Holon-S1 Technical Report

AUTHORS
Hassan Muhammad, Benjamin Warner, Paul S. Scotti, Tanishq Mathew Abraham

DATE
October 13, 2026

ABSTRACT

- Introducing **Holon-S1** our breakthrough model that acts as a flashlight into the "black box" of pathology foundation models, transcribing latent representations of morphology into spatially resolved genomic profiles, signalling pathways & biological programs.
- **Holon-S1** outperforms Path2Space and DeepSpot-M on gene expression prediction from H&E.
- We demonstrate **Holon-S1** applications for biomarker discovery and precision medicine across several case studies.
- Holon-S1 uncovered a novel chemotherapy sensitivity biomarker in triple negative breast cancer, previously only known to play a role in other cancer types.

1 Introduction

Tumors and surrounding healthy tissue form an interconnected biological system of gene expression, protein activity, cellular structure, and tissue mechanics. Multiple technologies have been developed to capture and quantify aspects of this system, ranging from multiplex imaging to single cell sequencing. However, in routine clinical care, a method originating from 1877 is still the gold standard: H&E-stained microscopic tissue slides qualitatively reviewed by pathologists. Today, these tissue slides can be digitized into high-resolution whole-slide images, creating a rich computational view of tissue morphology. Pathology foundation models can then transform these images into powerful embeddings for downstream applications [1, 2], but the biology encoded within those representations remains largely opaque [3].

Spatial transcriptomics (ST) provides a complementary view of tissue biology by measuring genomic expression profiles while preserving spatial organization. These assays can resolve molecular activity at the level of multicellular spots or individual cells, but are expensive, technically demanding, and difficult to scale [4].

Different ST platforms balance spatial resolution against measurement stability. Single-cell assays such as Xenium/Atera (10x Genomics, Inc.) and CosMx (Bruker Corp.) provide the most granular view, but their measurements can be sparse and sensitive to limited transcript capture, dropout, and segmentation errors. Conventional spot-based assays such as Visium

(10x Genomics, Inc.) produce more stable measurements by aggregating signal across multiple cells, but can obscure fine-grained tissue structure. Intermediate platforms such as Visium HD (10x Genomics, Inc.) offer a useful middle ground by measuring denser regions, combining near-cellular resolution with sufficient molecular signal to produce coherent measurements across large tissue regions.

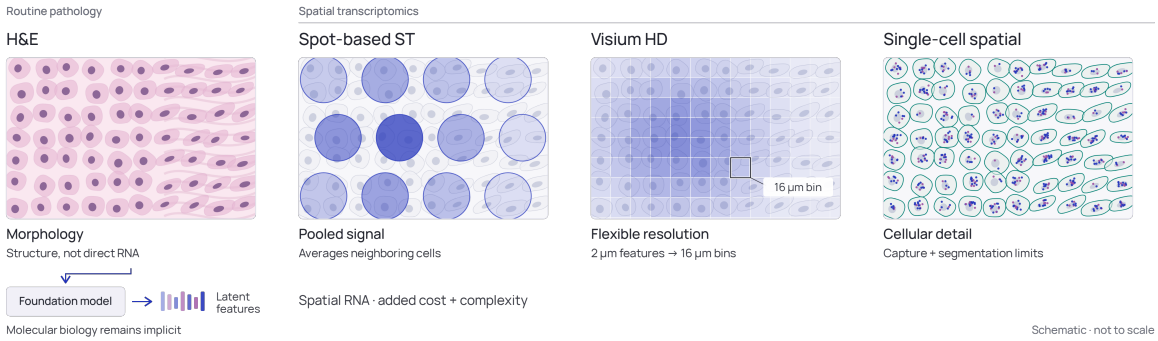


Figure 1 H&E-stained histopathology captures tissue phenotype through morphology, while spatial transcriptomics (ST) captures genotype through spatially resolved gene expression. They offer different views into the same biological system. Differing ST technologies come with different benefits and limitations, including resolution, noise level, and panel size.

Holon-S1 predicts spatially resolved gene expression directly from routine digital pathology H&E images, offering a scalable alternative to ST in real-world workflows. We trained the model using a pan-cancer strategy on data spanning multiple assay platforms and spatial resolutions, including both spot-level and single-cell measurements. This approach was designed to capture the relationship between tissue morphology and gene expression across diverse ST panels while accounting for the trade-offs inherent to each platform.

During training, **Holon-S1** integrated fine-grained cellular morphology with broader slide-level spatial context, capturing both local cell-cluster phenotypes and the larger tissue programs in which they are embedded. These complementary signals enabled the model to represent molecular activity within tumours and their surrounding microenvironments.

We evaluated **Holon-S1** across a diverse set of genomic assays, including ST platforms absent from training, to assess its generalization across unseen measurement technologies. We then explored its utility in retrospective disease characterization, biomarker discovery, and translational research. In an osteosarcoma case study, we tested whether predictions derived from routine whole-slide images could recover findings previously obtained through molecular profiling, despite Holon-S1 not being trained on osteosarcoma. We also examined whether the spatial organization of predicted gene expression could reveal biomarkers associated with treatment response and prognosis, identifying novel markers for triple negative breast cancer and localized colorectal cancer.

2 Case Study: Retrospective analysis of osteosarcoma in a real patient

Our goal in developing **Holon-S1** was to provide patients with clinically relevant insights. To illustrate this, we retrospectively analyzed the case of Sid Sijbrandij [5], co-founder of GitLab, who was diagnosed with osteosarcoma. After exhausting standard-of-care options, Sijbrandij assembled an unusually comprehensive longitudinal profile of his disease, integrating genomic, single-cell, spatial and other data modalities to identify potential therapies targeting the specific biology of his tumor. He subsequently made these data publicly available [6], providing a framework for a more patient-centered approach to precision oncology.

Inspired by Sijbrandij's efforts to accelerate and democratize precision medicine, we tested whether **Holon-S1** could use pathology images to reproduce findings and therapeutic recommendations derived from the specialized genomic panels available to him and his research team. Our second aim was to assess if **Holon-S1** may potentially generate insights beyond genomic profiling alone. His analyses identified a distinct fibroblast program associated with remodelling of the extracellular matrix within the tumor microenvironment, potentially creating a physical barrier to immune-cell infiltration. Consistent with this hypothesis, his team measured low levels of T-cell infiltration. His tumor also exhibited pronounced MDM2 overexpression, placing it in the 99.7th percentile relative to the comparison cohort. Together with additional molecular and cellular findings, these observations provided a rationale for considering immune-checkpoint inhibition, FAP-targeted radioligand therapy directed at the fibroblast program, and MDM2 inhibition, among other therapeutic strategies. We encourage readers to review Sijbrandij's story directly on his website.

2.1 Deploying Holon-S1 on Routine Diagnostic Pathology

Fourteen digitized pathology whole-slide images (WSIs) were collected from Sijbrandij's public database. Each WSI was partitioned into a grid of 56×56 μm regions, for which **Holon-S1** inferred genomic profiles, producing a spatially resolved transcriptomic map of each specimen. Limiting our analysis to only the eleven initial diagnostic specimens, unsupervised clustering was used to identify tumor and its microenvironment from surrounding healthy tissue, before verification by quantifying proliferation, immune, stromal markers known to distinguish tumor landscapes. Gene-set enrichment and differential gene expression analyses were then used to compare the tumor and its microenvironment with histologically normal tissue from the same specimens, as well as with two independent healthy bone and bone-marrow reference samples collected from Sijbrandij's database. Because matched sequencing data from Sijbrandij's healthy tissue were unavailable, his original analyses relied on publicly available normal-tissue expression profiles from other patients. By contrast, **Holon-S1** enabled tumour and adjacent normal tissue to be evaluated within the same pathology specimen.

Holon-S1 predicted higher expression of fibroblast- and fibrosis-associated genes in cancer-like regions, including *FAP* (fibroblast activation protein), the target of the radioligand therapy Sijbrandij received and to which the remission has been tentatively attributed. To prioritize

candidate therapeutic targets, we overlaid a pre-specified panel of ten genes chosen from osteosarcoma clinical-trial and literature context; most are targets of agents in osteosarcoma trials registered on ClinicalTrials.gov. We further highlighted targets of therapies in Sijbrandij's clinical history or in development as backup options. Predicted expression was higher in cancer-like regions for *MDM2* (MDM2 inhibition), *KDR*/VEGFR2 (vaccine) and *CD276*/B7-H3 (antibody–drug conjugate), and *CLEC14A* (CAR-T). Figure 2 summarizes these findings in volcano plots: the x-axis shows the $\log 2$ fold change in predicted expression between tumor-associated regions and healthy reference tissue, while the y-axis shows $-\log 10$ of the false discovery rate–adjusted p-value. Positive fold changes indicate higher predicted expression in tumor-associated regions, and higher points indicate stronger statistical evidence for differential expression.

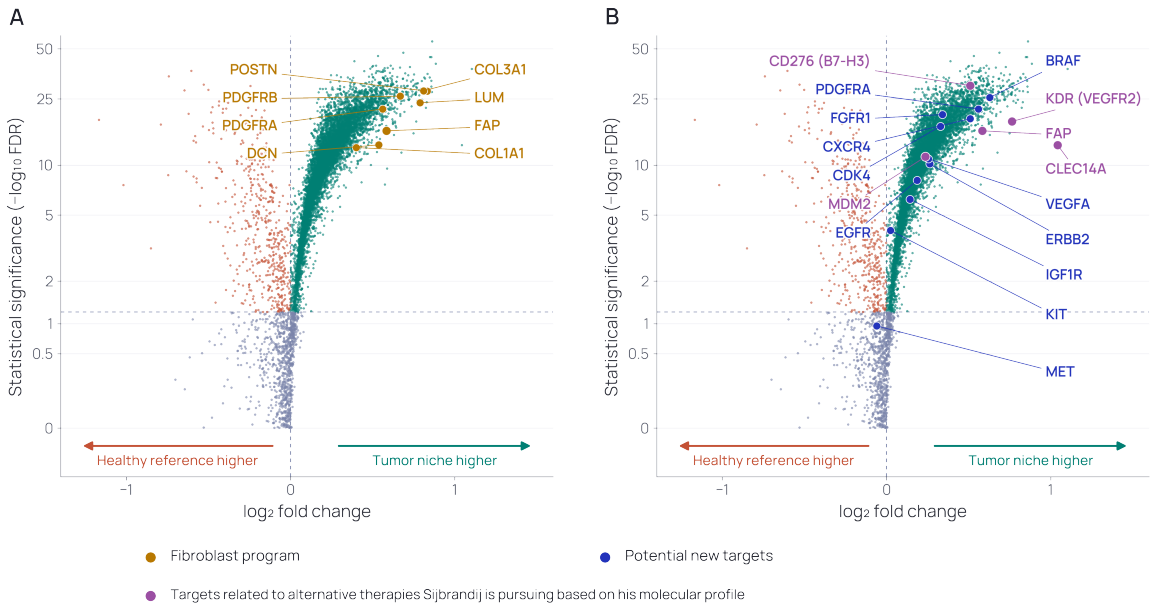


Figure 2 (Left) Fibroblast-associated markers, including FAP, are highlighted in this differential gene expression analysis comparing tumor and non-tumor regions. Fibroblast-associated markers consistently showed higher expression in the tumor niche. (Right) Therapeutically relevant markers are highlighted in the same analysis. Although some show modest expression differences or limited statistical significance, others are significantly differentially expressed, including markers associated with alternative therapies being pursued by Sijbrandij.

To assess whether **Holon-S1** recapitulated the elevated MDM2 expression observed in Sijbrandij's molecular profile, we compared his inferred MDM2 expression with that of an independent osteosarcoma cohort. This cohort comprised 270 publicly available WSIs from ten pediatric patients treated with adjuvant chemotherapy, accessed through Osteosarcoma Explorer [7, 8]. Whereas his clinical team contextualized his molecular measurements against a reference cohort of 2,600 patients with bulk sequencing, our analysis compared Holon-S1-inferred expression across histological samples. Inferred MDM2 expression in Sijbrandij's samples exceeded nine of the 10 comparator patients.

2.2 Tracking Treatment Response Through Longitudinal Pathology

Next, we used **Holon-S1** to analyze WSIs from Sijbrandij's initial resection and a subsequent resection following treatment with several immune-checkpoint inhibitors. We quantified the spatial association between predicted immune- and cancer-marker expression (canonical programs defined by MSigDB Hallmark [9] and PanglaoDB [10]) to assess immune infiltration within tumor regions. The initial specimen showed limited T-cell infiltration, consistent with his molecular profiling, whereas the post-treatment specimen showed substantially greater immune infiltration. Paired immunohistochemistry (IHC) data from the post-treatment resection provided an independent check of the tumor regions identified by **Holon-S1**. Together, these findings corroborated the molecular analyses and illustrated the model's potential to detect treatment-associated changes directly from longitudinal H&E pathology specimens.

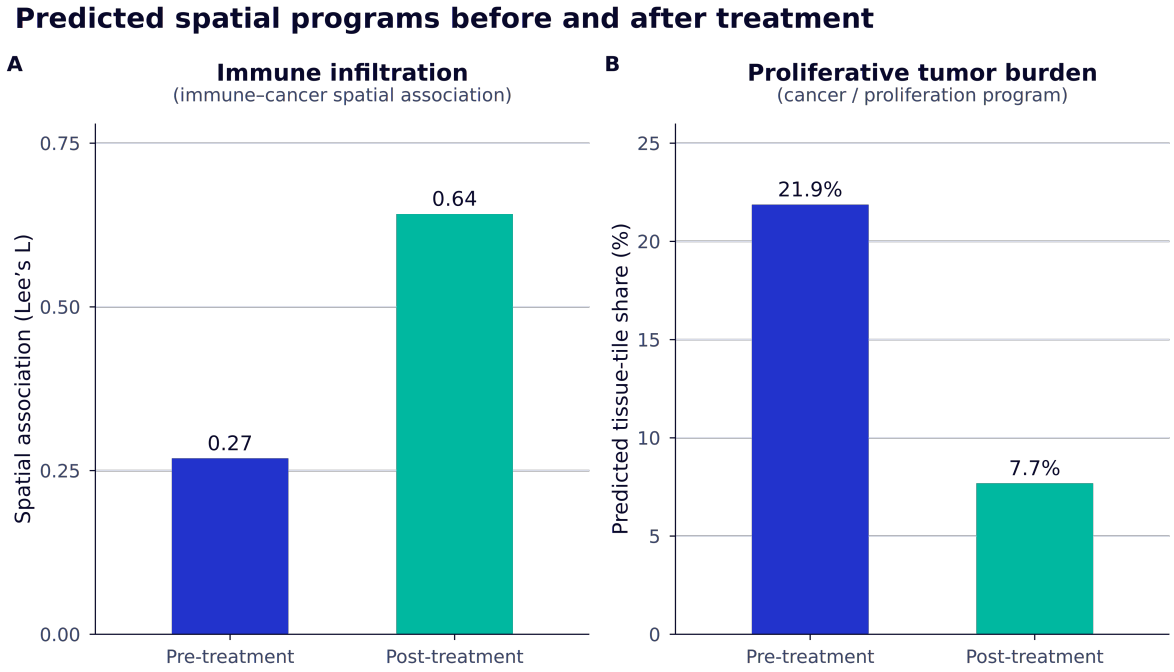


Figure 3 Holon-S1's assessment of pre- and post-treatment resections showed an increase in immune infiltration, as measured by Lee's L spatial coherence between cancer markers and immune markers, and a decrease in cancer proliferation.

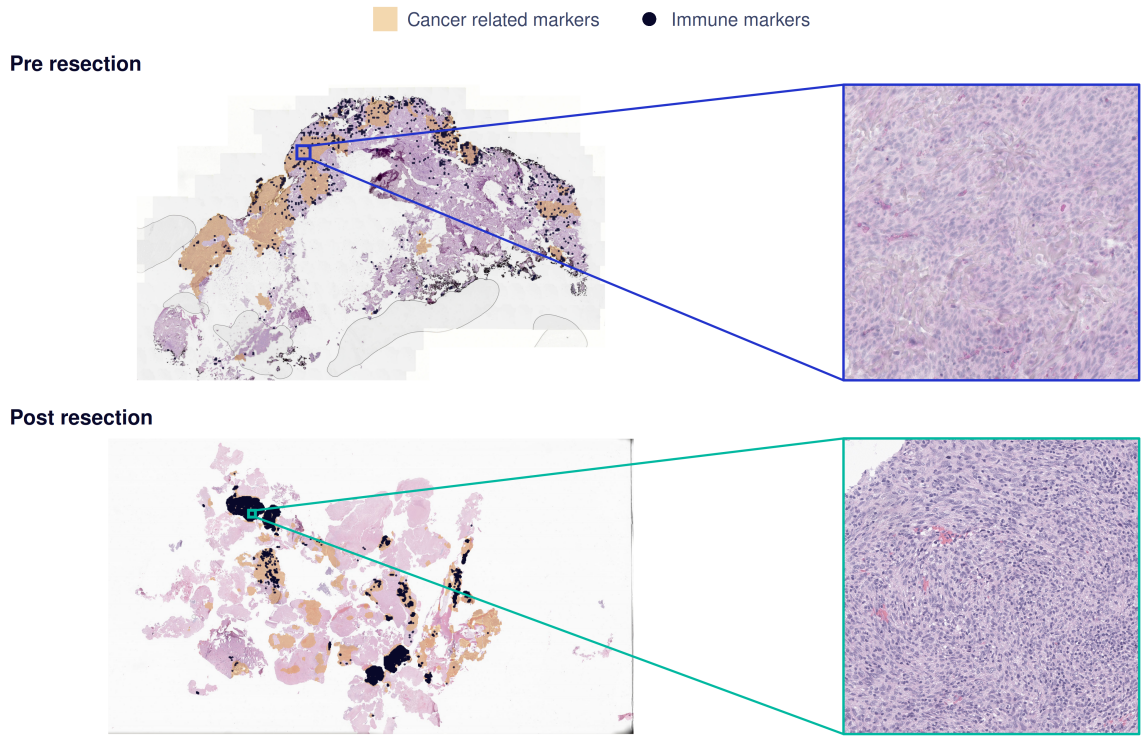


Figure 4 Pre-treatment resection exhibited limited immune presence within the cancer as compared to the post-treatment example. We depict representative regions to help visualize immune infiltration.

2.3 Extending Immune Profiling Beyond Targeted Gene Panels

Sijbrandij's team recently published an in-depth analysis of his Xenium spatial transcriptomic data [11]. Although the targeted Xenium panel enabled a broad range of analyses, its ability to identify tertiary lymphoid structures (TLS) was limited by insufficient representation of B-cell-associated genes. We deployed **Holon-S1** to address this gap. Because B and T cells exhibit similar morphology in conventional histopathology, we first evaluated the model's ability to distinguish between these cell types (see section 4.3). We then applied **Holon-S1** to Sijbrandij's WSIs to identify candidate immune-cell regions and extend the TLS analysis that could not be resolved using the targeted Xenium panel alone.

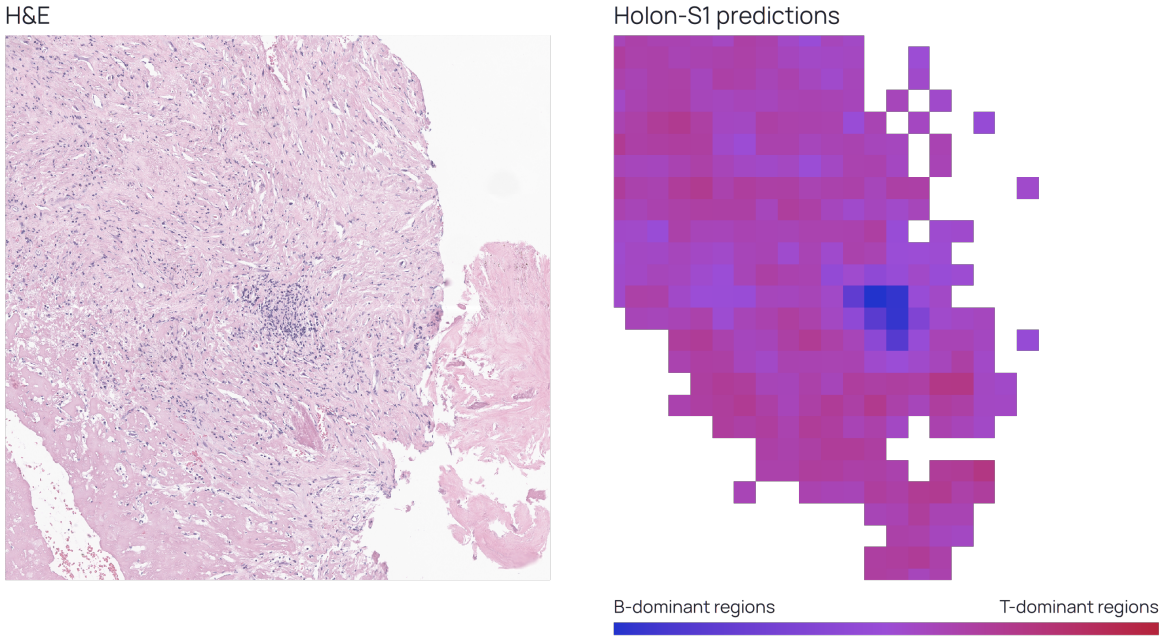


Figure 5 Holon-S1 was able to predict B-cell markers missing from Sijbrandij's original Xenium panels.

2.4 Reflection and Limitations

Despite the absence of osteosarcoma from its training data, **Holon-S1** recovered several findings from Sijbrandij's molecular analyses and generated additional insights directly from routine pathology images. Many specimens had been collected years before the corresponding molecular findings emerged, raising the possibility that relevant biological signals could have been identified earlier in the disease course. This retrospective analysis supports the model's potential to generalize to tumor and tissue types not represented in training and to inform personalized precision medicine workflows. Although **Holon-S1** is currently a research-only tool and not yet approved for patient diagnosis, these findings motivate further evaluation. We are designing additional retrospective and prospective studies to assess reproducibility and establish the model's clinical utility.

3 Applications in Biomarker Discovery and Translational Research

We believe Holon-S1 will play a role beyond direct patient care. Particularly in the development of companion diagnostics and translational biopharma research by accelerating interpretable biomarker discovery. Here, we demonstrate the discovery of simple predictive and prognostic markers which deliver actionable insights into triple negative breast cancer and localized colorectal cancer.

3.1 Spatial Organization for Biomarker Discovery

3.1.1 Predicting Response to Neoadjuvant Chemotherapy in Triple Negative Breast Cancer

Patients with triple negative breast cancer (TNBC) are often candidates for neoadjuvant chemotherapy (NAC). Currently, pathologists quantify immune markers in a patient's tissue to help determine if the patient will respond to NAC or if they should consider alternative treatments. Unfortunately, this immune assessment lacks precision and is subject to inter- and intra-observer variability [12].

We assembled TNBC cohorts with WSIs from IMPRESS [13] and TransNEO [14], with response to NAC defined as pathologic complete response (pCR) at surgery. IMPRESS served as the discovery cohort, while TransNEO was reserved as a held-out test set.

We investigated whether the spatial organization of gene expression predicted by **Holon-S1** could identify markers of treatment response in TNBC. We restricted the analysis to genes meeting our prediction-confidence criteria in breast cancer (Section 4) and quantified local expression variation (gradient), spatial clustering (Getis–Ord), and spatial autocorrelation (Moran's I; Figure 6). In the IMPRESS discovery cohort, we evaluated each gene–metric combination using univariate models and five-fold cross-validation. EREG gradient (gEREG) emerged as a candidate marker associated with pathological complete response (pCR) and was subsequently evaluated in the TransNEO test cohort.

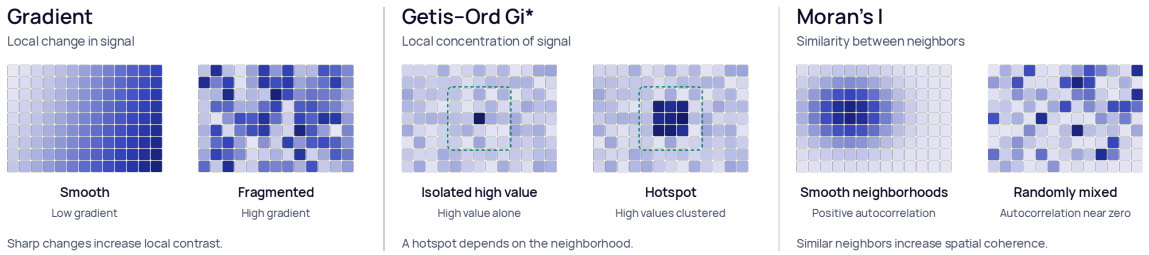


Figure 6 Schematic illustration of the spatial gene-expression metrics evaluated: gradient, Getis–Ord, and Moran's I.

To provide a pathology-based clinical baseline, we generated computational estimates of stromal tumor-infiltrating lymphocyte (sTIL) percentages, a measure conventionally estimated visually by pathologists [15]. We used CytoFormer predictions [16] to delineate stromal regions within the inferred tumor boundary, including stroma between tumor nests while excluding tumor-cell territories. Within these regions, we estimated the percentage of stromal area occupied by lymphocyte and plasma-cell footprints using HoVer-NeXt predictions [17]. These automated scores served as a proxy for manual sTIL assessment.

We summarized precision–recall performance using average precision, with five-fold out-of-fold predictions in IMPRESS and IMPRESS-trained models applied to TransNEO. Separately, we fitted univariate and multivariate logistic regressions within each cohort to estimate odds ratios (OR), 95% confidence intervals, and likelihood-ratio p-values. Multivariate models included both gEREG and sTIL. Odds ratios were expressed per one discovery cohort-computed standard deviation increase in each feature.

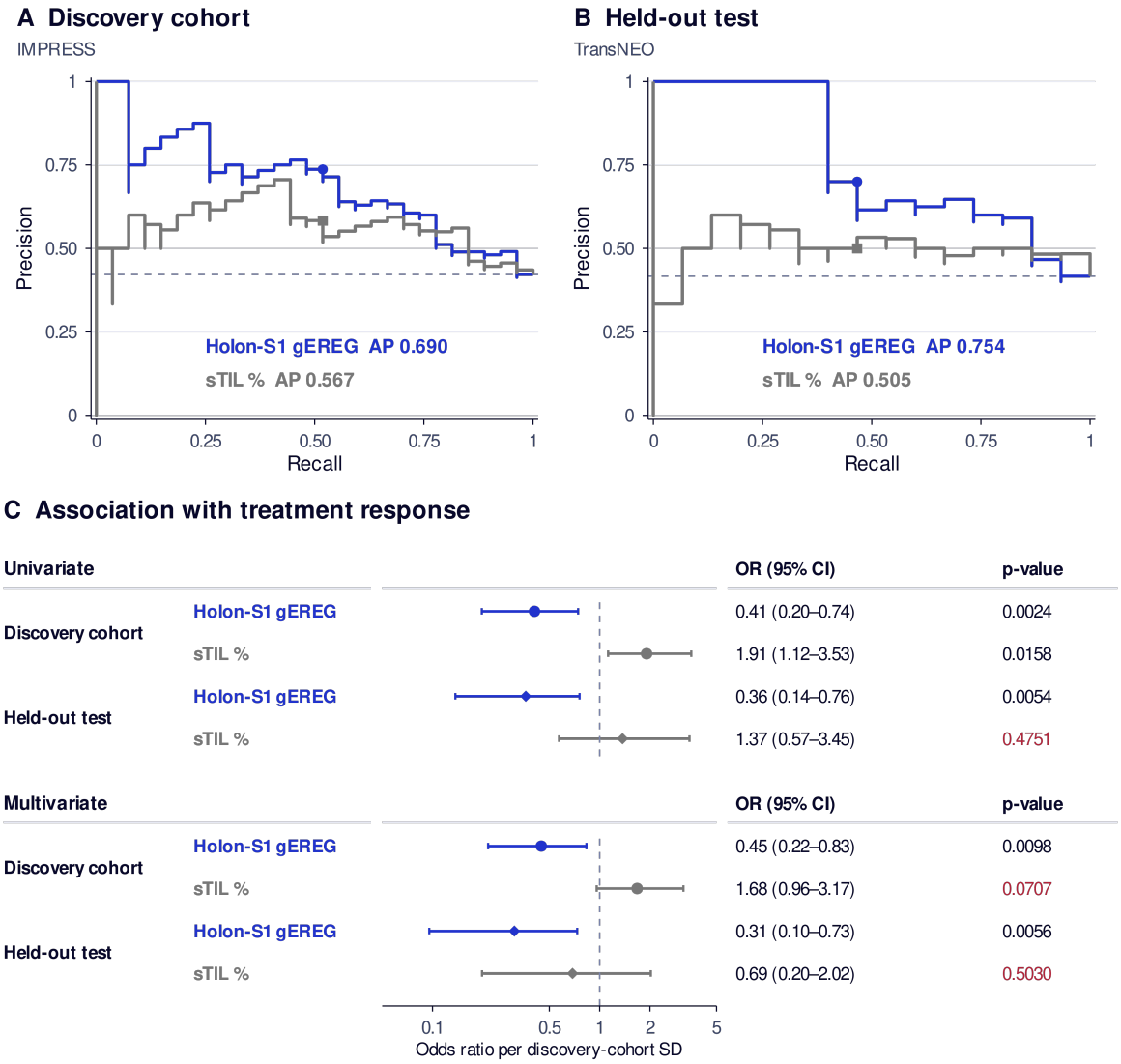


Figure 7 **(A–B)** Precision–recall curves in the discovery and held-out test cohorts for Holon-S1-predicted EREG spatial organization (gEREG) and percent of stroma tumor infiltrating lymphocytes (sTIL); AP indicates average precision. **(C)** Univariate and multivariate odds ratios per discovery-cohort standard deviation, with 95% confidence intervals.

gEREG remained a significant predictor in univariate modelling and in multivariate modelling when adjusting for sTIL. sTIL was unable to significantly predict pCR with the held-out test cohort and did not maintain significance in multivariate modelling. Previous studies [18–20] have implicated EREG expression in sensitivity to chemotherapy and anti-EGFR therapies in other cancer types, even spurring a clinical trial [21] to assess real-world therapeutic implications of these findings. Our findings motivate further investigation into whether the spatial organization of EREG may similarly inform treatment response in breast cancer, highlighting **Holon-S1** as a tool for developing novel insights into cancer and cancer therapy. Future work may include combining this EREG spatial marker with other signatures to further improve clinical decision support for TNBC patients, as well as comparing to sTIL as quantified by expert pathologists.

3.2 Interpretable Spatial Biology for Translational Research

3.2.1 Prognosticating Localized Colorectal Cancer

Spatial colocalization describes whether two biological programs, represented by gene-expression signatures, are active within the same tissue neighborhoods or occupy separate regions. We quantified this relationship using Lee's L, with positive values indicating spatial alignment and negative values indicating segregation. Using **Holon-S1**-predicted expression maps, we evaluated predefined pairs of programs spanning hypoxia, metabolism, inflammatory signaling, vascular biology, proliferation, stroma/EMT, cytotoxic immunity, and cancer epithelium for associations between their spatial colocalization and disease-specific survival in non-metastatic colorectal cancer (CRC). Programs were computed by averaging genes at each spatial location following canonical MSigDB Hallmark [9] and PanglaoDB [10] databases.

We used TCGA-COADREAD [22] as the discovery cohort and reserved SurGen [23] for independent validation. When multiple slides were available, we averaged slide-level colocalization values to obtain one score per patient for each program pair.

Within the discovery cohort, we assessed each pair's association with disease-specific survival using a continuous Cox proportional hazards (CPH) model and five-fold cross-validation. This analysis identified proliferation—stroma/EMT (epithelial-mesenchymal transition) colocalization as a prognostic feature: greater spatial segregation of these programs was associated with better disease-specific survival. The association showed the same risk association in the held-out SurGen cohort.

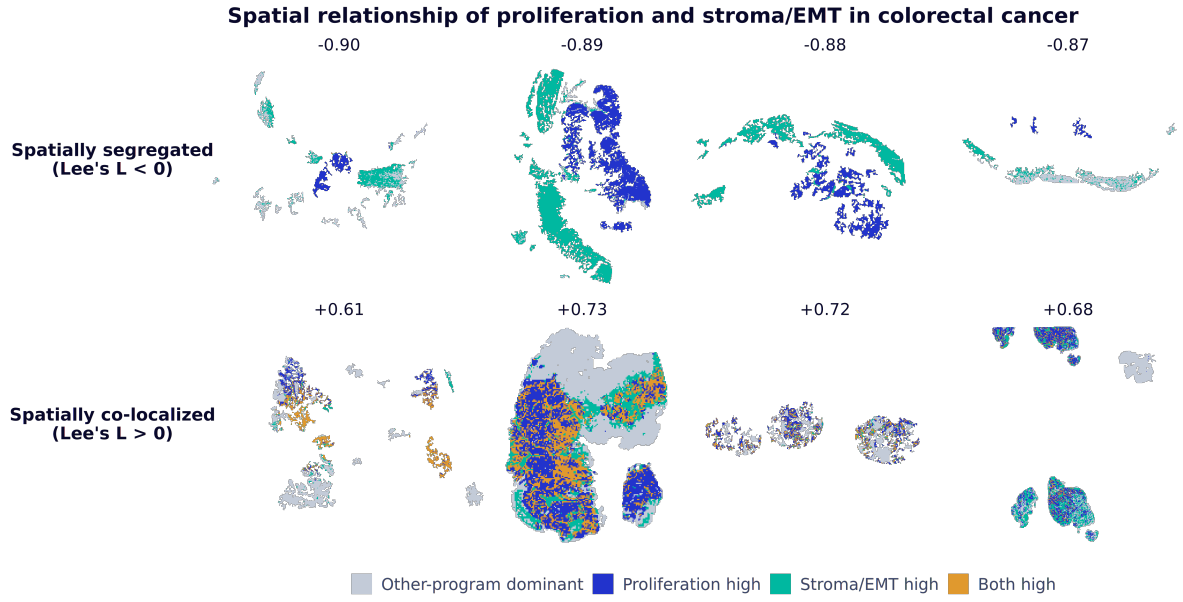


Figure 8 Spatial segregation and co-localization of proliferation and stroma/EMT programs, predicted by Holon-S1

To visualize survival differences, we divided patients into lower- and higher-colocalization groups based on their Lee's L scores for the proliferation–stroma/EMT pair and constructed Ka-

plan–Meier curves and used CPH to compute categorical hazard ratios . Log-rank tests indicated significant differences in disease-specific survival between the groups.

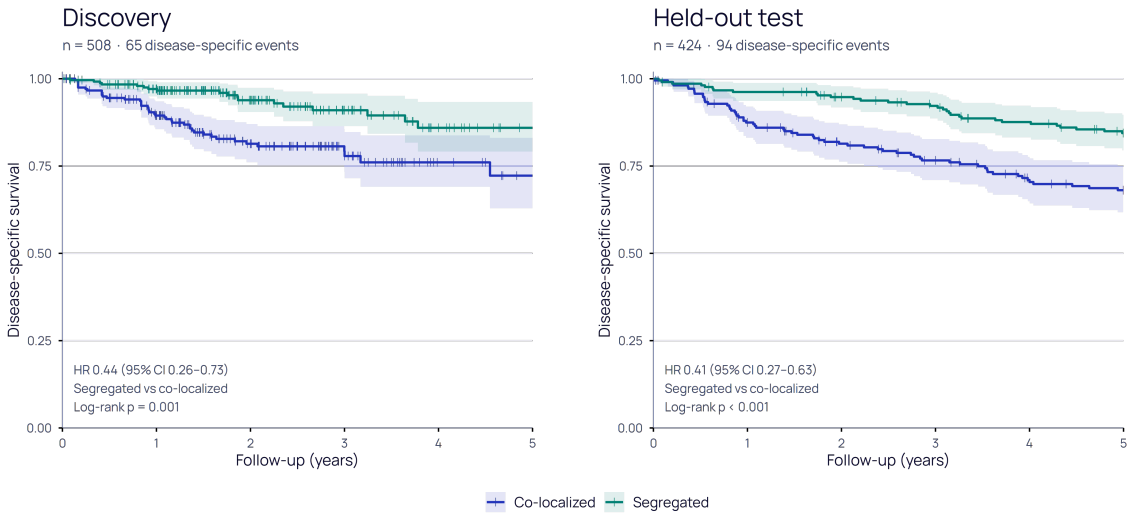


Figure 9 Spatial co-localization/segregation of proliferation and stroma/EMT programs predicted by Holon-S1 significantly stratify (log-rank test $p < 0.05$) patients by disease-specific survival within five years.

Our findings identify spatial colocalization of proliferation and stroma/EMT signatures as a candidate prognostic marker in localized colorectal cancer. This observation adds a spatial perspective to established evidence linking stromal activation and mesenchymal expression programs to poor prognosis [24, 25]. Mechanistic studies have shown that TGF-β-activated cancer-associated fibroblasts can promote metastatic tumor-cell survival through IL-11 secretion and downstream GP130/STAT3 signaling [26]. These studies provide a biological rationale for the association identified using **Holon-S1**. Together, our results illustrate how translational research teams can use **Holon-S1** to identify biologically interpretable candidate biomarkers and generate hypotheses about tumor–microenvironment interactions directly from routine pathology images. Further investigation is needed to determine the cellular interactions and signaling mechanisms underlying this spatial pattern and its prognostic significance.

4 Validating the Language of Holon-S1

This year, several studies have demonstrated the promise of predicting transcriptome-scale gene expression directly from H&E images [27–29]. Still, robust generalization across the full transcriptome remains an open challenge. This may reflect limitations in the scale and diversity of available training data [30] or a more constraining possibility: not every molecular process produces a visible morphological signature. Consequently, virtual transcriptome models predict some genes far more reliably than others. Identifying which genes are recoverable from tissue morphology is therefore as important as improving aggregate performance and model robustness, especially when gene expression predictions are used in clinically important down-

stream modeling. To define this boundary, we validated Holon-S1 on WSIs paired with ST data spanning multiple tissue types and assay platforms. These results allowed us to establish a confidence-filtered set of genes that comprise Holon-S1's translation "vocabulary."

4.1 Technical Validation of Gene Expression Prediction and Gene Curation

Our first finding was a consistent trend between the technological generation of the spatial transcriptomics platform and measured concordance: predictions evaluated against newer, more advanced platforms consistently achieved higher accuracy. Figure 10 demonstrates this observation. Although Xenium 5K (10x Genomics) was the most recent platform represented in Holon-S1's training data, validation on Atera (10x Genomics), released this year, yielded a higher median Pearson correlation coefficient (PCC) than any Xenium sample. This suggests that lower performance against older platforms may partly reflect noise or uncertainty in the reference measurements rather than weaker model predictions. Comparisons between virtual transcriptomics models must therefore account for the sensitivity, resolution, and measurement characteristics of the assays used for validation.

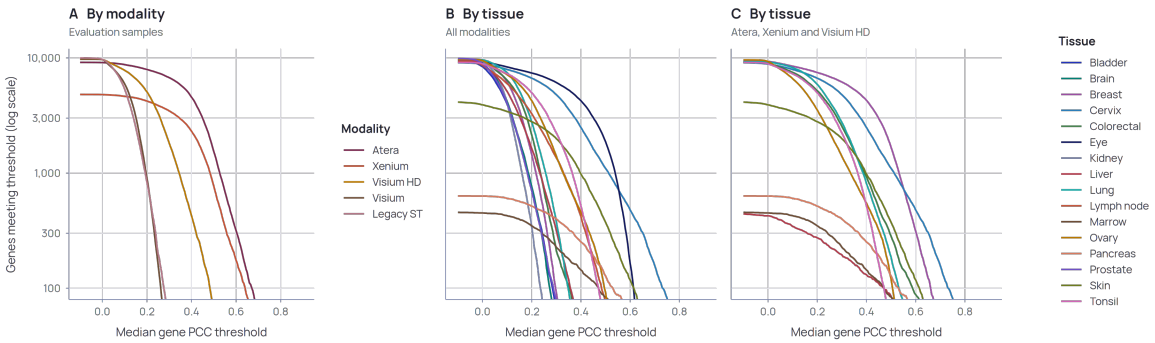


Figure 10 Number of genes meeting increasing prediction-accuracy thresholds across held-out test paired H&E-stained slides and associated spatial transcriptomics (ST) samples unseen during model training, grouped by ST assay technology (A) and tissue type (B - C).

We also observed differences in performance across organ types. However, because our held-out cohorts were profiled using different assay technologies, tissue-specific effects could not be fully separated from platform-specific measurement limitations. We therefore define a confidence-based gene vocabulary for each organ, restricting predictions to genes that are reliably supported by currently available validation data. We expect these vocabularies to expand as Holon-S1 is continually evaluated on larger cohorts generated with newer, higher-sensitivity assays.

4.2 Comparison to SOTA multimodal model

DeepSpot-M [28] was recently released as a multimodal improvement to the previous DeepSpot [31], citing state-of-the-art performance in predicting gene expression from digital pathology at time of release. We compared the ability of Holon-S1 and DeepSpot-M to predict gene expression on slides paired with Atera whole transcriptome sequencing. Holon-S1 outperformed DeepSpot-

M, despite having significantly less breast tissue samples in its training corpus. Both models (to our knowlege) did not include cervical tissue during training, suggesting Holon-S1's ability to better generalize to tissue types unseen in training data.

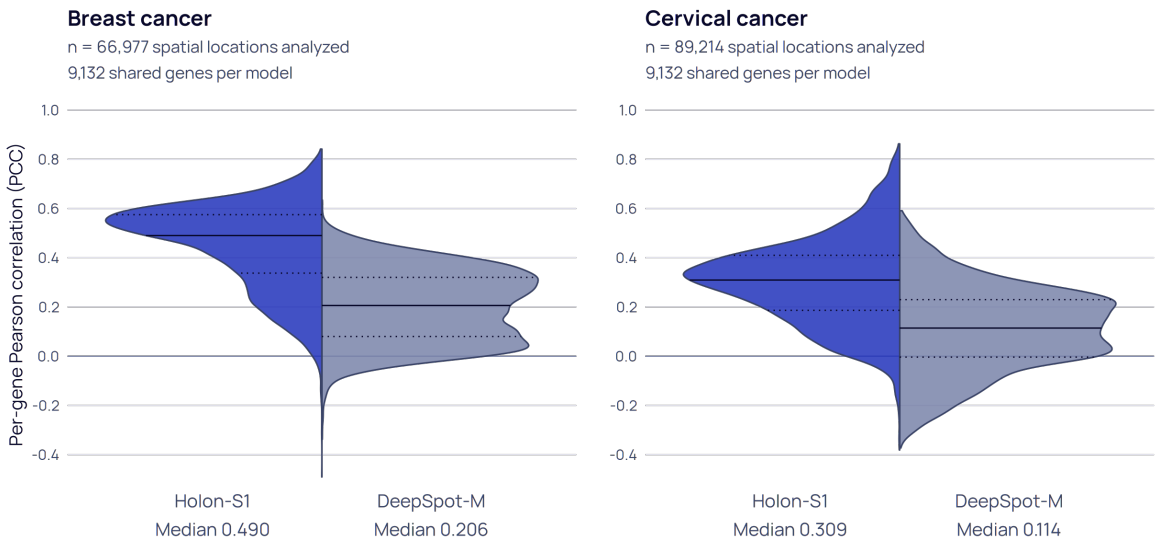


Figure 11 When evaluated on the same samples and same regions within those samples, Holon-S1 (blue) outperformed DeepSpot-M (grey) in gene prediction accuracy as quantified by median per-gene pearson correlation.

4.3 Distinguishing B-cell Dominant Regions in Immune Landscape

To support the identification of B-cell-dominant regions in Sijbrandij's samples, we first evaluated the ability of Holon-S1 to distinguish B- and T-cell lineages from H&E images. These lymphocyte populations are difficult to distinguish reliably by morphology alone, and their identification typically relies on lineage-specific markers assessed by immunohistochemistry or multiplex immunofluorescence.

We used the OrionCRC dataset [32], which comprises H&E-stained whole-slide images paired with multiplex immunofluorescence (IF) images of the same tissue sections. Ground truth was derived from IF, defining B cells as $CD45^{+}CD20^{+}CD3E^{-}$ and T cells as $CD45^{+}CD3E^{+}CD20^{-}$ after excluding autofluorescence-high objects. Cell counts were aggregated into 56-µm tiles without smoothing, with ≥75% lineage representation defining B- or T-dominance. Holon-derived predictions were compared with these labels at matched spatial locations (held-out balanced accuracy: 88.5%; AUROC: 0.943).

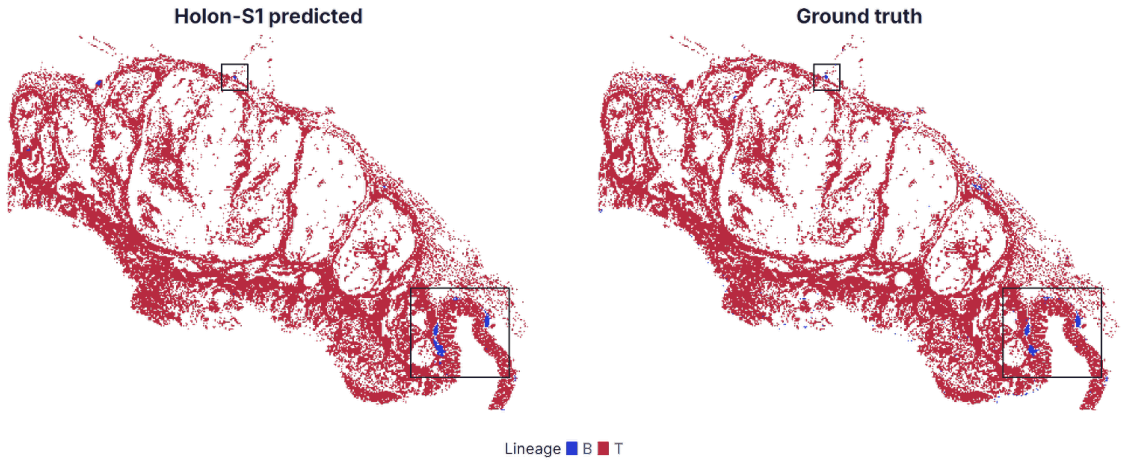


Figure 12 Holon-S1 predictions (left) and immunofluorescence-derived ground truth (right) show B-cell (blue) and T-cell (red) dominance.

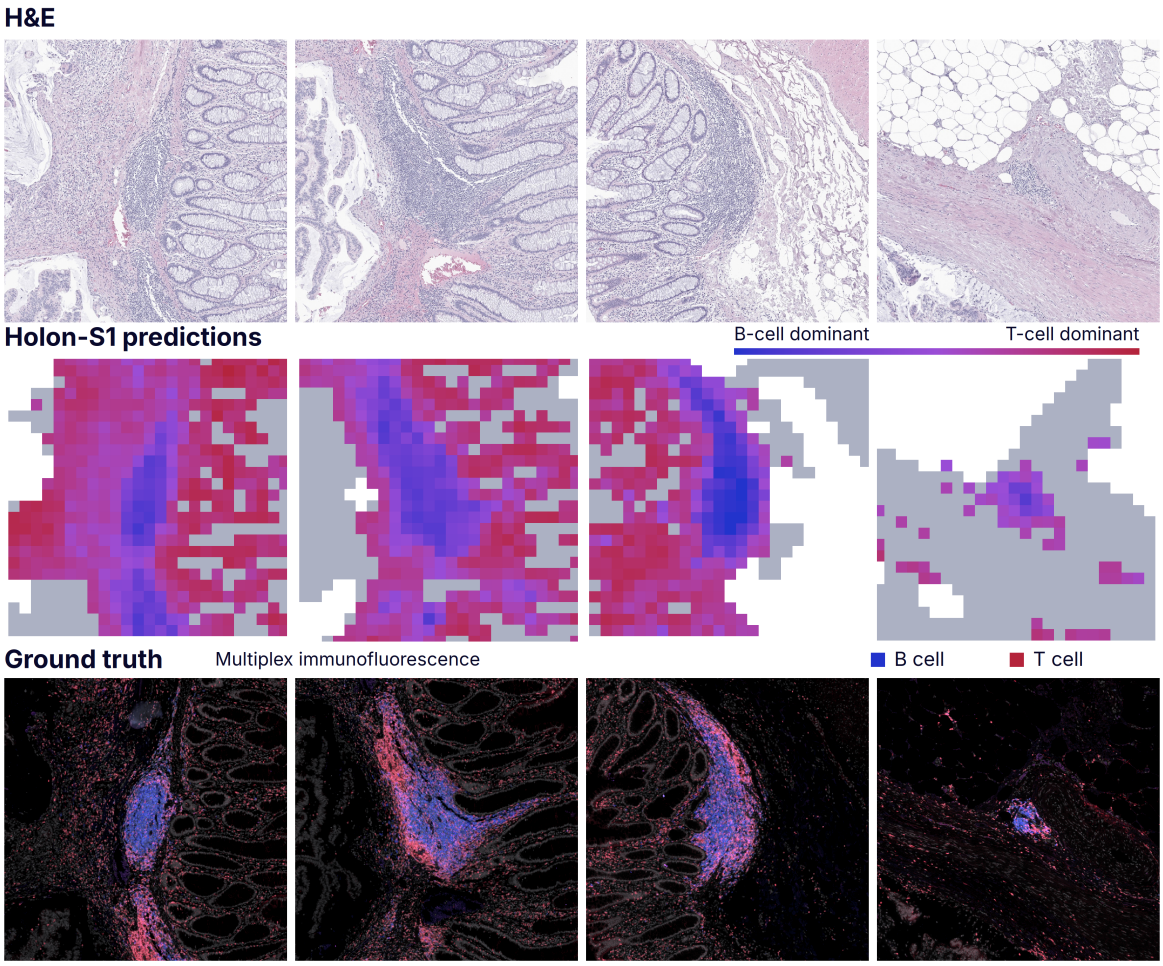


Figure 13 (Top) Selected regions from H&E-stained whole slide image which show dense immune regions with B- and T-lineage immune cells. (Middle) Holon-S1 predictions visualizing model probability towards B- or T-cell regional dominance. (Bottom) Immunofluorescence-derived ground truth.

Because **Holon-S1** infers gene expression at the regional rather than single-cell level, we evaluated its ability to distinguish T-cell-dominant from B-cell-dominant regions. Although these predictions do not resolve individual cell identities, they accurately distinguish regions by their dominant lymphocyte lineage. Current future directions include cell-level prediction models.

4.4 Analytical Validation and Robustness

Variations in tissue preparation, staining, and scanning across institutions pose a persistent challenge to the robustness of pathology foundation models. We evaluated **Holon-S1** on slides with substantially degraded image quality and assessed prediction variability when the same biological specimens were scanned using different protocols.

Path2Space, published in Cell earlier this year [27], reported improved gene-expression prediction relative to more than 20 existing methods. One benchmark cohort included 40 breast cancer slides profiled using Visium across formalin-fixed paraffin-embedded (FFPE) and fresh-frozen specimens, collected from the Human Tumor Atlas Network [33]. Following strict quality control metrics, Path2Space retained a small subset of samples for evaluation.

We wanted to evaluate Holon-S1 across all samples in the HTAN dataset, including those deemed to be too low quality by the Path2Space framework. For each slide, we summarized Holon-S1 prediction accuracy as the median per-gene Pearson correlation against log1p-transformed raw counts. We compared FFPE ($n = 8$ slides) and fresh-frozen ($n = 31$ slides) performance using a two-sided exact Mann–Whitney U test, without adjustment for within-patient clustering.

Holon-S1 achieved a higher median correlation than the reported Path2Space reference on the evaluated subset and maintained comparable performance on slides excluded by Path2Space's quality-control criteria. Performance did not differ significantly between FFPE and fresh-frozen slides ($p = 0.184$).

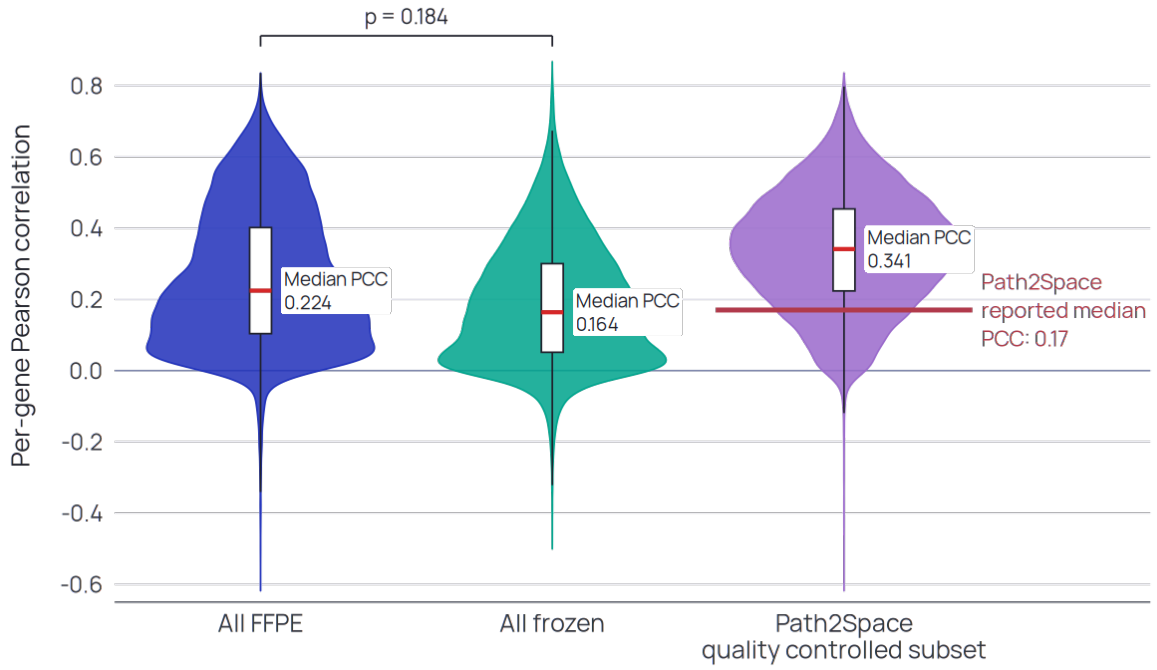


Figure 14 We ran Holon-S1 on the HTAN dataset which included slides with severely degraded quality, prepared from FFPE and frozen tissue, and paired with Visium (10x Genomics, Inc.) panels. There was no statistically siginificance (p=0.184) of median per-gene pearson correlation between slides produced from the two different preperation methods. Holon-S1 was also run on the quality-curated subset of slides analyzed by Shulman et. al. to directly compare against Path2Space. Reported performance on these slides is annotated.

The various **Holon-S1** gene-expression prediction evaluations presented above included slides acquired using more than 15 scanner models. To directly assess prediction consistency when the same tissue section was scanned on different systems, we used the PAR dataset [34], which comprises prostate biopsy slides from Erbil, Iraq, each digitized using Grundium, Hamamatsu, and Leica scanners.

We randomly selected ten specimens and evaluated a fixed panel of 50 genes with strong prediction performance in both **Holon-S1** and DeepSpot-M determined in section 4.2. Both models were evaluated at identical tissue coordinates within each scan, produced by the same preprocessing pipeline, with shared artifact exclusions.

Because the scans across different scanners were not spatially registered, we compared predicted expression distributions using Wasserstein-1 distance (WD) [35]. Predictions were normalized by the median within-scan standard deviation across the three scanners for each specimen, model, and gene before distances were calculated. Each specimen was summarized by the median of its 150 gene–scanner-pair distances. **Holon-S1** achieved lower normalized WD in nine of ten specimens, with an overall median of 0.190 compared with 0.523 for DeepSpot-M. These results assess distributional consistency; without matched RNA measurements or spatial registration, they do not establish expression accuracy or spatial localization accuracy.

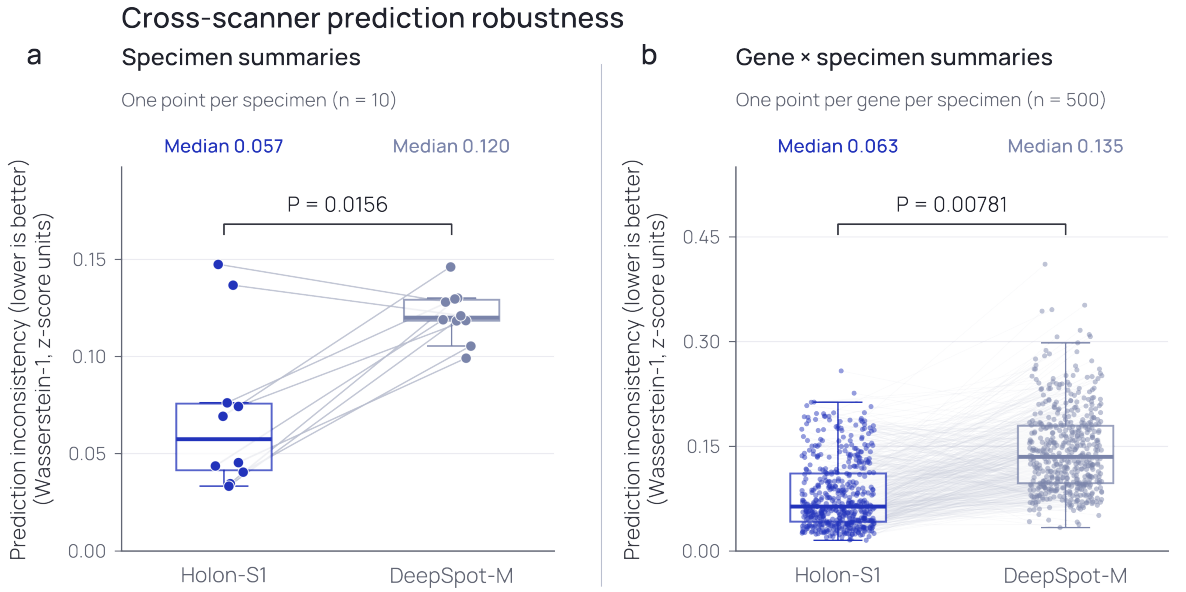


Figure 15 Holon-S1 produced more consistent standardized expresasion distributions than DeepSpot-M across three scanners for 50 genes in 10 PAR specimens. Points represent individual specimens (a) or gene–specimen combinations (b). Higher Wasserstein-1 distances indicate greater inconsistency. Lines connect matched comparisons. P-values are computed by two-sided paired permutation tests that preserve specimen grouping.

For qualitative assessment, we visualized DDR1 predictions in a separate illustrative specimen. DDR1 was selected before inspecting the PAR maps because it had the highest minimum prediction accuracy across both models. Maps were normalized to within-scan z-scores and displayed on a shared color scale without spatial smoothing. **Holon-S1** showed more consistent regional patterns across rescans, whereas DeepSpot-M showed greater variation in relative regional contrast, particularly in the Leica scan.

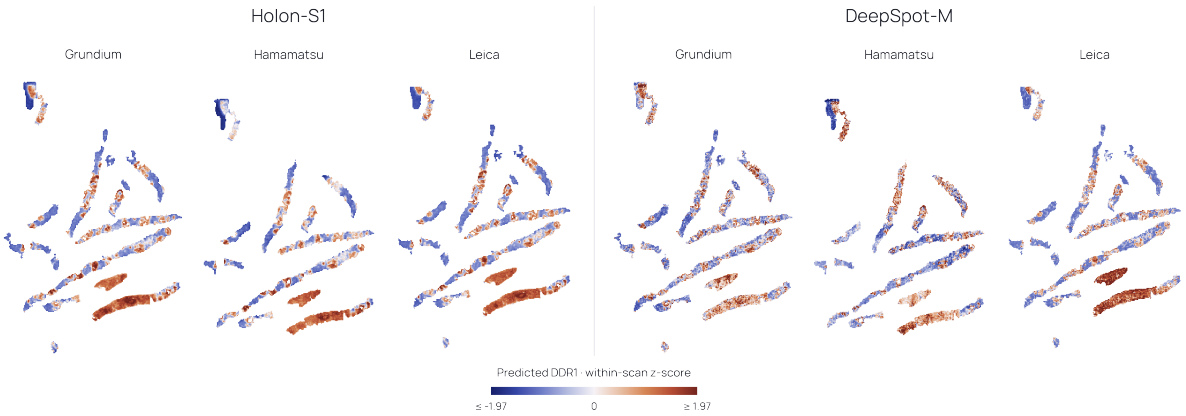


Figure 16 DDR1 expression was selected as a qualitative example to demonstrate Holon-S1 and DeepSpot-M prediction consistency. Both models have a high prediction accuracy of DDR1 expression per validation but differ in prediction consistency when evaluated on paired scanners.

Citation

For attribution in research contexts, please cite this work as:

> Hassan Muhammad, Benjamin Warner, Paul S. Scotti, Tanishq Mathew Abraham. Holon-S1 Technical Report. Sophont Technical Report, 2026. https://sophont.med/blog/holon/.

BIBTEX CITATION

```
@techreport{muhammad2026holon,
  author = {Muhammad, Hassan and Warner, Benjamin and Scotti, Paul S. and Abraham,
    Tanishq Mathew},
  title = {{Holon-S1} Technical Report},
  institution = {Sophont},
  year = {2026},
  month = oct,
  url = {https://sophont.med/blog/holon/},
}
```

References

[1] Mieko Ochi, Daisuke Komura, and Shumpei Ishikawa. Pathology foundation models. *JMA journal*, 8(1):121–130, 2025.

[2] Daniel Kaplan, Ratna Sagari Grandhi, Connor Lane, Benjamin Warner, Tanishq Mathew Abraham, and Paul S Scotti. How to train a state-of-the-art pathology foundation model with $1.6 k. *Sophont Blog*, 2025.

[3] Nhat Minh Le, Ciyue Shen, Neel Patel, Chintan Shah, Darpan Sanghavi, Blake Martin, Alfred Eng, Harshith Padigela, Raymond Biju, Sean Grullon, et al. Evaluating the utility of sparse autoencoders for interpreting a pathology foundation model. In *NeurIPS 2025 Workshop for Imageomics: Discovering Biological Knowledge from Images Using AI*, 2025.

[4] Kelly D Smith, David K Prince, James W MacDonald, Theo K Bammler, and Shreeram Akilesh. Challenges and opportunities for the clinical translation of spatial transcriptomics technologies. *Glomerular Diseases*, 4(1):49–63, 2024.

[5] Sid Sijbrandij. Cancer. https://sytse.com/cancer/.

[6] Sid Sijbrandij's Osteosarcoma Data. https://osteosarc.com/.

[7] Donghan M Yang, Qinbo Zhou, Lauren Furman-Cline, Xian Cheng, Danni Luo, Hongyin Lai, Yueqi Li, Kevin W Jin, Bo Yao, Patrick J Leavey, et al. Osteosarcoma explorer: a data commons with clinical, genomic, protein, and tissue imaging data for osteosarcoma research. *JCO Clinical Cancer Informatics*, 7:e2300104, 2023.

[8] Ka Yaw Teo, Ovidiu Daescu, Kevin Cederberg, Anita Sengupta, and Patrick J Leavey. Correlation of histopathology and multi-modal magnetic resonance imaging in childhood

osteosarcoma: predicting tumor response to chemotherapy. *PLoS One*, 17(2):e0259564, 2022.

[9] Arthur Liberzon, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P Mesirov, and Pablo Tamayo. The molecular signatures database hallmark gene set collection. *Cell systems*, 1(6):417–425, 2015.

[10] Oscar Franzén, Li-Ming Gan, and Johan LM Björkegren. Panglaodb: a web server for exploration of mouse and human single-cell rna sequencing data. *Database*, 2019:baz046, 2019.

[11] Sid Sijbrandij's Osteosarcoma Data. Osteosarcoma xenium: Analysis and interpretation. https://osteosarc.com/xenium/analysis/. Accessed 27 March 2026.

[12] Nurkhairul Bariyah Baharun, Mohamed Afiq Hidayat Zailani, Afzan Adam, Qiaoyi Xu, Muaatamarulain Mustangin, and Reena Rahayu Md Zin. Assessment of tumor-infiltrating lymphocytes in triple-negative breast cancer: Interobserver variability and contributing factors. *Diagnostics*, 15(19):2492, 2025.

[13] Zhi Huang, Wei Shao, Zhi Han, Ahmad Mahmoud Alkashash, Carlo De la Sancha, Anil V Parwani, Hiroaki Nitta, Yanjun Hou, Tongxin Wang, Paul Salama, et al. Artificial intelligence reveals features associated with breast cancer neoadjuvant chemotherapy responses from multi-stain histopathologic images. *NPJ precision oncology*, 7(1):14, 2023.

[14] Stephen J. Sammut et al. TransNEO neoadjuvant breast cancer study. European Genome-phenome Archive, 2021. URL https://ega-archive.org/studies/EGAS00001004582. Data set, accession EGAS00001004582.

[15] Xia Yang, Jia Rao, Wentao Yang, and Ruohong Shui. Evaluation of the predictive and prognostic values of stromal tumor-infiltrating lymphocytes in her2-positive breast cancers treated with neoadjuvant chemotherapy: X. yang et al. *Targeted oncology*, 13(6):757–767, 2018.

[16] Jialu Yao, Songhao Li, Alina Yu, and Zhi Huang. Cytoformer: A molecularly supervised cell foundation model for histopathology cell classification. *Precision Pathology*, page 100006, 2026.

[17] Elias Baumann, Bastian Dislich, Josef Lorenz Rumberger, Iris D Nagtegaal, Maria Rodriguez Martinez, and Inti Zlobec. Hover-next: A fast nuclei segmentation and classification pipeline for next generation histopathology. *Proceedings of machine learning research*, 250:61–86, 2024.

[18] Daniela Conticelli, Marco Volante, Filippo Pietrantonio, Claudia Orrù, Martina Olivero, Alessia Nottegar, Felice Borghi, Gian L Baiocchi, Giovanni Crotti, Uberto Fumagalli Romario, et al. Areg and ereg are predictive biomarkers of response to egfr inhibition in gastroesophageal cancer. *Cancer research*, 85(16):3111–3122, 2025.

[19] Yujia Zhang, Fengjun Qiu, Tingjie Ye, Sau Har Lee, Jiatuo Xu, Lingyan Jia, Rui Zeng, Xiaoling Wang, Xudong Hu, Xiaofeng Yan, et al. Epiregulin increases stemness-associated genes expression and promotes chemoresistance of non-small cell lung cancer via erk signaling. *Stem cell research & therapy*, 13(1):197, 2022.

[20] Wei Xu, Yujia Zhang, Fengjun Qiu, Tingjie Ye, Sau Har Lee, Jiatuo Xu, Lingyan Jia, Rui Zeng, Xiaoling Wang, Xudong Hu, et al. Ereg increases stemness associated genes expression and promotes chemoresistance of non-small cell lung cancer via erk signaling. 2022.

[21] University of Leeds. Association between tumour amphiregulin, epiregulin and epidermal growth factor receptor (EGFR) expression and response to anti-EGFR agents in colorectal cancer. ClinicalTrials.gov, 2019. URL https://clinicaltrials.gov/study/NCT03986541. Clinical trial registration NCT03986541. Updated May 28, 2024. Accessed September 15, 2026.

[22] John N Weinstein, Eric A Collisson, Gordon B Mills, Kenna R Shaw, Brad A Ozenberger, Kyle Ellrott, Ilya Shmulevich, Chris Sander, and Joshua M Stuart. The cancer genome atlas pan-cancer analysis project. *Nature genetics*, 45(10):1113–1120, 2013.

[23] Craig Myles, In Hwa Um, Craig Marshall, David Harris-Birtill, and David J Harrison. Surgen: 1020 h&e-stained whole-slide images with survival and genetic markers. *GigaScience*, 14: giaf086, 2025.

[24] Alexandre Calon, Enza Lonardo, Antonio Berenguer-Llergo, Elisa Espinet, Xavier Hernando-Momblona, Mar Iglesias, Marta Sevillano, Sergio Palomo-Ponce, Daniele V. F. Tauriello, Daniel Byrom, Carme Cortina, Clara Morral, Carles Barceló, Sebastien Tosi, Antoni Riera, Camille Stephan-Otto Attolini, David Rossell, Elena Sancho, and Eduard Batlle. Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nature Genetics*, 47(4): 320–329, 2015. doi: 10.1038/ng.3225.

[25] Claudio Isella, Andrea Terrasi, Sara Erika Bellomo, Consalvo Petti, Giovanni Galatola, Andrea Muratore, Alfredo Mellano, Rebecca Senetta, Adele Cassenti, Cristina Sonetto, Giorgio Inghirami, Livio Trusolino, Zsolt Fekete, Mark De Ridder, Paola Cassoni, Guy Storme, Andrea Bertotti, and Enzo Medico. Stromal contribution to the colorectal cancer transcriptome. *Nature Genetics*, 47(4):312–319, 2015. doi: 10.1038/ng.3224.

[26] Alexandre Calon, Elisa Espinet, Sergio Palomo-Ponce, Daniele V. F. Tauriello, Mar Iglesias, María Virtudes Céspedes, Marta Sevillano, Cristina Nadal, Peter Jung, Xiang H.-F. Zhang, Daniel Byrom, Antoni Riera, David Rossell, Ramón Mangues, Joan Massagué, Elena Sancho, and Eduard Batlle. Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. *Cancer Cell*, 22(5):571–584, 2012. doi: 10.1016/j.ccr.2012.08.013.

[27] Eldad D Shulman, Emma M Campagnolo, Roshan Lodha, Youngmin Chung, Amos Stemmer, Thomas Cantore, Beibei Ru, Tian-Gen Chang, Sumona Biswas, Saugato Rahman Dhruba, et al. Ai-predicted spatial transcriptomics unlocks breast cancer biomarkers from pathology. *Cell*, 2026.

[28] Kalin Nonchev, Sebastian Dawo, Karina Silina, Viktor H Koelzer, and Gunnar Rätsch. Deepspot-m: a multimodal foundation model for transcriptome-wide virtual spatial transcriptomics from histology. *medRxiv*, pages 2026–06, 2026.

[29] Abhishaike Mahajan. Tario-2: A whole-transcriptome foundation model from he alone. Noetik, February 2026. URL https://www.noetik.blog/p/tario-2-a-whole-transcriptome-foundation.

[30] Abhishaike Mahajan. Scaling behavior of TARIO. Noetik, February 2026. URL https://www.noetik.blog/p/scaling-behavior-of-tario.

[31] Kalin Nonchev, Sebastian Dawo, Karina Silina, Holger Moch, Sonali Andani, Tumor Profiler Consortium, Viktor H Koelzer, and Gunnar Rätsch. Deepspot: Leveraging spatial context for enhanced spatial transcriptomics prediction from h&e images. *MedRxiv*, pages 2025–02, 2025.

[32] Jia-Ren Lin, Yu-An Chen, Daniel Campton, Jeremy Cooper, Shannon Coy, Clarence Yapp, Juliann B Tefft, Erin McCarty, Keith L Ligon, Scott J Rodig, et al. High-plex immunofluorescence imaging and traditional histology of the same tissue section for discovering image-based biomarkers. *Nature cancer*, 4(7):1036–1052, 2023.

[33] Orit Rozenblatt-Rosen, Aviv Regev, Philipp Oberdoerffer, Tal Nawy, Anna Hupalowska, Jennifer E Rood, Orr Ashenberg, Ethan Cerami, Robert J Coffey, Emek Demir, et al. The human tumor atlas network: charting tumor transitions across space and time at single-cell resolution. *Cell*, 181(2):236–249, 2020.

[34] Peshawa J Muhammad Ali, Navin Vincent, Saman S Abdulla, Han N Mohammed Fadhl, Anders Blilie, Kelvin Szolnoky, Julia Anna Mielcarz, Xiaoyi Ji, Kimmo Kartasalo, Abdulbasit K Al-Talabani, et al. The par dataset: Prostate biopsy whole slide images from an underrepresented middle eastern population. *Scientific Data*, 13(1):1061, 2026.

[35] Victor M Panaretos and Yoav Zemel. Statistical aspects of wasserstein distances. *Annual review of statistics and its application*, 6(1):405–431, 2019.