

细胞交响曲：利用单细胞病理组学解读细胞的分子和表型协同作用

Paul H. Acosta^{1,2,*} Pingjun Chen^{1,2} Simon P. Castillo^{1,2}
Maria Esther Salvatierra¹ Yinyin Yuan^{1,2,*} Xiaoxi Pan^{1,2,*}

¹Translational Molecular Pathology Department, Division of Pathology and Laboratory Medicine,
The University of Texas MD Anderson Cancer Center, Houston, TX, United States

²Institute for Data Science in Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

*Corresponding authors: phacosta@mdanderson.org, yyuan6@mdanderson.org, xpan7@mdanderson.org

Abstract

Xenium, 一个新的空间转录组学平台, 可以对复杂肿瘤组织进行亚细胞分辨率的分析。尽管组织学图像中包含丰富的形态信息, 但提取稳健的细胞水平特征并将其与空间转录组数据整合仍然是一个关键挑战。我们介绍了 CellSymphony, 这是一个灵活的多模态框架, 利用了从 Xenium 转录组特征和组织学图像中派生的基础模型嵌入, 在真正的单细胞分辨率上运作。通过学习将空间基因表达与形态背景融合的联合表征, CellSymphony 实现了精确的细胞类型注释, 并揭示了三种癌症类型中的不同微环境区域。该工作强调了基础模型和多模态融合在解读复杂组织生态系统中细胞的生理和表型协作方面的潜力。

1. 介绍

肿瘤微环境是一个复杂的生态系统, 其中不同的细胞在空间和分子层面进行互动, 影响癌症的进展 [2, 7]。最近的空间转录组学进展使得高分辨率的基因表达分析成为可能, 揭示了肿瘤的空间组织和分子异质性 [13, 15]。在这些技术中, Xenium 平台允许多达 5000 个基因的亚细胞转录组分析, 开辟了研究细胞表型、它们的交流以及新兴组织结构的新途径 [11]。

虽然组织学图像提供了关于细胞和组织结构的丰富背景和形态信息, 但它们缺乏分子分辨率。相反, Xenium 提供了分子见解, 但在异质环境中准确的细胞分类时常面临困难, 特别是在细胞边界模糊或表达谱稀疏时 [17]。因此, 利用这两种模式得到的结果对在原位全面理解细胞表型显得尤为重要。目前的方法通常依赖于基因组改变的间接推断 [19] 或基于标记基因 [12] 的识别, 这可能在复杂的肿瘤微环境中不可靠。这突显了整合正交模态、组织学和空间转录组学以改善肿瘤细胞注释的必要性。最近的一些研究尝试整合组织学图像和空间转录组学数据, 特别是来自 Visium 平台的数据, 以加强下游分析, 如从形态学推测基因 [8, 21, 22, 24]、细胞表型分析 [12, 25] 和组织特征化

[4, 9, 10]。然而, 这些努力主要集中在斑点-补丁或斑点-细胞水平上, 并没有充分利用丰富的细胞级分辨率。因此, 细胞级的组织学与转录组学整合仍然很大程度上未被充分探索。

基础模型在形态学 [5, 23] 和分子分析 [6, 20] 方面都有了进展。通过对多种癌症类型的广泛数据进行训练, 这些模型学习了一般化的表示形式, 从而提升了下游任务的效果。基于病理学和单细胞基础模型 UNI2 [5] 和 scGPT [6], 我们引入了 CellSymphony (图 1), 这是一个多模态框架, 能够在单细胞分辨率下整合形态学和转录组信息, 以增强空间组织的特征描述, 并获得对高度协同的肿瘤生态系统的生物学洞察力。我们的实验结果表明, 整合基因和形态嵌入能够提高细胞类型注释的准确性, 并加强空间组织亚型和微环境生态位的发现。

2. 方法

CellSymphony 框架专为 Xenium Prime 5k 空间转录组学数据与 H & E 图像配对而设计, 首先使用基础模型从每种模式中提取深度特征嵌入。这些嵌入随后驱动两项核心能力: (1) 基于多模态变压器的细胞类型分类, 以及 (2) 对比学习以对齐形态学和转录组学。

我们为每个细胞推导了配对的、模态特定的嵌入 (图 1a)。在转录组学层面, 我们将细胞的基因计数向量输入到预训练的 scGPT 模型中, 产生固定的转录组嵌入 (图 1a, 上)。在形态学层面, 以该细胞核为中心的 224 x 224 H & E 图块 (0.5µm/px) 通过 UNI2, 生成形态学嵌入。我们未使用模型默认的图瓦特征表示, 而是从 UNI2 中提取中间空间标记, 并聚合最接近细胞核的标记, 以生成 1536 维的细胞特异性形态嵌入, 从而允许灵活的空间背景 (图 1a, 下)。这些匹配的形态和 RNA 向量共同构成我们跨模态分析的联合输入。

我们训练了基于 Transformer 的模型, 以对来自 Xenium Prime 数据的基因表达和形态嵌入的细胞类型进行分类 (图 1b, 上)。细胞注释的来源有: (1) SingleR [3], 其使用基于参考的转录组分类, 或 (2) AI-sTIL [1, 18], 一个基于 H & E 的分割和分类流程。模型依

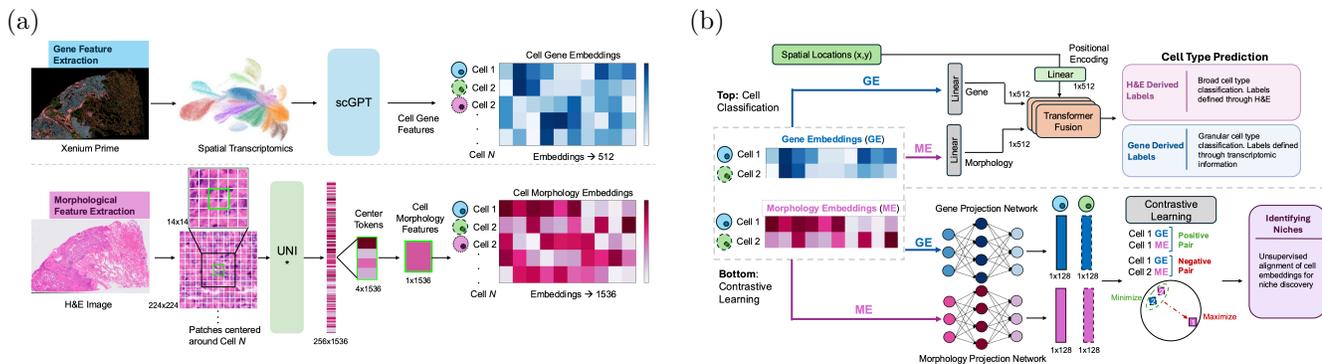


Figure 1. CellSymphony 框架概述。(a) 细胞嵌入是通过使用 scGPT 从 Xenium 转录组中得出的 (顶部), 并通过修改的 UNI* 模型从组织学图像中得出 (底部)。(b) 两种深度学习方法的结构: 基于 transformer 的细胞类型注释 (顶部) 和用于跨模态对齐转录组和形态特征的对比如学习 (底部)。

据标签来源分别进行训练, 排除了模糊或置信度低的注释, 并将数据分为 80 % 训练和 20 % 测试。

我们评估了四种 Transformer 架构: (1) 仅在基因嵌入 (GEB) 上训练的单模态 Transformer; (2) 通过使用正弦编码为空间坐标添加第二个输入 token 的空间 Transformer; (3) 处理基因和形态嵌入 (MEB) 作为独立输入 token 的双模态 Transformer, 每个 token 都标记有特定模态的标签; (4) 使用相对位置注意力来指导 token 之间相互作用的多输入 Transformer, 将所有三个输入 (基因、形态、空间) 整合在一起。每个嵌入被投影到一个共享的维度空间, 然后通过一个 6 层 Transformer 融合编码器。输出经过均值池化后使用线性层进行分类。所有模型使用 AdamW 训练 20 个周期 (学习率 3×10^{-5} , 批尺寸 64, 权重衰减 0.01), 采用交叉熵损失和类别权重来处理标签不平衡问题。

为了在共享空间中对齐转录组和形态嵌入, 我们实现了一种无监督的对比如学习框架 (图 1b, 底部)。如前所述, 使用 UNI2 和 scGPT 分别从 H & E 和 Xenium 数据中提取细胞特异性特征。每种模态都通过一个单独的投影网络映射到一个 128 维的潜在空间。

网络使用 InfoNCE 损失 [16] 进行了训练, 该损失将匹配的嵌入 (相同的细胞) 拉近, 同时将不匹配的对分开。损失是在每批细胞上双向计算的, 从形态学到转录组学, 反之亦然。后续分析集中在训练后的个体投影上。

我们使用 UMAP [14]、聚类和空间可视化来评估所得特征, 以评估在生物信号和生态位检测中的改进, 特别是在形态学衍生表示中的改进。

我们使用了来自 10x Genomics 的肺 ($n_{\text{cells}} = 244,659$)、乳腺 ($n_{\text{cells}} = 461,094$) 和前列腺 ($n_{\text{cells}} = 151,665$) Xenium Prime 5k 数据集。每个样本包括源自人类全组织 & 通路面板的基因表达、空间细胞坐标和匹配的高分辨率 H & E 图像。

Tissue	P_{fib}	R_{fib}	$F1_{\text{fib}}$	P_{lym}	R_{lym}	$F1_{\text{lym}}$	P_{tum}	R_{tum}	$F1_{\text{tum}}$
Unimodal Transformer (GEB Trained)									
Breast	0.71	0.65	0.68	0.55	0.65	0.60	0.85	0.85	0.85
Lung	0.80	0.79	0.79	0.56	0.65	0.60	0.83	0.80	0.82
Prostate	0.88	0.83	0.86	0.28	0.09	0.14	0.78	0.85	0.82
Spatial Transformer (GEB Trained)									
Breast	0.70	0.70	0.70	0.58	0.59	0.59	0.85	0.84	0.85
Lung	0.79	0.82	0.81	0.56	0.63	0.59	0.87	0.79	0.82
Prostate	0.88	0.85	0.86	0.31	0.09	0.14	0.80	0.85	0.82
Dual-Modality Transformer (GEB + MEB Trained)									
Breast	0.90	0.90	0.90	0.90	0.87	0.88	0.92	0.95	0.93
Lung	0.90	0.93	0.91	0.94	0.81	0.87	0.91	0.91	0.91
Prostate	0.93	0.89	0.91	0.71	0.66	0.68	0.86	0.92	0.89
Multi-Input Transformer (GEB + MEB Trained)									
Breast	0.93	0.85	0.89	0.84	0.93	0.88	0.92	0.95	0.93
Lung	0.93	0.87	0.90	0.77	0.96	0.85	0.90	0.92	0.91
Prostate	0.93	0.87	0.90	0.30	0.85	0.44	0.85	0.91	0.88

Table 1. Performance metrics for AI-sTIL predictions across tissues. Abbreviations: P = Precision, R = Recall, F1 = F1-score; fib = Fibroblast, lym = Lymphocyte, tum = Tumor.

2.1. 细胞类型分类的性能

我们在三种不同的组织类型中评估了分类模型: 乳腺癌, 其中侵袭性和原位癌的大簇与正常导管交错分布; 肺癌, 其中恶性上皮细胞与基质、免疫细胞和管状结构 (例如支气管) 混合; 以及前列腺癌, 以低免疫浸润和形态上相似的正常和癌性腺体为特征。

针对基于形态学的基准真值 (AI-sTIL 标签; 表格 1), 空间编码相比单模态 Transformer 仅提供了微小的改进。相反, 结合了基因表达和形态学的双模态 Transformer 表现最为优异。通过多输入 Transformer 加入空间输入并没有带来进一步的提升, 这表明一旦稳健的基因和形态学嵌入确立后, 空间信息对单细胞标注提供的额外收益有限。

使用转录组学的真实标记 (SingleR; 表格 2), 双模态 Transformer 在所有细胞类型的分类上都持续改进, 达到了肺部 B 细胞 (0.91)、肺部 T 细胞 (0.94) 和前列腺成纤维细胞 (0.86) 的最高 F1 得分。结合空间上下文 (多输入 Transformer), 在肺部的表现进一步提升——尤其是上皮细胞 ($F1 = 0.99$)、B 细胞 ($F1 =$

Tissue	P_{b_c}	R_{b_c}	$F1_{b_c}$	P_{end}	R_{end}	$F1_{end}$	P_{epi}	R_{epi}	$F1_{epi}$	P_{fib}	R_{fib}	$F1_{fib}$	P_{mac}	R_{mac}	$F1_{mac}$	P_{t_c}	R_{t_c}	$F1_{t_c}$	
Unimodal Transformer (GEB Trained)																			
Breast	0.77	0.79	0.78	0.80	0.76	0.78	0.94	0.97	0.95	0.85	0.80	0.82	0.82	0.83	0.82	0.89	0.87	0.88	
Lung	0.94	0.87	0.90	0.94	0.93	0.94	0.97	0.97	0.97	0.93	0.92	0.92	0.88	0.95	0.92	0.89	0.94	0.93	
Prostate	0.54	0.55	0.55	0.84	0.76	0.80	0.92	0.97	0.94	0.89	0.85	0.87	0.89	0.80	0.84	0.84	0.77	0.81	
Spatial Transformer (GEB Trained)																			
Breast	0.81	0.75	0.78	0.81	0.76	0.78	0.94	0.97	0.95	0.81	0.82	0.82	0.82	0.82	0.82	0.88	0.89	0.88	
Lung	0.88	0.92	0.90	0.94	0.93	0.93	0.97	0.98	0.97	0.92	0.93	0.93	0.93	0.93	0.93	0.94	0.92	0.93	
Prostate	0.76	0.49	0.59	0.88	0.75	0.81	0.93	0.96	0.95	0.88	0.87	0.88	0.80	0.85	0.82	0.84	0.79	0.81	
Dual-Modality Transformer (GEB + MEB Trained)																			
Breast	0.78	0.83	0.81	0.94	0.68	0.79	0.94	0.98	0.96	0.84	0.82	0.83	0.85	0.82	0.83	0.90	0.89	0.89	
Lung	0.91	0.92	0.91	0.95	0.92	0.94	0.97	0.98	0.97	0.94	0.91	0.93	0.91	0.94	0.93	0.93	0.95	0.94	
Prostate	0.78	0.49	0.60	0.94	0.74	0.83	0.92	0.98	0.95	0.91	0.88	0.89	0.90	0.82	0.86	0.88	0.76	0.82	
Multi-Input Transformer (GEB + MEB Trained)																			
Breast	0.78	0.82	0.80	0.69	0.82	0.75	0.97	0.95	0.96	0.79	0.84	0.82	0.79	0.86	0.82	0.93	0.84	0.88	
Lung	0.95	0.97	0.96	0.98	0.98	0.98	1.00	0.99	0.99	0.92	0.95	0.93	0.96	0.98	0.97	0.98	0.97	0.97	
Prostate	0.21	0.74	0.32	0.82	0.77	0.79	0.95	0.93	0.94	0.94	0.83	0.88	0.74	0.85	0.79	0.81	0.76	0.78	

Table 2. SingleR 预测在不同组织中的性能指标。缩写: P = 精确度, R = 召回率, F1 = F1 得分; 下标: b_c = B 细胞, end = 内皮细胞, epi = 上皮细胞, fib = 成纤维细胞, mac = 巨噬细胞, t_c = T 细胞。GEB: 基因嵌入, MEB: 形态嵌入。

0.96)、T 细胞 (F1 = 0.97) 和巨噬细胞 (F1 = 0.97), 这突显了空间组织在该组织中的重要性。相比之下, 在乳腺和前列腺中, 空间输入的影响有限, 性能提升微乎其微或不一致。值得注意的是, 前列腺中 B 细胞的分类显著下降 (n = 313), 这可能反映了类别不平衡或对空间噪声的敏感性。

2.2. 通过对比学习提升微环境特定区域的发现

虽然有监督分类达到了高准确率, 但其对预定义标签的依赖限制了对细微生物学结构的发现。为了解决这个问题, 我们应用了自监督对比学习框架, 将基因表达和形态对齐到一个共享空间, 并评估了由此产生的形态推导投影。

在不同的组织中, 对比训练的形态投影揭示了细化的空间和表型结构。在肺癌中, 富含淋巴细胞的区域变得更加清晰, 特别是在群集 2 中, 空间一致性提高, 并且 B 细胞和 T 细胞的富集更加明显 (图 2a)。在乳腺癌中, 投影更好地界定了巨噬细胞和成纤维细胞的区域, 并在多个群集中进一步分层上皮亚型 (图 2b)。在前列腺癌中, 投影捕捉了上皮区域之间的分化梯度, 并提供了腺体结构的更清晰分离, 这与已知的组织学变异相一致 (图 2c)。

总的来说, 这些发现表明对比学习提高了形态学衍生特征解析复杂组织结构的能力, 使得无需在推断时借助基因表达, 即可无监督识别富含免疫的区域和肿瘤异质性。

3. 讨论和结论

在这项研究中, 我们引入了 CellSymphony, 这是一种在单细胞水平上整合空间转录组学和组织形态学的多模态框架。应用于来自肺、乳腺和前列腺癌组织的 Xenium Prime 5k 数据, CellSymphony 利用基于基础模型 (scGPT 和 UNI2) 的多模态融合, 支持监督分类和对比学习。

CellSymphony 的核心优势在于其在真正的单细胞分辨率下操作的能力, 同时能够灵活地融合和对齐

态和基因表达。该框架在注释来源 (基于分子的 SingleR 和基于形态的 AI-sTIL) 中实现了高分类性能, 而对比学习提升了单靠形态捕捉到的生物结构, 特别是在富含淋巴细胞和肿瘤上皮区域。

CellSymphony 为未来的多模态空间分析提供了一个灵活的基础。其模块化设计使得能够整合额外的数据类型, 如蛋白成像或双重 RNA 检测, 并且可以适应融入空间图或生物先验信息, 如基因本体论和调控网络。通过在单细胞分辨率上对齐多样的模态, CellSymphony 为研究肿瘤微环境和空间细胞组织开启了新的方向。

References

- [1] Khalid AbdulJabbar, Shan E Ahmed Raza, Rachel Rosenthal, Mariam Jamal-Hanjani, Selvaraju Veeriah, Ayse Akarca, Tom Lund, David A Moore, Roberto Salgado, Maise Al Bakir, et al. Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. *Nature medicine*, 26(7):1054–1062, 2020.
- [2] Nicole M Anderson and M Celeste Simon. The tumor microenvironment. *Current biology*, 30(16):R921–R925, 2020.
- [3] Dvir Aran, Agnieszka P Looney, Leqian Liu, Esther Wu, Valerie Fong, Austin Hsu, Suzanna Chak, Ram P Naikawadi, Paul J Wolters, Adam R Abate, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature immunology*, 20(2):163–172, 2019.
- [4] Feng Bao, Yue Deng, Sen Wan, Susan Q Shen, Bo Wang, Qionghai Dai, Steven J Altschuler, and Lani F Wu. Integrative spatial analysis of cell morphologies and transcriptional states with muse. *Nature biotechnology*, 40(8):1200–1209, 2022.
- [5] Richard J Chen, Tong Ding, Ming Y Lu, Drew FK Williamson, Guillaume Jaume, Andrew H Song, Bowen Chen, Andrew Zhang, Daniel Shao, Muhammad Shaban, et al. Towards a general-purpose foun-

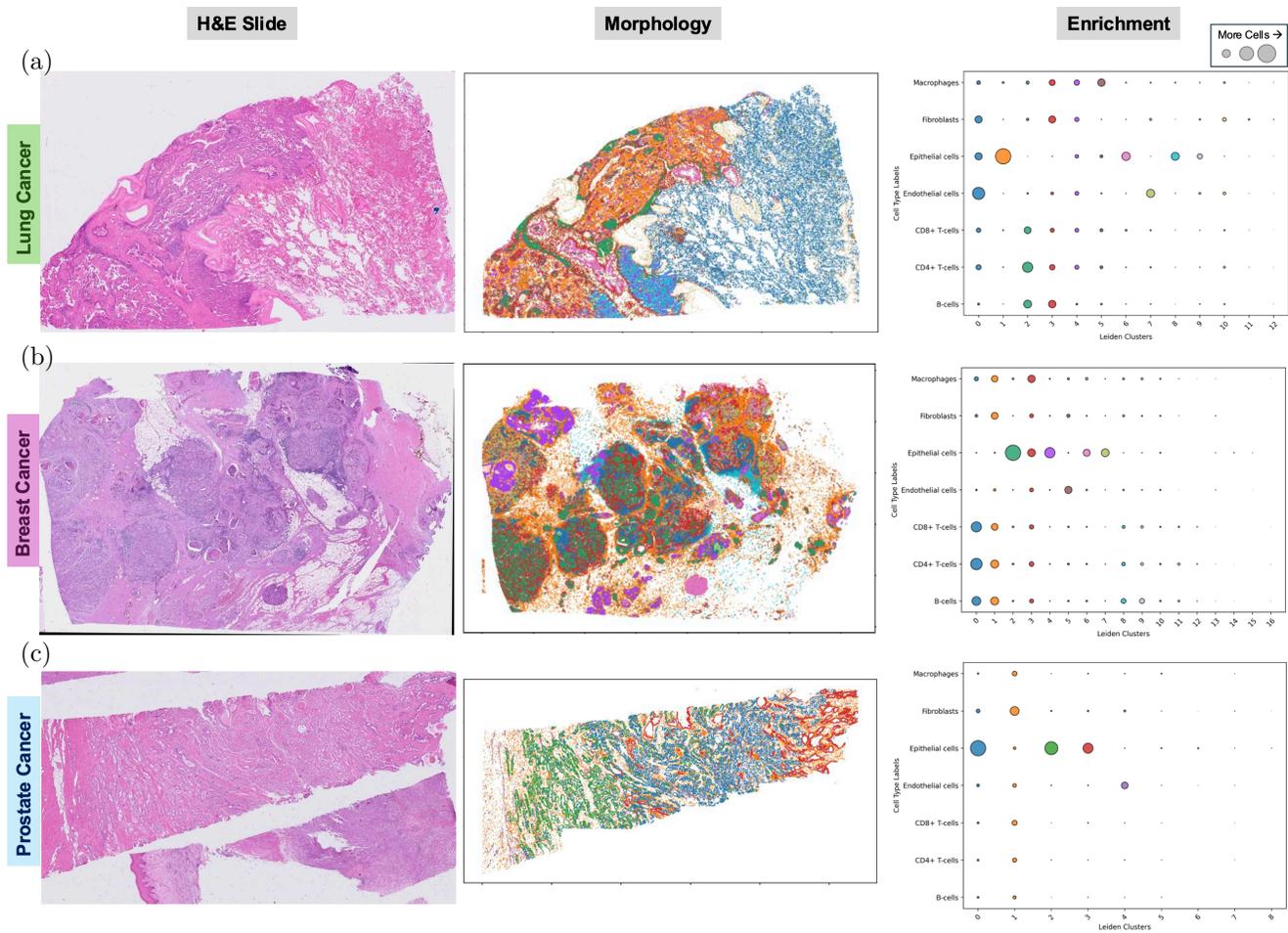


Figure 2. 形态投影聚类在对比训练后的肺 (a), 乳腺 (b) 和前列腺 (c) 样本。每个面板显示 (从左到右): H & E 图像、每个聚类定义的形态模式, 以及这些聚类内细胞类别的富集。

- dation model for computational pathology. *Nature Medicine*, 30(3):850–862, 2024.
- [6] Haotian Cui, Chloe Wang, Hassaan Maan, Kuan Pang, Fengning Luo, Nan Duan, and Bo Wang. scgpt: toward building a foundation model for single-cell multi-omics using generative ai. *Nature Methods*, 21(8):1470–1480, 2024.
- [7] Karin E De Visser and Johanna A Joyce. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer cell*, 41(3):374–403, 2023.
- [8] Bryan He, Ludvig Bergenstråhle, Linnea Stenbeck, Abubakar Abid, Alma Andersson, Åke Borg, Jonas Maaskola, Joakim Lundeberg, and James Zou. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nature biomedical engineering*, 4(8):827–834, 2020.
- [9] Siyu He, Yinuo Jin, Achille Nazaret, Lingting Shi, Xueer Chen, Sham Rampersaud, Bahawar S Dhillon, Izabella Valdez, Lauren E Friend, Joy Linyue Fan, et al. Starfish integrates spatial transcriptomic and histologic data to reveal heterogeneous tumor-immune hubs. *Nature Biotechnology*, 43(2):223–235, 2025.
- [10] Jian Hu, Kyle Coleman, Daiwei Zhang, Edward B Lee, Humam Kadara, Linghua Wang, and Mingyao Li. Deciphering tumor ecosystems at super resolution from spatial transcriptomics with tesla. *Cell systems*, 14(5):404–417, 2023.
- [11] Amanda Janesick, Robert Shelansky, Andrew D Gottscho, Florian Wagner, Stephen R Williams, Morgane Rouault, Ghezal Beliakoff, Carolyn A Morrison, Michelli F Oliveira, Jordan T Sichertman, et al. High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nature communications*, 14(1):8353, 2023.
- [12] Jiahui Jiang, Yunhe Liu, Jiangjiang Qin, Jianfeng Chen, Jingjing Wu, Melissa P Pizzi, Rossana Lazcano, Kohei Yamashita, Zhiyuan Xu, Guangsheng Pei, et al. Meti: deep profiling of tumor ecosystems by integrating cell morphology and spatial transcriptomics. *Nature communications*, 15(1):7312, 2024.
- [13] ELHAM Karimi, N Simo, N Milet, W TE, A ALSH,

- ND QU, L AIL, R ABS, A ALIND, ND MORRIS GOODMA, et al. Method of the year 2024: spatial proteomics. *Nat Methods*, 21:2195–2196, 2024.
- [14] Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. Umap: Uniform manifold approximation and projection. *Journal of Open Source Software*, 3(29):861, 2018.
- [15] Lambda Moses and Lior Pachter. Museum of spatial transcriptomics. *Nature methods*, 19(5):534–546, 2022.
- [16] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. arXiv preprint arXiv:1807.03748, 2018.
- [17] Nejla Ozirmak Lermi, Max Molina Ayala, Sharia Hernandez, Wei Lu, Khaja Khan, Alejandra Serrano, Idania Lubo, Leticia Hamana, Katarzyna Tomczak, Sean Barnes, et al. Comparison of imaging-based single-cell resolution spatial transcriptomics profiling platforms using formalin-fixed, paraffin-embedded tumor samples. *bioRxiv*, pages 2024–12, 2024.
- [18] Xiaoxi Pan, Maria E Salvatierra, Caner Ercan, Lakshmi Kakarala, Wei Lu, Ou Shi, Idania C Lubo Julio, Ignacio I Wistuba, Luisa M Solis Soto, and Yinyin Yuan. Tmeseg: Connecting histopathology with spatial transcriptomics through tumor microenvironment segmentation for lung cancer. *Cancer Research*, 85(8_Supplement_1):2426–2426, 2025.
- [19] Beibei Ru, Jinlin Huang, Yu Zhang, Kenneth Aldape, and Peng Jiang. Estimation of cell lineages in tumors from spatial transcriptomics data. *Nature Communications*, 14(1):568, 2023.
- [20] Christina V Theodoris, Ling Xiao, Anant Chopra, Mark D Chaffin, Zeina R Al Sayed, Matthew C Hill, Helene Mantineo, Elizabeth M Brydon, Zexian Zeng, X Shirley Liu, et al. Transfer learning enables predictions in network biology. *Nature*, 618(7965):616–624, 2023.
- [21] Hongyi Wang, Xiuju Du, Jing Liu, Shuyi Ouyang, Yen-Wei Chen, and Lanfen Lin. M2ost: Many-to-one regression for predicting spatial transcriptomics from digital pathology images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 7709–7717, 2025.
- [22] Ronald Xie, Kuan Pang, Sai Chung, Catia Perciani, Sonya MacParland, Bo Wang, and Gary Bader. Spatially resolved gene expression prediction from histology images via bi-modal contrastive learning. *Advances in Neural Information Processing Systems*, 36:70626–70637, 2023.
- [23] Hanwen Xu, Naoto Usuyama, Jaspreet Bagga, Sheng Zhang, Rajesh Rao, Tristan Naumann, Cliff Wong, Zelalem Gero, Javier González, Yu Gu, et al. A whole-slide foundation model for digital pathology from real-world data. *Nature*, 630(8015):181–188, 2024.
- [24] Daiwei Zhang, Amelia Schroeder, Hanying Yan, Haochen Yang, Jian Hu, Michelle YY Lee, Kyung S Cho, Katalin Susztak, George X Xu, Michael D Feldman, et al. Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. *Nature biotechnology*, 42(9):1372–1377, 2024.
- [25] Weiqin Zhao, Zhuo Liang, Xianjie Huang, Yuanhua Huang, and Lequan Yu. Hist2cell: Deciphering fine-grained cellular architectures from histology images. *bioRxiv*, pages 2024–02, 2024.