Systematic Benchmark of H&E-to-Transcriptomics Prediction in Skin Tissue

R Shaikhutdinova, S Gansberger, J Staller, N Singh, I Oyarzun, M Simon, B Sterniczky, P Tschandl, J Griss

Department of Dermatology, Medical University of Vienna

Objective

H&E stained tissue sections are widely used in clinical pathology due to their ability to provide fast, reliable, and inexpensive assessment of tissue structure and cellular organization[1]. Recent advances in machine learning have substantially expanded their utility, enabling prediction of spatial omics directly from H&E images[2]. Yet independent benchmark assessing whether published state-of-the-art methods can accurately predict single-cell gene expression from H&E images is still missing.

Methods

We benchmarked three published H&E-to-expression methods (GHIST[3], Pixel2Gene[4], SpatialEx[5]) against ridge-regression baselines on the same image-encoder embeddings (16 models total), across three skin datasets with matched H&E and Xenium spatial transcriptomics. Models were trained on one section and tested on a held-out section from a different individual. Predictions were evaluated per gene using Pearson correlation coefficient (PCC), Spearman correlation coefficient (SCC), and Structural Similarity Index Measure (SSIM), and by whether cell-level annotations were preserved.

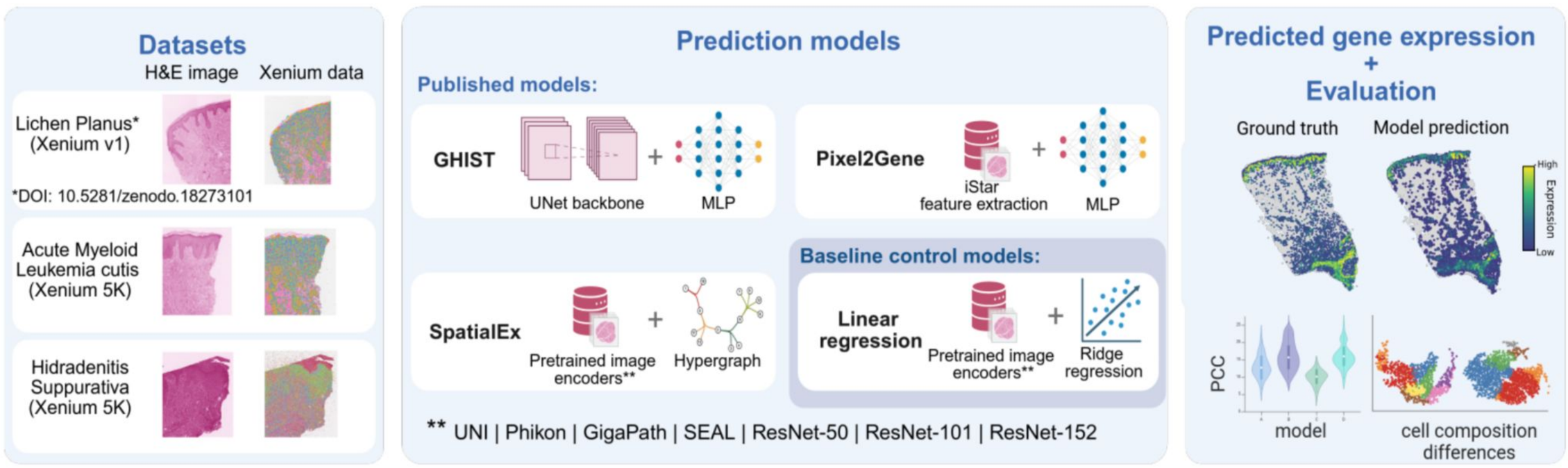

Figure 1. Graphical summary of the study's workflow.

Results

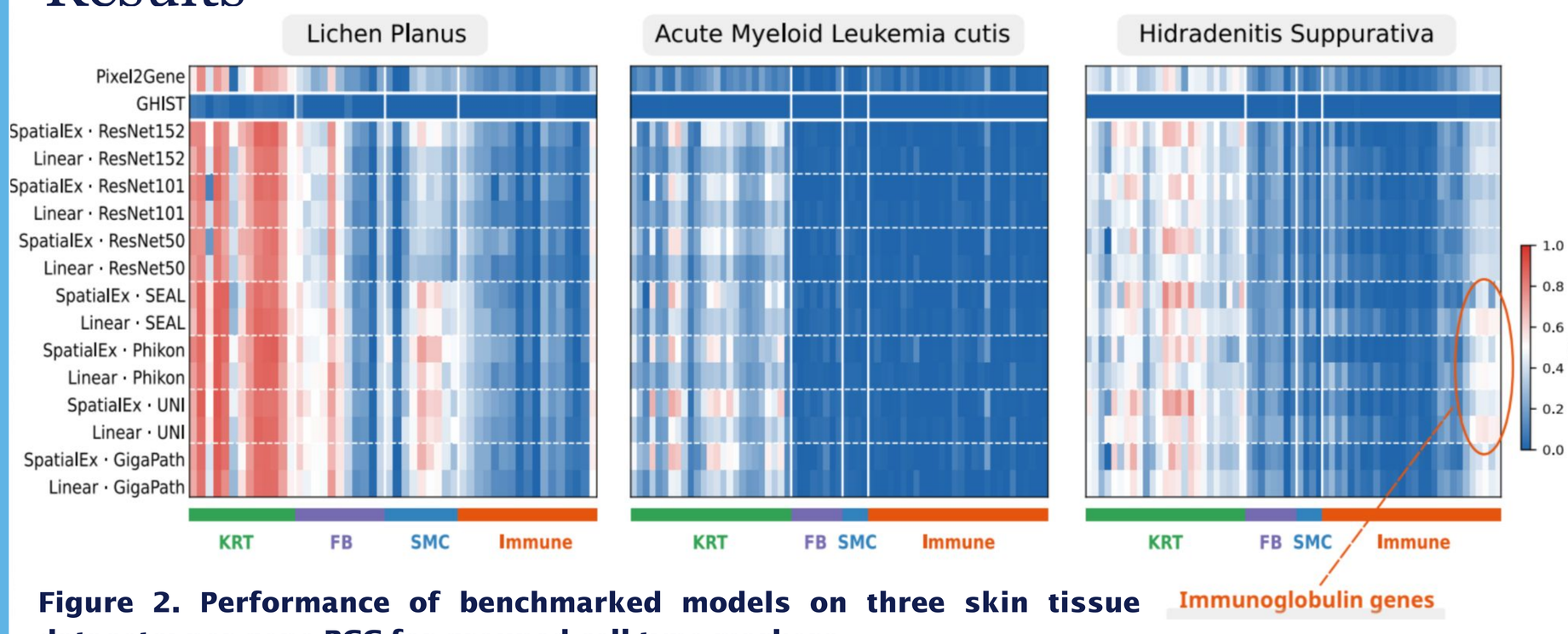

Figure 2. Performance of benchmarked models on three skin tissue datasets: per-gene PCC for grouped cell-type markers.

Most genes are poorly predicted; reliable predictions are keratinocyte-specific. Across all models, the majority of genes showed PCC values near zero. The few well predicted genes were almost exclusively keratinocyte-associated, reflecting their morphologically distinct location in the epidermis. Notably, simple ridge regression on image-encoder embeddings performed as well as the more complex published models.

Predicted expression fails to preserve cell-type identity. Upon clustering and annotation based on predicted expression values, only keratinocytes formed distinct, correctly labeled clusters, while immune, fibroblast, and other dermal populations collapsed into indistinct groups, indicating that current methods cannot reliably resolve cell-level structure beyond the epidermis.

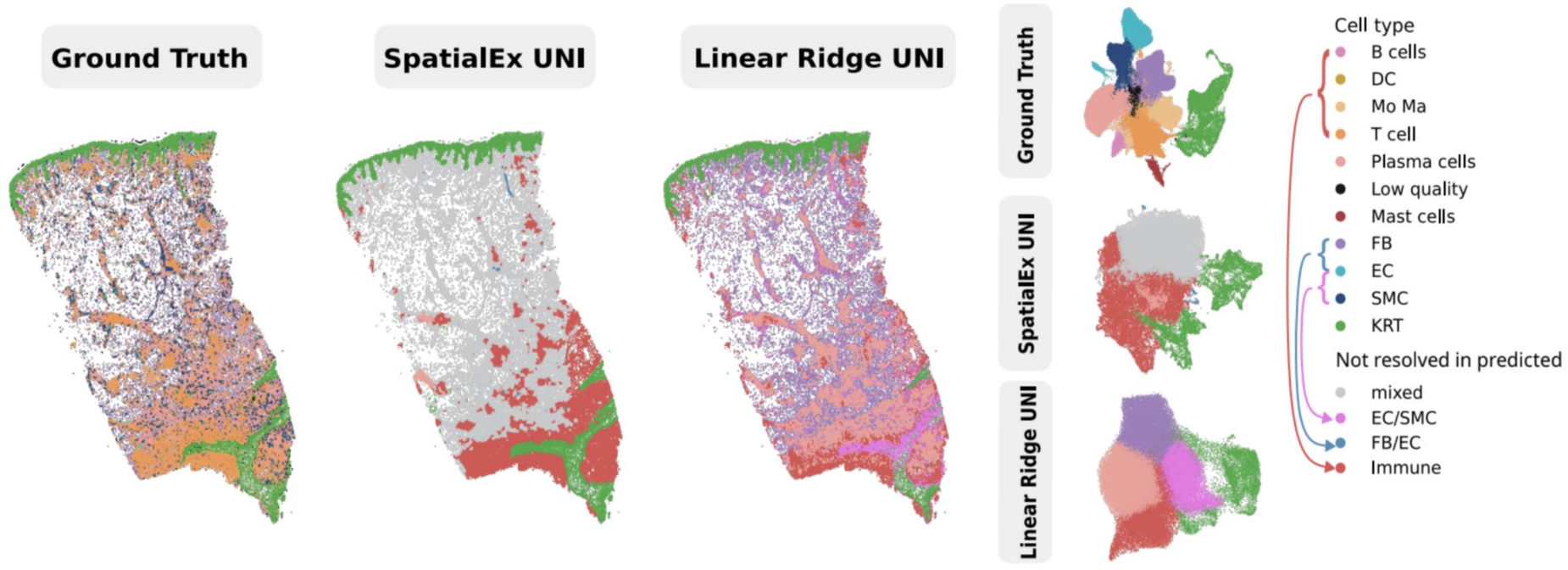

Figure 3. Cell-type structure in predicted gene expression for the two best-performing models, Hidradenitis Suppurativa dataset.

Conclusion

- We provide an independent benchmark of state-of-the-art H&E-based single-cell gene expression prediction methods in human skin.
- Current methods accurately predict only a limited subset of genes, predominantly those associated with keratinocytes.
- Complex deep learning architectures provide little improvement over simple linear models built on pretrained image embeddings.
- Biological cell-type structure and spatial organization are not reliably preserved in predicted expression profiles.

References

1. Fischer AH, Jacobson KA, Rose J et al. Hematoxylin and eosin staining of tissue and cell sections. CSH Protoc 2008;2008:pdb.prot4986. https://doi.org/10.1101/pdb.prot4986.
2. Hao N, Yang X, Yan B et al. Histopathology-centered Computational Evolution of Spatial Omics: Integration, Mapping, and Foundation Models, arXiv:2601.07826. Preprint, arXiv, 12 Jan. 2026. https://doi.org/10.48550/arXiv.2601.07826.
3. Fu X, Cao Y, Bian B et al. Spatial gene expression at single-cell resolution from histology using deep learning with GHIST. Nat Methods 2025 2025;22(9):1–11. https://doi.org/10.1038/s41592-025-02795-z.
4. Yao S, Schroeder A, Jiang S et al. Pixel2Gene enables histology-guided reconstruction and prediction of spatial gene expression. Preprint, bioRxiv, 23 Feb. 2026, 2026.02.21.707168. https://doi.org/10.64898/2026.02.21.707168.
5. Liu Y, Wang C, Wang Z et al. High-parameter spatial multi-omics through histology-anchored integration. Nat Methods 2026;23(2):373–86. https://doi.org/10.1038/s41592-025-02926-6.