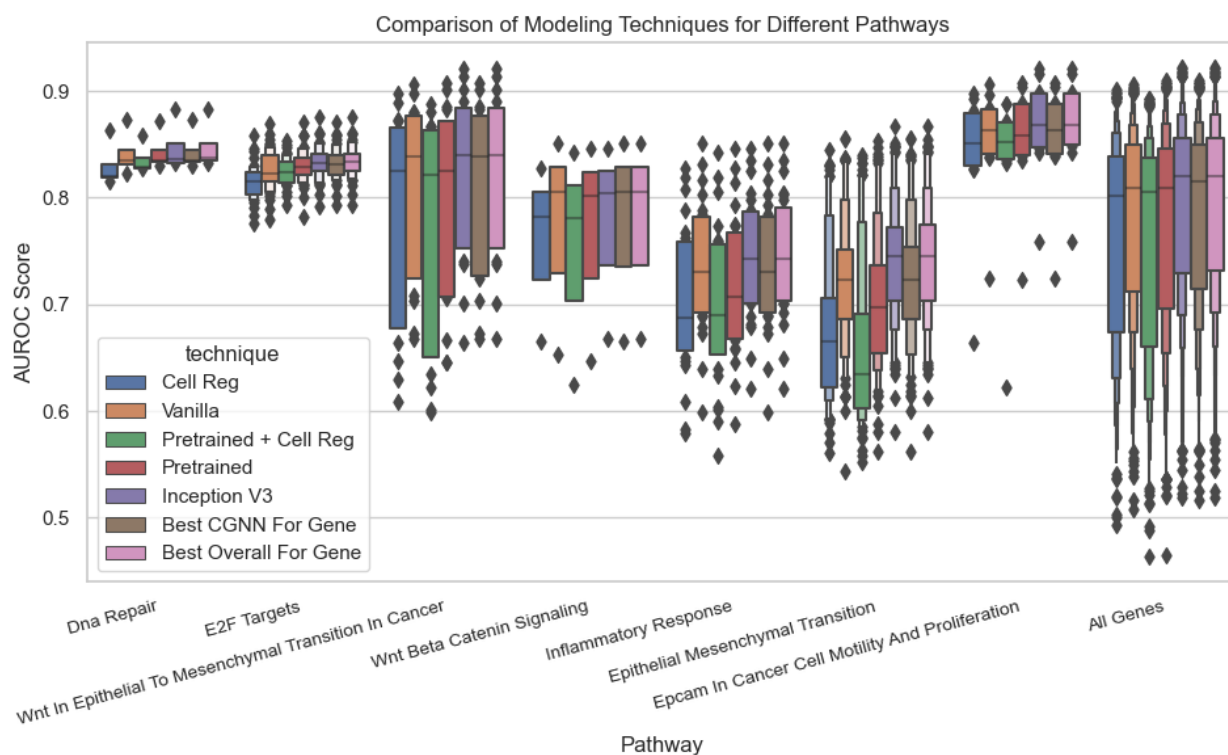
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056



1057  
1058  
1059  
1060

**Supplementary Figure 3: Correlation of predicted versus true single cell log-expression for different deciles of genes, ranked based on performance at the Visium spot level**

1061



1062

1063 **Supplementary Figure 4: Boxplot of AUROC scores for selected pathways for all modeling**  
 1064 **approaches.**

1065

1066 **Supplementary Table 2: Top 10 pathways ranked by adjusted p-value, sourced from the Elsevier**  
 1067 **Pathway Collection via Enricher, for each cell-level modeling strategy.**

	Rank	Term	Overlap	Adjusted P-value
<b>No regularization</b>	1	Desmosomes Role in Dilated Cardiomyopathy	5/15	1.72e-06
	2	Desmosome Assembly	5/18	2.19e-06
	3	Corneodesmosomes in Atopic Dermatitis	5/19	2.19e-06
	4	Desmosome Dysfunction in Cardiomyocyte	4/12	1.49e-05
	5	EPCAM in Cancer Cell Motility and Proliferation	5/36	3.61e-05
	6	Proteins with Altered Expression in Cancer Metastases	7/106	3.61e-05
	7	WNT in Epithelial to Mesenchymal Transition in Cancer	5/43	7.05e-05
	8	Proteins with Altered Expression in Endometrial Cancer	4/23	1.28e-04

	9	Epithelial to Mesenchymal Transition in Cancer: Overview	6/90	1.40e-04
	10	Clear Cell Endometrial Cancer and Papillary Serous Endometrial Cancer	5/65	3.88e-04
<b>Contrastive Pretraining</b>	1	Desmosome Dysfunction in Cardiomyocyte	4/12	4.26e-05
	2	Desmosomes Role in Dilated Cardiomyopathy	4/15	5.17e-05
	3	Proteins with Altered Expression in Cancer Metastases	7/106	5.17e-05
	4	Desmosome Assembly	4/18	6.44e-05
	5	Corneodesmosomes in Atopic Dermatitis	4/19	6.50e-05
	6	Epithelial to Mesenchymal Transition in Cancer: Overview	6/90	1.50e-04
	7	EPCAM in Cancer Cell Motility and Proliferation	4/36	6.61e-04
	8	WNT in Epithelial to Mesenchymal Transition in Cancer	4/43	0.0012
	9	Proteins Involved in HPV Infection	3/19	0.0019
	10	Proteins Involved in Pathogenesis of Arrhythmogenic Right Ventricular Cardiomyopathy	3/21	0.0023
<b>Single Cell Regularization</b>	1	Desmosomes Role in Dilated Cardiomyopathy	5/15	1.61e-06
	2	Desmosome Assembly	5/18	2.04e-06
	3	Corneodesmosomes in Atopic Dermatitis	5/19	2.04e-06
	4	Proteins with Altered Expression in Cancer Metastases	8/106	2.92e-06
	5	Desmosome Dysfunction in Cardiomyocyte	4/12	1.10e-05
	6	Epithelial to Mesenchymal Transition in Cancer: Overview	7/90	1.10e-05
	7	EPCAM in Cancer Cell Motility and Proliferation	5/36	2.65e-05
	8	WNT in Epithelial to Mesenchymal Transition in Cancer	5/43	5.76e-05
	9	Proteins with Altered Expression in Endometrial Cancer	4/23	1.06e-04
	10	Alkalosis	3/9	1.96e-04
<b>Both regularization methods</b>	1	Desmosomes Role in Dilated Cardiomyopathy	5/15	1.52e-06
	2	Desmosome Assembly	5/18	1.93e-06
	3	Corneodesmosomes in Atopic Dermatitis	5/19	1.93e-06
	4	Desmosome Dysfunction in Cardiomyocyte	4/12	1.31e-05

---

5	EPCAM in Cancer Cell Motility and Proliferation	5/36	3.18e-05
6	Proteins with Altered Expression in Cancer Metastases	7/106	3.18e-05
7	WNT in Epithelial to Mesenchymal Transition in Cancer	5/43	6.21e-05
8	Proteins with Altered Expression in Endometrial Cancer	4/23	1.13e-04
9	Epithelial to Mesenchymal Transition in Cancer: Overview	6/90	1.23e-04
10	Clear Cell Endometrial Cancer and Papillary Serous Endometrial Cancer	5/65	3.42e-04

---

1068

1069

1070

1071

1072

1073

1074