

SSC2025 Category/Theme:: Health — Health for a Sustainable Tomorrow

Multimodal Tissue Models: Integrating Imaging and Molecular Data for Discovery

Tianyi Zhang¹, Dwivedi Naman², Malay Singh¹, Timothy Antoni², Shimazawa Kei¹, Jing Yang Tham³, Ju Can², Firas Elbayoumi¹, Dhanav Handa², Tiantian He⁵, Iain Bee Huat Tan^{2,3,4}, Hwee Kuan Lee^{1,6}, Shyam Prabhakar^{2,7,8}, Grace Hui Ting Yeo^{1,2}

¹Agency for Science, Technology and Research (A*STAR), Bioinformatics Institute, Singapore,

²Agency for Science, Technology and Research (A*STAR), Genome Institute of Singapore, Singapore, Singapore, Singapore,

³National Cancer Centre, Division of Medical Oncology, Singapore, Singapore,

⁴National University of Singapore, Duke-National University of Singapore Medical School, Singapore, Singapore,

⁵Agency for Science, Technology and Research (A*STAR), Institute for High Performance Computing, Singapore, Singapore,

⁶National University of Singapore, School of Computing, Singapore, Singapore,

⁷Nanyang Technological University, Lee Kong Chian School of Medicine, Singapore, Singapore,

⁸National University of Singapore, Cancer Science Institute of Singapore, Singapore, Singapore

Keywords: Spatial Omics, Digital Pathology, Artificial Intelligence

Background

Self-supervised foundation models for computational histopathology have shown great promise for predicting clinical phenotypes from routine hematoxylin and eosin (H&E)-stained images¹⁻³. However, their application in discovery, for example identifying novel biomarkers or therapeutic targets, remains limited. This is because while H&E images contain rich morphological information, extracting molecular insights from them remains a major challenge.

Spatial omics (SO) is a novel technology that enables high-throughput measurement of RNAs and proteins in intact tissues, capturing hundreds to thousands of molecular features while preserving spatial context⁴. In this way, SO provides a molecular map that can guide discovery efforts. We hypothesize that SO measurements offer a rich source of auxiliary information for improving model performance, generalizability and interpretability.

Methodology

We designed a general modeling framework that predicts aggregated gene expression in 224x224 pixel tiles using tile-level features extracted via pre-trained pathology foundation models. We evaluated three training strategies for incorporating SO information: direct fine-tuning, contrastive embedding and linear/K-nearest neighbor probing. Models were trained and evaluated using a 10X Xenium colorectal cancer (CRC) dataset with paired H&E and SO data (29 slides, 19 patients) under five-fold cross-validation. Generalizability was further assessed using independent datasets, including public 10X Xenium CRC datasets, The Cancer Genome Atlas Adenocarcinoma cohort (TCGA COAD-READ), and the METU-CCTGS CRC Tumor Grade Segmentation Benchmark⁵.

Results and Discussion

We demonstrate that tile-level models can accurately predict spatial molecular profiles from H&E alone. Direct fine-tuning outperformed contrastive embedding and linear/KNN probing, achieving a mean within-slide Spearman correlation of 0.624 ± 0.035 across 10 marker

genes, and 0.455 ± 0.025 across the full set of 329 genes. Tile-level features extracted using pathology foundation models consistently outperformed naïve image features. Data titration experiments indicated that model performance continues to improve with additional training data. Performance on individual genes correlated with mean gene expression, suggesting that sensitivity of current SO assays is another limiting factor.

On independent datasets, fine-tuned models generalized well, achieving comparable performance on public 10X Xenium CRC datasets, and showing good concordance with protein immunohistochemistry (IHC) staining in a clinical cohort. Interestingly, survival predictions based on model-predicted gene expression outperformed those using actual bulk RNA-sequencing in TCGA COAD-READ, while providing interpretable molecular features. Model predictions also enabled *de novo* domain segmentation, correctly identifying normal and tumor regions in METU-CCTGS without requiring labor-intensive pathology annotations for training.

Conclusion

Bridging the gap between histopathology imaging and molecular data offers a powerful approach for improving digital pathology workflows and accelerating biomarker discovery.

References

- [1] Vorontsov, E. et al. A foundation model for clinical-grade computational pathology and rare cancers detection. *Nat. Med.* 30, 2924–2935 (2024).
- [2] Xu, H. et al. A whole-slide foundation model for digital pathology from real-world data. *Nature* 630, 181–188 (2024).
- [3] Chen, R. J. et al. Towards a general-purpose foundation model for computational pathology. *Nat. Med.* 30, 850–862 (2024).
- [4] Marx, V. Method of the Year: spatially resolved transcriptomics. *Nat. Methods* 18, 9–14 (2021).
- [5] Arslan, D. et al. Colorectal cancer tumor grade segmentation: A new dataset and baseline results. *Heliyon* 11, e42467 (2025).