DeepSpot: Leveraging Spatial Context for Enhanced Spatial Transcriptomics Prediction from H&E Images

Kalin Nonchev[1, 2, 3], Sebastian Dawo[3], Karina Silina[3], Holger Moch[4], Sonali Andani[1,2,5], Tumor Profiler Consortium, Viktor H Koelzer[4,5,6†], Gunnar Rätsch[1,2,7,8†]

[1]Department of Computer Science, ETH Zurich, Zurich Switzerland.
[2]Swiss Institute of Bioinformatics, Zurich Switzerland.
[3]Institute of Pharmaceutical Sciences, Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich Switzerland.
[4]Department of Pathology and Molecular Pathology, University Hospital Zurich, University of Zürich, Zurich Switzerland.
[5]Computational and Translational Pathology Group, Department of Biomedical Engineering, University of Basel, Basel Switzerland.
[6]Institute of Medical Genetics and Pathology, University Hospital Basel, Basel Switzerland.
[7]AI Center, ETH Zurich, Zurich Switzerland.
[8]Medical Informatics Unit, University Hospital Zurich, Zurich Switzerland.

Abstract

Spatial transcriptomics technology remains resource-intensive and unlikely to be routinely adopted for patient care soon. This hinders the development of novel precision medicine solutions and, more importantly, limits the translation of research findings to patient treatment. Here, we present DeepSpot, a deep-set neural network that leverages recent foundation models in pathology and spatial multi-level tissue context to effectively predict spatial transcriptomics from H&E images. DeepSpot substantially improved gene correlations across multiple datasets from patients with metastatic melanoma, kidney, lung, or colon cancers as compared to previous state-of-the-art. Using DeepSpot, we generated 1 792 TCGA spatial transcriptomics samples (37 million spots) of the melanoma and renal cell cancer cohorts. We anticipate this to be a valuable resource for biological discovery and a benchmark for evaluating spatial transcriptomics models. We hope that DeepSpot and this dataset will stimulate further advancements in computational spatial transcriptomics analysis.

Keywords: machine learning, pathology, histology, spatial transcriptomics prediction, deep sets

†Corresponding authors. E-mails: viktor.koelzer@usz.ch; gunnar.raetsch@inf.ethz.ch;

1 Introduction

Spatial transcriptomics provides valuable insights into tissue-specific properties both in normal physiology [1–3] and in disease progression and treatment [4–8]. However, the technology is limited to sequencing small tissue regions (e.g., 6.5x6.5mm capture area for Visium, 10x Genomics™) and it is both resource-intensive and error-prone, making its routine adaption in clinical setting unlikely in the near future [9, 10]. This hinders the development of new precision medicine solutions and, more importantly, limits the translation of basic research findings to patient diagnosis and treatment. Therefore, robust and cost-effective methods are urgently needed to enable simple, affordable, and reliable analysis of spatially resolved gene expression data in routine clinical workflows and retrospective cohorts.

Recent advances in deep learning demonstrated that high-resolution hematoxylin and eosin (H&E)-stained slides can be used as input for computational models to efficiently predict bulk RNA expression [11–13]. Building on this, the increased availability of spatial transcriptomics data [5, 14–16] enabled the development of models to predict spatial transcriptomics from H&E images. However, accurately predicting spatial transcriptomics profiles remains challenging. Existing frameworks either struggle to effectively leverage morphological details and spatial context - especially those adapted from bulk RNA predictions - are computationally expensive or are limited to predicting only a small number of genes. For example, ST-Net uses a convolutional neural network (DenseNet-121 [17], pre-trained on ImageNet [18]) followed by a fully connected linear layer to predict the expression of 250 genes [19]. While Jaume and Doucet et al. [14] utilized recent pathology foundation models (e.g., UNI [20], Phikon [21]), pre-trained on extensive H&E datasets, to extract tile spot features, they focused on predicting only 50 genes using either ridge regression or random forest. Other methods, such as BLEEP [22], use contrastive learning to create a low-dimensional joint embedding and require access to the full training data during inference to perform k-nearest neighbor mapping. Furthermore, alternative methods employing vision transformers propose compressing the H&E image as a sequence of tiles and integrating their spatial location by encoding absolute pixel coordinates through positional embeddings to predict ∼750 genes [23–25]. However, the gigapixel resolution of pathology slides makes full-image processing impractical, requiring tile subsampling and potentially leading to the loss of important details. By contrast, the spatial tissue context, proven effective for representing spatial transcriptomics data [5], remains underutilized for predicting spatial transcriptomics from H&E images. Therefore, a new and reliable method is needed to overcome the aforementioned challenges, advance spatial transcriptomics prediction and enable its effective integration into clinical practice.

To address this, we developed DeepSpot, a novel deep-learning model that utilizes recent pathology foundation models and spatial multi-level tissue context to effectively predict spatial transcriptomics from H&E images. We compared its performance against previous state-of-the-art methods on multiple datasets sequenced using Visium from 10x Genomics™: 24 samples with kidney cancer, 18 with metastatic melanoma, 8 with colon cancer, and 6 samples enriched with tertiary lymphoid structures (TLS) from kidney and lung cancer. We demonstrated substantially improved gene correlations which enabled downstream spatial transcriptomics analysis leading to biomarker and pathway discovery. Furthermore, we applied DeepSpot to slide images from 1 792 TCGA patients with corresponding bulk RNA-seq. This not only served as an out-of-distribution validation but also generated a large multi-modal spatial transcriptomics dataset with over 37 million spots from patients with melanoma or kidney cancer. This dataset represents a unique resource that significantly enriches the available spatial transcriptomics data for TCGA samples, providing valuable insights into the molecular landscapes of cancer tissues (Figure 1).

2 Results

2.1 DeepSpot leverages pathology foundation models and spatial tissue context

We introduce DeepSpot, a novel deep-learning model that leverages recent foundation models in pathology to effectively predict spatial transcriptomics from H&E images. DeepSpot employs a deep-set neural network [26] to model transcriptomic spots as bags of sub-spots and integrates

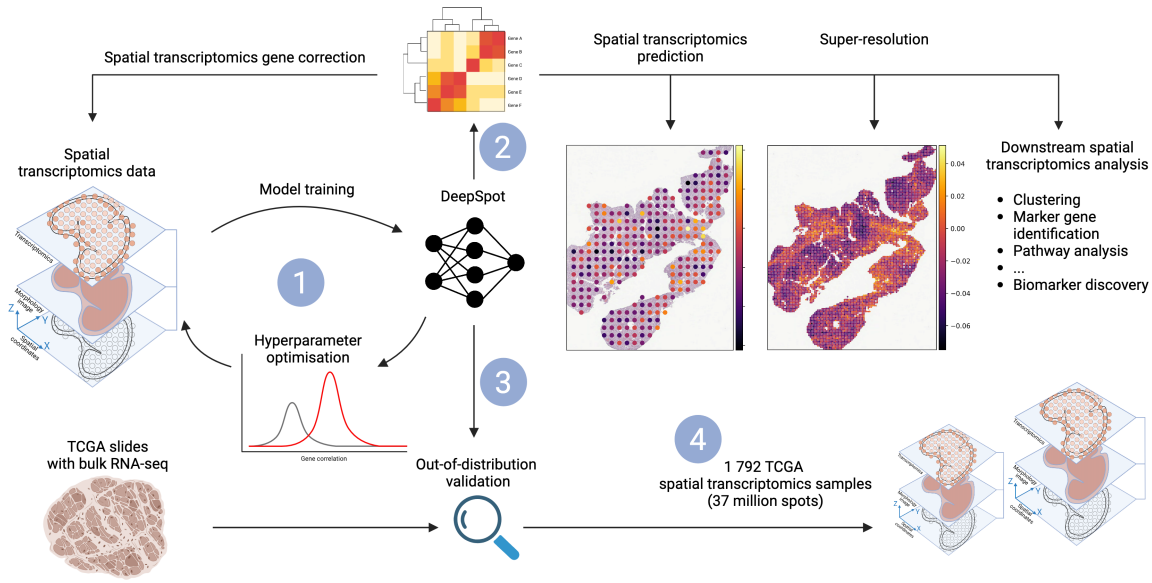

Fig. 1 Graphical summary.
DeepSpot predicts spatial transcriptomics from H&E images by leveraging recent foundation models in pathology and spatial multi-level tissue context. **1**: DeepSpot is trained to predict 5 000 genes, with hyper-parameters optimized using cross-validation. **2**: DeepSpot can be used for de novo spatial transcriptomics prediction or for correcting existing spatial transcriptomics data. **3**: Validation involves nested leave-one-out patient cross-validation and out-of-distribution testing. We predicted spatial transcriptomics from TCGA slide images, aggregated the data into pseudo-bulk RNA profiles, and compared them with the available ground truth bulk RNA-seq. **4**: DeepSpot generated 1 792 TCGA spatial transcriptomics samples with over 37 million spots from melanoma or kidney cancer patients, enriching the available spatial transcriptomics data for TCGA samples and providing valuable insights into the molecular landscapes of cancer tissues.

multi-level tissue details along with spatial neighborhood morphology. This integration, supported by the robust foundation from pre-trained H&E models, significantly enhances the accuracy and granularity of gene predictions from H&E images at minimal extra cost (Figure 2, S1). The model is trained to predict the 5 000 most variable genes in a multiregression setting using spatial transcriptomics data. First, the H&E slides are split into tiles each corresponding to a transcriptomic spot. For each spot, we create a bag of sub-spots by splitting it into non-overlapping sub-tiles to capture local tissue morphology and a bag of neighboring spots to capture the global tissue environment. A pre-trained pathology foundation model (e.g., UNI [20], Phikon [21], H-optimus-0 [27]) is then used to extract tile feature representations. These features input DeepSpot's ϕ_{spot} module, whose learned representations are concatenated and fed into the gene head predictor, ρ_{gene}, to predict the spatial transcriptomics profiles. Ultimately, the predicted digital spatial transcriptomics can be used for various downstream tasks, including spot phenotyping, gene expression analysis, and pathway and biomarker discovery.

The motivation for modeling transcriptomic spots as bags of sub-spots is inspired by the fact that each spot contains between 1-10 cells (10x Genomics™, Visium). This approach captures finer details within the target spot effectively overcoming the resolution limitations of the sequencing technology while also learning the contributions of individual sub-spots. Additionally, DeepSpot integrates the global tissue environment by pooling neighboring spots and jointly learning the tissue landscape. Building on the strong foundation of recent models for H&E images, the deep sets architecture is particularly well-suited for representing spatial transcriptomics data due to its permutation invariance and permutation equivariance properties [26]. These characteristics make it a robust framework for modeling spatial transcriptomics, as it effectively captures the relationships between spots by focusing on biologically meaningful spatial patterns rather than the exact order of the spots. Furthermore, the shared weights of the ϕ_{spot} module are optimized simultaneously across multiple tasks, leading to better generalization and more efficient learning in the limited data context of spatial transcriptomics data (Figure 6A).

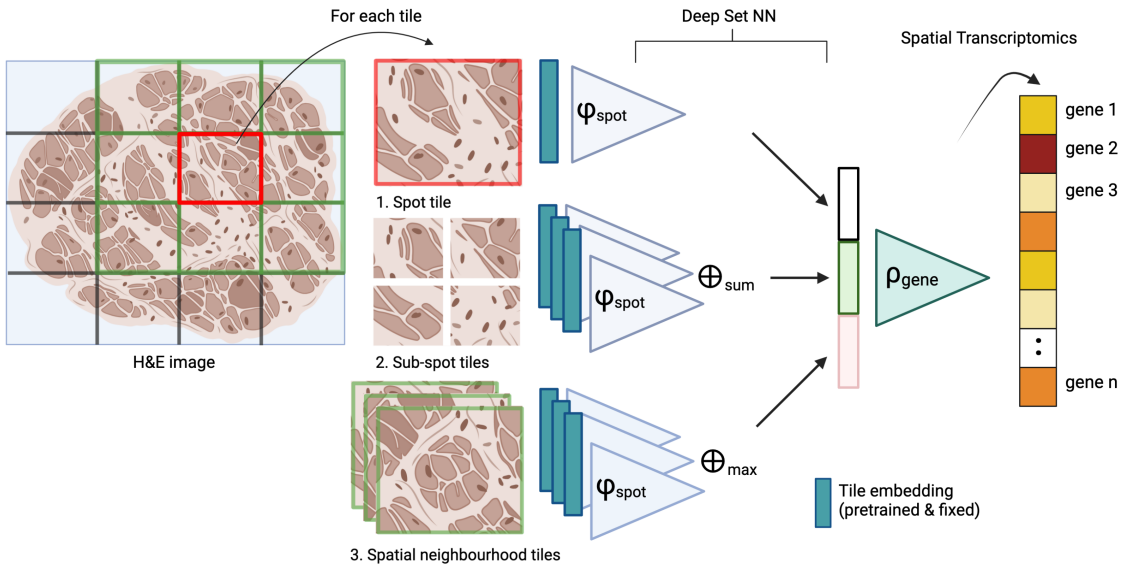

Fig. 2 DeepSpot leverages pathology foundation models and spatial tissue context. Workflow of DeepSpot: H&E slides are first divided into tiles, each corresponding to a spot. For each spot, we create a bag of sub-spots by dividing it into sub-tiles that capture the local morphology, and a bag of neighboring spots to represent the global context. A pre-trained pathology foundation model extracts tile features, which are input to the model. The concatenated representations are then fed into the gene head predictor, ρ_{gene}, to predict spatial gene expression.

2.2 DeepSpot improves gene expression prediction

We benchmarked DeepSpot's performance on multiple spatial transcriptomics datasets sequenced using Visium from 10x Genomics™ to predict the top 5 000 most variable genes as defined by Scanpy with the Seurat v3 flavor [28]. In line with the methodology of [14, 19, 22–25], we adopted Pearson correlation to measure the similarity between the predicted and the ground truth gene expression. We calculated correlations for the top 100 to 5 000 most variable genes (Figure 3A). To avoid hyperparameter overfitting, we employed nested leave-one-out patient cross-validation, with an internal cross-validation loop for hyperparameter selection. Specifically, for each fold, we trained a model on the training set, selected hyperparameters on the validation set, and evaluated performance on the test set, ensuring that each set contains distinct patients. This process was iterated over all folds, and the resulting median Pearson correlation, along with the standard error, is reported (Figure 3A).

DeepSpot consistently predicted improved spatial gene expression with higher Pearson correlation across all datasets compared to previous state-of-the-art methods (Figure 3A). For example, on the Tumor Profiler dataset [5], DeepSpot improved the Pearson correlation across the top 500 genes by 29%, increasing it from 0.24 (best competitor, MLP) to 0.31 (DeepSpot). The Tumor Profiler dataset comprises 18 tissue slices from 7 patients with metastatic melanoma. This improvement highlights the superior performance of DeepSpot in predicting spatial transcriptomics by leveraging recent foundation models in pathology and spatial multi-level tissue context. The vision transformer-based models (HisToGene, Hist2ST, and THItoGene) underperformed across all datasets, likely because their architecture depends on learning pathology features from scratch. In contrast, leveraging transfer learning with prior histopathology knowledge from pathology foundation models trained on extensive collections of clinical slides could offer a significant advantage (Figure S1).

To qualitatively illustrate the predicted spatial transcriptomics, we compare the expression for *SOX10*, a known melanoma marker gene [29, 30] for slice MELIPIT-1-2 with available ground truth pathology annotations (Figure 3B, C). The gene expression predictions from BLEEP ($\rho = 0.57$) and ST-Net ($\rho = 0.61$) exhibit low spatial resolution, with areas of high expression intensity concentrated at tumor boundaries. In contrast, linear regression ($\rho = 0.64$) and MLP ($\rho = 0.64$), utilizing pathology foundation model features, show greater contrast between normal and tumor tissues suggesting they are more effective at distinguishing between these conditions. However, the

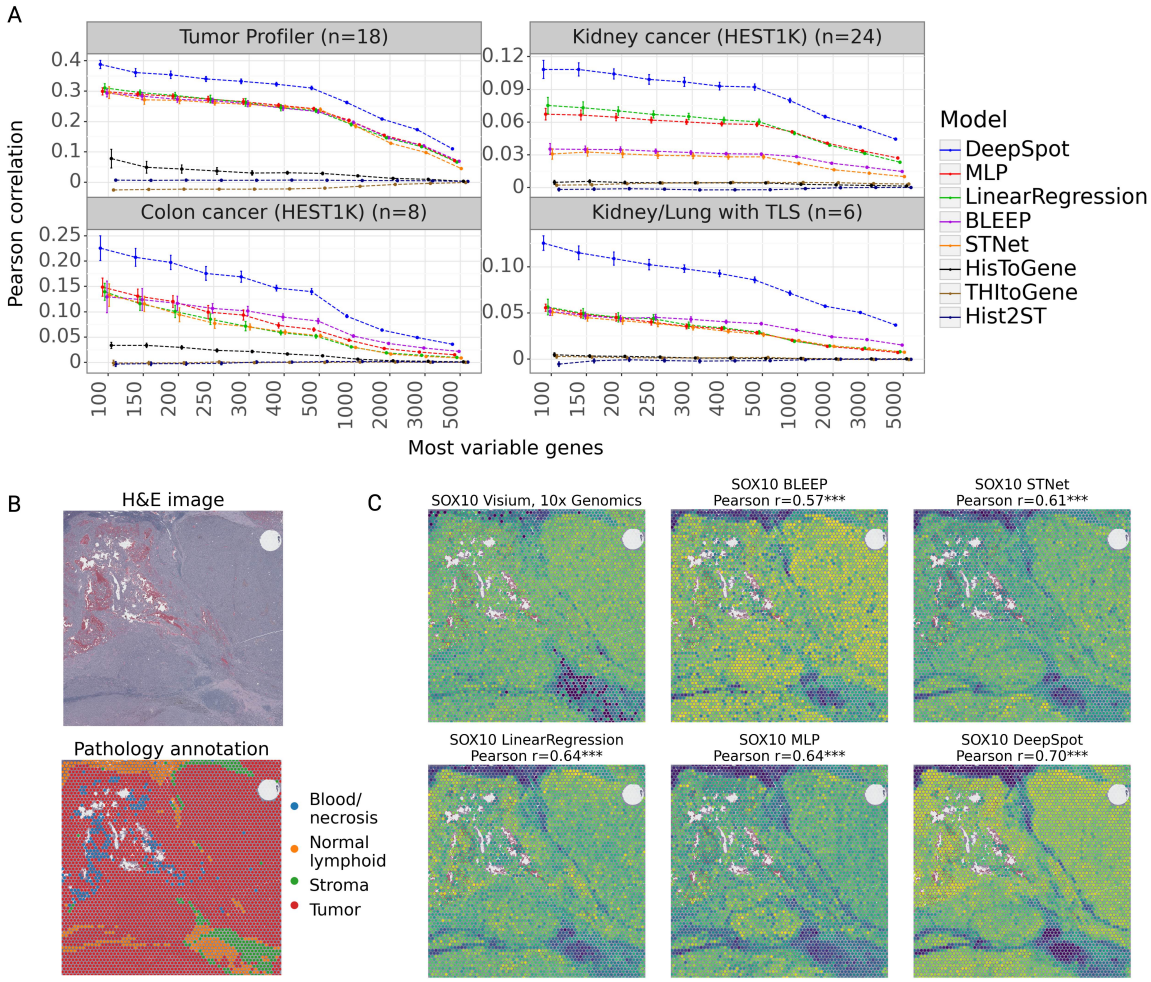

Fig. 3 DeepSpot improves gene expression prediction.
A: Benchmark of DeepSpot and previous state-of-the-art methods on four spatial transcriptomics datasets generated using Visium from 10x Genomics™. The y-axis represents the Pearson correlation between the predicted and ground truth gene expression. The x-axis represents Pearson correlation computed on the top N highly variable genes as defined by Scanpy with the Seurat v3 flavor [28]. Models are ordered based on their relative rank across datasets. **B**: H&E image and pathology annotation for slice MELIPIT-1-2 from the Tumor Profiler dataset [5] and **C**: Comparison of gene expression for *SOX10* (melanoma marker [29, 30]). Stars denote the levels of statistical significance, with *** indicating p<0.001 as determined by bootstrapping.

gene expression within the tumor structures appears noisy and random, lacking a clear pattern and offering less detailed information about the tumor microenvironment. Notably, DeepSpot achieved the highest gene correlation to the ground truth ($\rho = 0.70$), effectively distinguishing between normal and tumor tissues. DeepSpot produced a more consistent gene expression pattern within the tumor, with the highest levels of expression concentrated in the tumor's core (Figure 3C).

2.3 DeepSpot enables digital spatial transcriptomics analysis

To demonstrate the utility of the digital spatial transcriptomics data, we use the DeepSpot predicted gene expression for slice MELIPIT-1-2 from the Tumor Profiler dataset to perform a common spatial transcriptomics analysis workflow [31]. First, we apply dimensionality reduction to visualize the data with randomly sampled morphology spots (Figure 4A). We observe an aggregation of tumor spots on the right side, with a gradual transition towards spots of necrosis, stroma, and normal lymphoid tissue on the left. Tumor spots appear significantly larger with enlarged nuclei; stroma cells have an elongated morphology and are noticeably more scattered; normal lymphoid cells are generally smaller and denser in line with the observations in [5].

Next, we performed a marker gene analysis using the Wilcoxon test. Among the identified tumor marker genes, several well-known melanoma-related genes (e.g., *S100B* [32], *TYR* [33], *CSPG4* [34, 35], *PMEL* [36] and *MLPH* [37]) were found (Figure 4B). They are involved in melanoma progression and associated with increased invasiveness of melanoma cells [33, 38–40]. To further validate our findings, we performed enrichment analysis using the top 100 tumor marker genes in conjunction with data from the Cancer Cell Line Encyclopedia [41] (Figure 4C). All of the top 15 cell lines were related to skin melanoma, indicating a strong association with tumor biology relevant to these malignancies [42–46]. Additionally, the significant enrichment of specific pathways involved in cell proliferation and survival suggests potential therapeutic targets for further investigation.

Finally, we conducted a pathway analysis (Figure 4D) using decoupleR [47] which revealed increased activity of the MAPK pathway in tumor tissues. This pathway is known for promoting processes such as cell proliferation, invasion, metastasis, migration, survival, and angiogenesis [48–50]. Additionally, consistent with the findings of [5], we observed predominant hypoxia signatures in regions of necrosis and hemorrhage, while JAK-STAT inflammatory signaling was dominant within the cancer microenvironment. Overall, these transcriptomics analyses highlight the accurate gene expression predictions made by DeepSpot and reinforce their interpretation in the context of underlying biological processes.

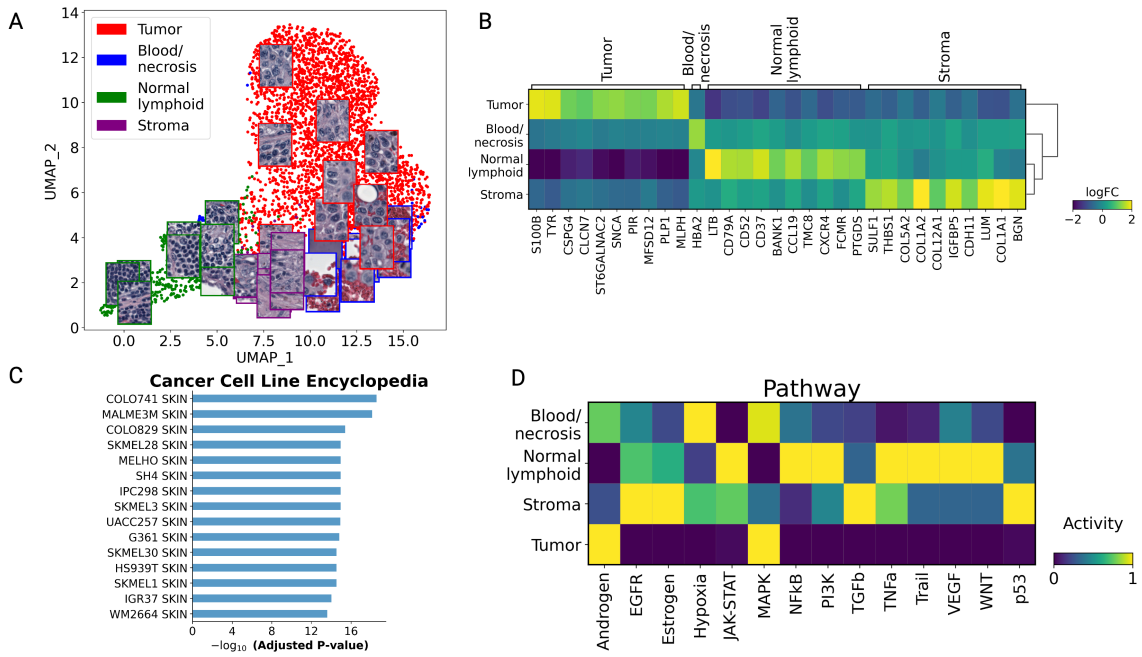

Fig. 4 DeepSpot enables digital spatial transcriptomics analysis.
A: UMAP plot of spatial transcriptomics predictions from DeepSpot for sample MELIPIT-1-2 from the Tumor Profiler dataset with sampled spot images. **B**: Matrix plot of the log fold change (logFC) of the top 10 marker genes per group derived from predicted spatial transcriptomics for the same sample. **C**: Gene set enrichment analysis using the top 100 tumor marker genes and the data from the Cancer Cell Line Encyclopedia [41]. **D**: Pathway analysis based on the predicted spatial transcriptomics using decoupleR [47].

2.4 DeepSpot overcomes spatial transcriptomics gene sensitivity

Higher resolution in spatial transcriptomics such as that produced by Visium from 10x Genomics often leads to lower gene sensitivity and spot contamination - known issues that result in poor downstream performance and misleading results [51]. These challenges are further compounded by the reduced RNA quality in archival samples. To demonstrate how DeepSpot addresses these issues, we present tissue slice KC2 (Figure 5A) excluded from the experiments due to low transcriptomics quality (Figure S2). Specifically, regions annotated as tertiary lymphoid structures (TLS) did not show expression of known TLS marker genes (Figure 5B, red boxes) [52, 53], and unexpectedly, these genes

were detected in other locations. This hold-out sample would typically have been discarded from downstream analysis. However, DeepSpot, trained on kidney and lung samples with TLS, successfully recovered this low-quality sample by effectively transferring tissue morphology patterns associated with TLS. As a result, DeepSpot accurately predicted the gene expression at positions corresponding to these morphological features (Figure 5C). Moreover, DeepSpot's architecture enables gene expression prediction at super-resolution by focusing on a single sub-spot at the center of the spot, further augmenting patient data with fine-grained details at minimal additional cost (Figure 5D). For example, we predicted spatial transcriptomics for the same patient by reducing spot distance (4x) and diameter (3x) in pixels, which enhanced data resolution and improved spatial accuracy. The super-resolution allowed us to pinpoint the exact locations in the H&E image corresponding to the highest gene expression (e.g., *LTB*, Figure 5D). Specifically, one can observe that the highest *LTB* expression occurs in the center of the TLS formations potentially indicating the TLS germinal centers.

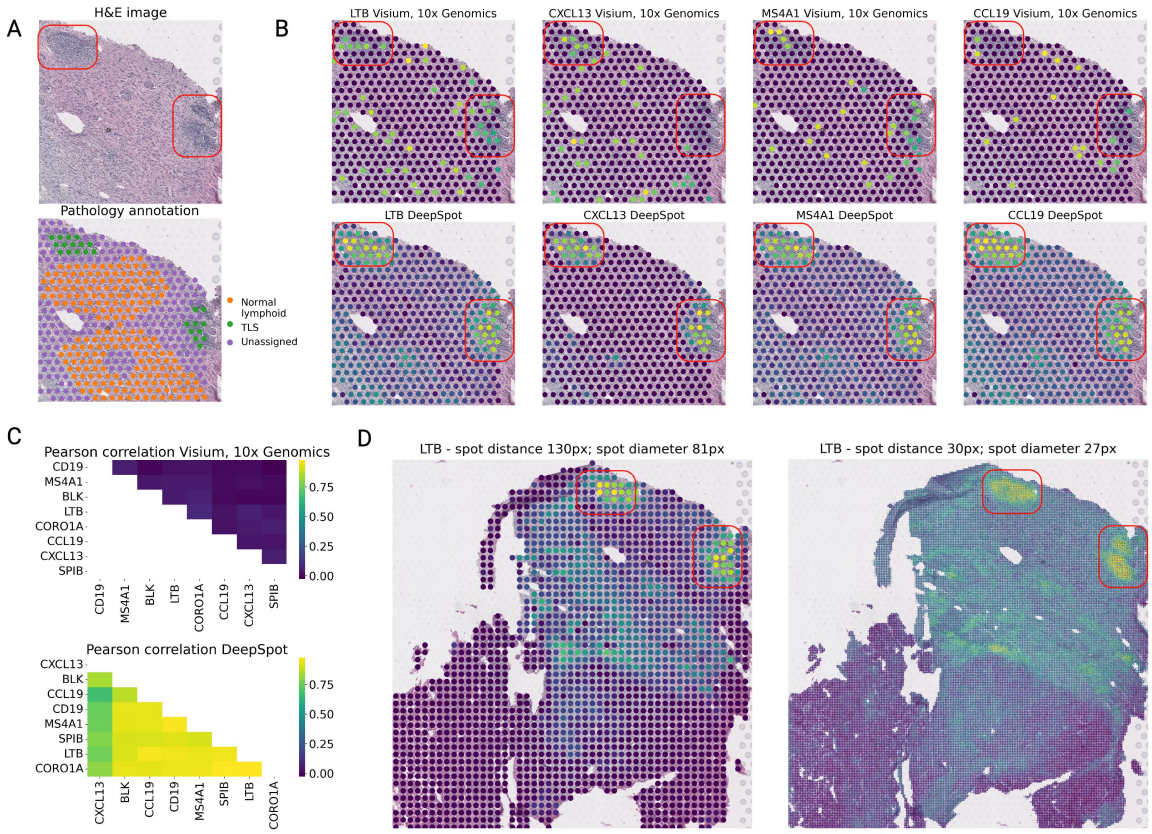

Fig. 5 DeepSpot overcomes spatial transcriptomics gene sensitivity.
A: H&E image and pathology annotation for slice KC2 from the Lung/Kidney with TLS dataset, zoomed into TLS formations. **B**: Comparison of gene expression for *LTB*, *CXCL13*, *MS4A1M*, and *CCL19* (known TLS markers) between ground truth data from 10x Genomics™, Visium and predictions from DeepSpot. **C**: Pearson correlation of gene expression for *LTB*, *CXCL13*, *MS4A1M*, *CCL19*, *CORO1A*, *CD19*, *SPIB*, and *BLK* across all spots in slice KC2. **D**: Spatial transcriptomics predictions for gene *LTB* with original spot distances of 130 px and spot diameter of 81 px and at super-resolution with spot distances of 30 px and spot diameter of 27 px in slice KC2.

2.5 DeepSpot predicts digital spatial transcriptomics for 1 792 TCGA slides

To assess out-of-distribution performance, we validated DeepSpot, trained on melanoma or kidney cancer spatial transcriptomics, using fresh frozen (FF) or formalin-fixed paraffin-embedded (FFPE) slides from 1 792 TCGA samples with corresponding bulk RNA-seq (Figure 6A). In summary, we predicted spatial gene expression for each slide in TCGA with melanoma (SKCM) or kidney cancer

(KIRC) and then aggregated the expression to generate pseudo-bulk RNA profiles. We compared this pseudo-bulk RNA to the available ground truth bulk RNA-seq (Figure 6B). Notably, DeepSpot accurately predicted gene expression and outperformed previous state-of-the-art models also in out-of-distribution setting. In melanoma, the DeepSpot pseudo-bulk RNA profile derived from lower-quality FF slides correlated more closely with the paired bulk RNA from the same tissue piece compared to the larger FFPE slide. Although DeepSpot was trained on FFPE melanoma slides, its design, leveraging a pathology foundation model, enabled effective generalization to FF slides, highlighting its potential for reliable gene expression predictions (Figure 6B). Conversely, DeepSpot trained on lower-quality FF slides from the kidney dataset underscored the importance of high-quality training data for learning robust pathology representations. Although the model outperformed previous state-of-the-art models, the higher correlation of FFPE pseudo-bulk over FF pseudo-bulk suggests limitations imposed by the quality of the FF slides in training data. Moreover, DeepSpot's predictions were more accurate for genes with greater variability in the training data. This highlights the need for larger spatial transcriptomics datasets that encompass a diverse range of genes to train more precise models.

To further demonstrate the quality and robustness of our predicted spatial transcriptomics data, we trained downstream models on the pseudo-bulk RNA for several tasks, including survival analysis (Figure 6C) and tumor type classification (Figure 6D). Notably, the gene expression generated from DeepSpot outperformed previous state-of-the-art methods in survival analysis and served as a better predictor than the ground truth bulk RNA-seq (Figure 6C, S5). For instance, in the SKCM dataset, DeepSpot improved the concordance index by 2%, increasing it from 0.677±0.001 (best competitor, STNet FFPE) to 0.691±0.001 (DeepSpot FFPE). When compared to the ground truth bulk RNA-seq (0.600±0.001), this represents a 13% improvement. Moreover, DeepSpot's gene expression improves tumor type classification in SKCM, increasing the F1 score by 4%, from 0.818±0.001 (best competitor, STNet FFPE) to 0.852±0.001 (DeepSpot FFPE), slightly surpassing the ground truth bulk RNA-seq - 0.846±0.001 (Figure 6D, S6). Finally, we compared the DeepSpot-predicted gene expression of known TLS marker genes [52, 53] in FFPE KIRC samples with available manual annotations [54] (Normal tissue or TLS, Figure 6F). Notably, TLS marker genes were significantly overexpressed in spots annotated as TLS compared to those in normal tissue (Figure 6G), underscoring the robustness of DeepSpot-predicted spot-level gene expression in providing enhanced contextual insights.

These experiments resulted in a large multi-modal spatial transcriptomics dataset with over 37 million spots from 1 792 samples with melanoma or kidney cancer. It represents a unique resource that significantly enriches the available spatial transcriptomics data for TCGA samples, providing unique insights into the molecular landscapes of cancer tissues. To further illustrate the richness and utility of our dataset compared to the limited information from TCGA bulk RNA-seq, we compared the expression of the melanoma marker *SOX10* in both bulk RNA data (Figure S3) and predicted spatial transcriptomics data (Figure S4) for SKCM. *SOX10* expression is notably higher on the left side, indicating the presence of tumor spots, and gradually decreases towards the right. Additionally, we provide visual examples of predicted spots expressing *SOX10* (melanoma), *CD37* (normal lymphoid), and *COL1A1* (stroma) (Figure 6E), which further validates the correlation between the morphological data and predicted transcriptomics. DeepSpot demonstrated to be a robust method for predicting spatial transcriptomics from H&E images, with the ability to generalize effectively to previously unseen and out-of-distribution pathology slides.

3 Discussion

In this work, we propose DeepSpot, a method that leverages pathology foundation models and spatial tissue context to effectively predict spatial transcriptomics from routine histology images. Our results consistently show improved spatial gene predictions over previous state-of-the-art methods across multiple datasets, including melanoma, colon, kidney and lung cancers sequenced using Visium from 10x Genomics™. DeepSpot demonstrated robustness to out-of-distribution H&E images from TCGA, producing accurate gene predictions. This led to the creation of a unique multimodal spatial transcriptomics dataset with over 37 million spots from 1 792 samples with melanoma or kidney cancer. We systematically demonstrated that the gene expression obtained from DeepSpot enables de novo spatial transcriptomics analysis, including clustering, marker gene identification,

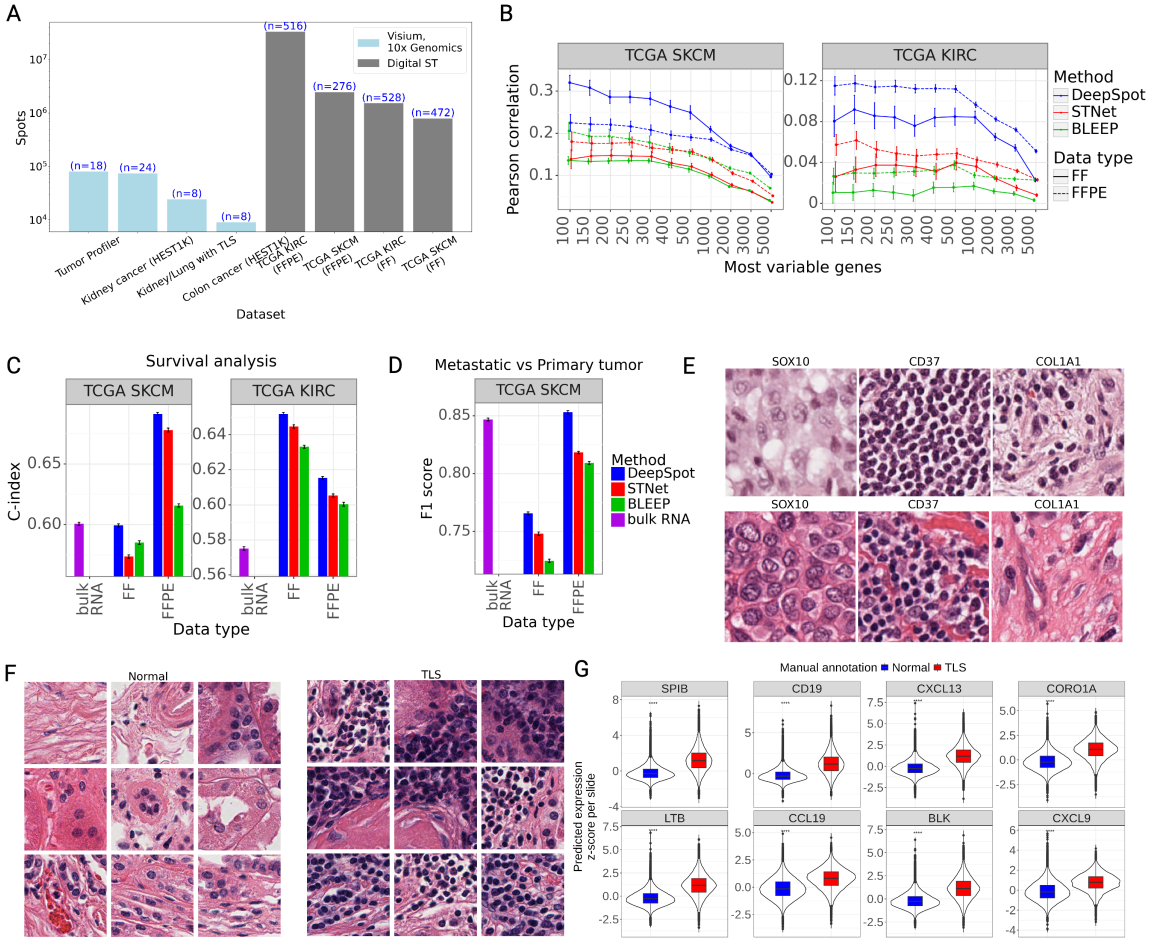

Fig. 6 DeepSpot predicts digital spatial transcriptomics from 1 792 TCGA slides.
A: Number of spots and samples per dataset. The x-axis represents the datasets, while the y-axis shows the number of spots. The number of samples is displayed above each dataset bar. **B**: Out-of-distribution performance on models trained on 10x Genomics™, Visium and gene expression predictions based on TCGA slides. The y-axis represents the Pearson correlation between the pseudo-bulk RNA and ground truth bulk RNA profile. The x-axis represents Pearson correlation computed on the top N highly variable genes as defined by Scanpy with the Seurat v3 flavor [28] on the training data. The line type indicates the data type used to generate the predictions (FF or FFPE). **C**: Survival analysis benchmark. The x-axis represents the concordance index (C-index), while the y-axis represents the data type. **D**: Tumor type classification benchmark. The x-axis represents the F1 score, while the y-axis represents the slide type. **E**: TCGA FFPE SKCM examples of spot images for *SOX10* (melanoma), *CD37* (normal lymphoid), and *COL1A1* (stroma). **F**: TCGA FFPE KIRC examples of spots manually labeled as normal tissue and TLS. **G**: The DeepSpot-predicted expression of known TLS marker genes in spots manually labeled as Normal or TLS. The y-axis represents the predicted gene expression, standardized per slide. Stars denote the levels of statistical significance, with **** indicating $p<0.0001$ as determined by the Wilcoxon rank-sum test.

and pathway discovery.

This improvement in spatial transcriptomics prediction resulted from incorporating multi-level tissue details and spatial neighborhood information along with utilizing pathology foundation models trained on extensive datasets of H&E slides. Our approach addresses the limitations of sequencing technology resolution by representing spots as bags, aggregating local spot information, and pooling global tissue context to improve the accuracy of spatial transcriptomics predictions from H&E images. This proved beneficial across diverse datasets (Figure 3A, S1). In contrast, the linear regression and MLP baseline performed comparably to the previous state-of-the-art models, ST-Net and BLEEP, which use only the target spot as input. Moreover, BLEEP inference requires complete access to the training data for k-nearest neighbor mapping in the latent space, which may become impractical due to the increasing data size and various ethical and privacy requirements in

clinical setting. Furthermore, the vision transformer models struggled to establish robust correlations between morphology and transcriptomics possibly due to their inability to capture the spatial tissue environment and their limitations in incorporating morphological features from pre-trained image models.

Several promising paths exist to enhance DeepSpot's gene expression prediction. **1)** Currently, we split each spot into a fixed number of non-overlapping sub-spots, but this approach could be improved by using a cell segmentation algorithm to detect cell locations and boundaries, enabling us to divide the spot based on identified cells. **2)** While we utilized general pathology foundation models to extract tile representations, it would be beneficial to use models specifically tailored to particular tissue types or tasks that are more relevant to spatial transcriptomics prediction (e.g., Hover-Net [55], cell nuclei segmentation and classification). Additionally, we anticipate that advancements in pathology foundation models could further enhance DeepSpot's performance. **3)** Currently, all neighboring spots are assigned the same weight; however, as the neighboring radius increases, this assumption may not hold. Therefore, it is important to investigate how to account for this variability such as by setting fixed radius parameters or learning a neighboring weight function.

In the future, DeepSpot could be applied to accurately predict spatial transcriptomics from H&E images and generate large-scale multi-modal spatial transcriptomics datasets at scale. This advancement would make spatial transcriptomics more accessible to scientists and healthcare professionals, enhancing its potential for biological discovery and accelerating clinical adoption. To foster these efforts, we have released the code for DeepSpot along with examples demonstrating its usage. Furthermore, we anticipate that the predicted TCGA spatial transcriptomics dataset, containing over 37 million spots, will serve as a valuable resource providing unique insights into the molecular landscapes of cancer tissues. It will also establish a benchmark for evaluating the performance and explainability of spatial transcriptomics models and support new model development. We hope that DeepSpot and this dataset will stimulate further advancements in computational spatial transcriptomics analysis.

4 Methods & Materials

4.1 Transcriptomics preprocessing

We normalized the transcriptomics count for each spot by total counts over all genes and then scaled it to a factor of 10 000, followed by a log1p transformation. To identify the most variable genes, we applied the *highly_variable_genes* function using Scanpy [28] with the Seurat v3 flavor, considering all spots in each dataset and using a batch key based on the tissue slice to reduce the impact of slice-specific variability. From this, we selected the top 5 000 most variable genes per dataset, which were used as the targets for prediction.

4.1.1 Spatial oversampling with AESTETIK

Due to the skewed nature of the transcriptomics data, where genes might be expressed only in a small number of spots, we perform oversampling on the training data as follows: We utilize AESTETIK [5], a recent deep-learning model for spatial transcriptomics representation learning, to jointly integrate the spatial and transcriptomics modalities and project them into a lower-dimensional space for each slide. We then perform clustering in AESTETIK's latent space using the Leiden algorithm with a resolution of 1 to obtain spot labels. Each cluster is subsequently oversampled to match the size of the largest cluster using resampling with replacement.

4.2 DeepSpot design

DeepSpot utilizes a deep-learning model based on deep-set neural networks [26]. The ϕ_{spot} module consists of a single fully connected layer with dropout regularization and ReLU activation. Its weights are shared across the three submodules—spot, sub-spot, and neighboring spots—enabling the model to perform multiple tasks simultaneously: 1) extracting features specific to each spot, 2) aggregating information from sub-spot features, and 3) maximum pooling features across the neighboring spots. The outputs from these submodules are concatenated and used as inputs to the ρ_{gene} module, which also consists of a single fully connected layer with dropout regularization and ReLU activation. The ρ_{gene} module predicts gene expression in a multiregression setting. To enhance network stability, we

employ an ensemble architecture for both ϕ_{spot} and ρ_{gene} modules using random LeCun initialization [56]. The ensemble output is obtained by averaging the predictions from an ensemble of 10 ϕ_{spot} and 10 ρ_{gene} modules. DeepSpot is implemented in Python using PyTorch and PyTorch Lightning.

4.2.1 Model input

For each spot tile X with array coordinates x_array_{target} and y_array_{target}, we compute 3x3 non-overlapping sub-spot tiles and select the closest neighboring spot tiles within a radius r from the target spot's location in the array ($x_array_{target} - r <= x_array <= x_array_{target} + r$; $y_array_{target} - r <= y_array <= y_array_{target} + r$). For each tile, we extract spot features using recent pathology foundation models (e.g., UNI [20], Phikon [21], H-optimus-0 [27]), which were pre-trained on extensive H&E datasets, and follow their recommended image preprocessing workflows. For each spot X, we construct three sets: one representing the spot itself, one representing the local tissue structure (3x3 - 9 sub-spots), and one representing the global tissue environment (neighboring spots located within a radius r).

4.2.2 Model output

We apply standard normalization per gene during training, using the mean and standard deviation computed from the training set. This process ensures that all genes are on the same scale, preventing the loss function from being dominated by a small subset of genes (Figure S1). In inference mode, we reverse the data transformation to restore the original gene ranges outlined in 4.1.

4.2.3 Training details

The model is optimized using the mean squared error loss function (Equation 1). It is trained on a single GPU with at least 12GB RAM with early stopping, using the Adam optimizer [57] with a learning rate of 1e-4, a weight decay of 1e-6 and a batch size of 1024. Computational data analysis was performed at Leonhard Med (https://sis.id.ethz.ch/services/sensitiveresearchdata/), a secure trusted research environment at ETH Zurich.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

4.3 Evaluation

We evaluate the model's performance using Pearson correlation per gene (Equation 2), employing nested leave-one-out patient cross-validation with an internal 3-fold cross-validation loop for hyper-parameter selection. Specifically, for each fold, we trained a model on the training set, selected hyperparameters on the validation set, and evaluated performance on the test set, ensuring that each set contained distinct patients. This process was repeated across all folds. We bootstrapped 10 000 times from the median Pearson correlation across the test folds and reported the resulting median Pearson correlation along with its standard error. For state-of-the-art methods, we utilized hyperparameter values recommended by the authors or discussed in their respective papers. The specific hyperparameter settings are included in the supplementary materials.

$$r = \frac{n\sum(xy) - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \tag{2}$$

4.3.1 Out-of-distribution evaluation on TCGA

Moreover, for out-of-distribution evaluation on TCGA, we generated spatial transcriptomics samples from the available TCGA image slides (FFPE and FF) and applied the pre-trained DeepSpot model (trained on the full spatial transcriptomics dataset for the specific condition) to predict gene expression. We then computed pseudo-bulk expression by averaging gene expression across all spots per slide and measured the correlation with the corresponding bulk RNA-seq data from TCGA. While the FF slides and their corresponding bulk RNA profiles came from the exact same tissue piece, the FFPE slides with larger tissue areas lacked exact bulk RNA profiles. Therefore, we matched them to a bulk RNA profile based on patient-level identification and tumor type classification. We applied

the pre-trained melanoma model to TCGA samples with skin melanoma (TCGA SKCM, n=472 FF; n=276 FFPE) and the kidney cancer to TCGA samples with kidney cancer (TCGA KIRC, n=528 FF, n=516 FFPE).

4.3.2 Ablation study

In the ablation study of DeepSpot, we follow the procedure outlined in 4.3, where we fix the hyper-parameter of interest and evaluate its effect on the Area under the Pearson gene correlation curve (Figure S1). We studied the impact of the neighbor spot radius, loss function, feature image model, normalization and oversampling.

4.4 Downstream applications

4.4.1 Predicted gene expression transformation

We enforce positive gene values by setting all values smaller than 0 to 0 and adding 1. This approach ensures compatibility with downstream tasks without introducing additional computational challenges (e.g., NaN values due to negative scores or floating point overflow) or requiring custom solutions.

4.4.2 Marker genes

For identifying marker genes, we use the *rank_genes_groups* function from Scanpy with the Wilcoxon signed-rank test. We select significant marker genes and rank them by their average log-fold change.

4.4.3 Gene set enrichment analysis

For the gene set enrichment analysis, we utilized the Python interface for EnrichR [58] in conjunction with data from the Cancer Cell Line Encyclopedia [41]. We selected the top 100 tumor marker genes based on their log-fold change.

4.4.4 Pathway analysis

For pathway analysis, we apply the multivariate linear model from the *decoupler* package [47] to assess regulatory pathway activities, using data from the *PROGENy* database [59].

4.4.5 Survival analysis on TCGA

We used the Cox proportional hazards model with an elastic net penalty [60] to conduct the survival analysis benchmark. We randomly sampled training sizes of 75, 100, and 125 patients, with the remaining patients used for evaluation. This process was repeated 1 000 times, and we reported the mean concordance index along with its standard error. When multiple samples were available per patient, we selected the larger sample based on tissue area.

4.4.6 Tumor type classification on TCGA

Following the procedure described in 4.4.5, we implemented logistic regression with an L2 penalty and evaluated the model using the F1 score (Figure S6).

4.4.7 Super-resolution gene expression prediction

To obtain super-resolution gene expression predictions, we utilize the already pre-trained DeepSpot. Rather than feeding all non-overlapping sub-spots into the model, we select only a sub-spot in the center and adjust the neighbor spot distances accordingly.

4.5 Spatial trancsriptomics data

We used spatial transcriptomics datasets generated using Visium from 10x Genomics™ with a capture area of 6.5x6.5mm and a spot diameter of 55 µm.

4.5.1 Tumor Profiler

The Tumor Profiler dataset [5] consists of 18 tissue slices from 7 patients with metastatic melanoma and includes pathology annotations with the following labels: Tumor, Normal lymphoid tissue, Blood-/Necrosis, Stroma, and Pigment. The dataset consists of FFPE sections processed using Visium technology. The H&E slices and corresponding pixel coordinates were scaled to 20x magnification.

4.5.2 Kidney cancer HEST-1K and Colon cancer HEST-1K

The kidney cancer HEST-1K [61] dataset consists of 24 tissue slices from 24 patients with kidney cancer. The colon cancer HEST-1K [62] consists of 8 tissue slices from 4 patients. Both datasets consist of fresh frozen sections processed using Visium technology and were downloaded from the HEST-1K [14]. Although the H&E slices were originally available at 40x magnification, to ensure compatibility with the TCGA slides, which are a mixture of 20x and 40x, we scaled the H&E slices and corresponding pixel coordinates to 20x magnification.

4.5.3 Kidney and Lung cancer with TLS

The kidney and lung cancer with tertiary lymphoid structures (TLS) dataset consists of 5 μm thick FFPE sections from kidney (3) and lung (5) tumors obtained from the Insitute of Pathology at the University Hospital of Zurich and mounted onto the Visium slides with the Human Probe Set v1. Samples were stained with hematoxylin and eosin and subsequently processed for sequencing following the manufacturers' recommendations. After library preparation, the samples were sequenced on an Illumina Novaseq 6000 and preprocessed with Space Ranger v2.1.0. Patient analyses were conducted according to the Declaration of Helsinki. Ethical approval for performing research on anonymized, archival patient material was obtained from the cantonal ethics commission Zurich (BASEC Nr. 2022-01854 and BASEC-Nr. 2024-01428). Spatial transcriptomics sequencing was performed at the Functional Genomics Center Zurich (FGCZ) of University of Zurich and ETH Zurich. The Visium spots were annotated in the corresponding H&E images by expert researchers (K.S. and S.D.) and included manual annotations with the following labels: TLS, Immune, Tumor, Normal, and Unassigned. One lung (LC4) and one kidney (KC2) slice were excluded from the analysis due to a misalignment between the spatial transcriptomics data and the expected morphological features (Figure S2).

4.5.4 TCGA

For out-of-distribution validation and de novo spatial transcriptomics prediction, we used patient metadata, image slides (fresh frozen - FF and formalin-fixed paraffin-embedded - FFPE) and paired bulk RNA-seq from the TCGA Research Network (https://www.cancer.gov/ccg/research/genome-sequencing/tcga) from samples with skin melanoma (TCGA SKCM), and kidney cancer (TCGA KIRC). When multiple FF sections were available for the same tissue sample, the top section was selected. All slides were scaled to 20x magnification. To distinguish between in-tissue spots and background, we computed the mean RGB value for each tile and discarded all tiles with a mean greater than 200 (close to white).

4.6 Data availability

The kidney cancer HEST-1K and colon cancer HEST-1K datasets were downloaded using the HEST-1K download pipeline outlined here: https://github.com/mahmoodlab/HEST/blob/main/tutorials/1-Downloading-HEST-1k.ipynb, with data available at Hugging Face https://huggingface.co/datasets/MahmoodLab/hes. The Tumor Profiler spatial transcriptomics dataset was obtained from [5]. The lung and kidney cancer with TLS dataset is available at https://zenodo.org/records/14620362.

TCGA spatial transcriptomics samples are available for download on Hugging Face. Access details can be found at https://github.com/ratschlab/DeepSpot.

4.7 Code availability

The open-source implementation of DeepSpot along with a tutorial is available at:

- https://github.com/ratschlab/DeepSpot

The Snakemake pipeline for reproducing the results is available at:

- https://github.com/ratschlab/he2st

Acknowledgments. This work was supported by the Swiss Federal Institutes of Technology (strategic focus area of personalized health and related technologies; 2021–367). The 10x spatial transcriptomics sequencing of a subset of the Tumor Profiler samples was made possible through a technology access program by 10x Genomics™, with special acknowledgments to Jacob Stern, James Chell, Rudi Schläfli, Laura Lipka, Mario Werner, Nikhil Rao, and Scott Brouilette for their invaluable contributions. The Tumor Profiler study was supported by a public-private partnership involving Roche Holding AG, ETH Zurich, University of Zurich, University Hospital Zurich, and University Hospital Basel.

Consent for publication. This manuscript has been seen and approved by all listed authors. The figures were created using BioRender.com and exported under a paid subscription.

Funding. We gratefully acknowledge funding from the Tumor Profiler Initiative and the Tumor Profiler Center (to V.H.K., G.R.). The Tumor Profiler study is jointly funded by a public-private partnership involving F. Hoffmann-La Roche Ltd., ETH Zurich, University of Zurich, University Hospital Zurich, and University Hospital Basel. We also acknowledge funding of K.N. by Swiss National Science Foundation (SNSF) grants 220127 (to G.R.) and 201656, ETH core funding (to G.R.), UZH core funding (to V.H.K.), funding by the Promedica Foundation grant F-87701-41-01 (to V.H.K.), SNSF Prima grant PR00P3-201656 (to K.S.) and funding from the Swiss Federal Institutes of Technology strategic focus area of personalized health and related technologies project 2021-367 (to G.R., V.H.K., S.A.).

Conflict of interest/Competing interests. V.H.K reports being an invited speaker for Sharing Progress in Cancer Care (SPCC) and Indica Labs; advisory board of Takeda; and sponsored research agreements with Roche and IAG, all unrelated to the current study. VHK is a participant in a patent application on the assessment of cancer immunotherapy biomarkers by digital pathology; a patent application on multimodal deep learning for the prediction of recurrence risk in cancer patients, and a patent application on predicting the efficacy of cancer treatment using deep learning all unrelated to the current work. GR is a participant in a patent application on matching cells from different measurement modalities which is not directly related to the current work. Moreover, G.R. is a cofounder of Computomics GmbH, Germany, and one of its shareholders.

TUMOR PROFILER CONSORTIUM. Rudolf Aebersold[5], Melike Ak[33], Faisal S Al-Quaddoomi[12,22], Silvana I Albert[10], Jonas Albinus[10], Ilaria Alborelli[29], Sonali Andani[9,22,31,36], Per-Olof Attinger[14], Marina Bacac[21], Daniel Baumhoer[29], Beatrice Beck-Schimmer[44], Niko Beerenwinkel[7,22], Christian Beisel[7], Lara Bernasconi[32], Anne Bertolini[12,22], Bernd Bodenmiller[11,40], Ximena Bonilla[9], Lars Bosshard[12,22], Byron Calgua[29], Ruben Casanova[40], Stéphane Chevrier[40], Natalia Chicherova[12,22], Ricardo Coelho[23], Maya D'Costa[13], Esther Danenberg[42], Natalie R Davidson[9], Monica-Andreea Dragan[7], Reinhard Dummer[33], Stefanie Engler[40], Martin Erkens[19], Katja Eschbach[7], Cinzia Esposito[42], André Fedier[23], Pedro F Ferreira[7], Joanna Ficek-Pascual[1,9,16,22,31], Anja L Frei[36], Bruno Frey[18], Sandra Goetze[10], Linda Grob[12,22], Gabriele Gut[42], Detlef Günther[8], Pirmin Haeuptle[3], Viola Heinzelmann-Schwarz[23,28], Sylvia Herter[21], Rene Holtackers[42], Tamara Huesser[21], Alexander Immer[9,17], Anja Irmisch[33], Francis Jacob[23], Andrea Jacobs[40], Tim M Jaeger[14], Katharina Jahn[7], Alva R James[9,22,31], Philip M Jermann[29], André Kahles[9,22,31], Abdullah Kahraman[22,36], Viktor H Koelzer[36,41], Werner Kuebler[30], Jack Kuipers[7,22], Christian P Kunze[27], Christian Kurzeder[26], Kjong-Van Lehmann[2,4,9,15], Mitchell Levesque[33], Ulrike Lischetti[23], Flavio C Lombardo[23], Sebastian Lugert[13], Gerd Maass[18], Markus G Manz[35], Philipp Markolin[9], Martin Mehnert[10], Julien Mena[5], Julian M Metzler[34], Nicola Miglino[35,41], Emanuela S Milani[10], Holger Moch[36], Simone Muenst[29], Riccardo Murri[43], Charlotte KY Ng[29,39], Stefan Nicolet[29], Marta Nowak[36], Monica Nunez Lopez[23], Patrick GA Pedrioli[6], Lucas Pelkmans[42], Salvatore Piscuoglio[23,29], Michael Prummer[12,22], Prélot, Laurie[9,22,31], Natalie Rimmer[23], Mathilde Ritter[23], Christian Rommel[19], María L Rosano-González[12,22], Gunnar Rätsch[1,6,9,22,31], Natascha Santacroce[7], Jacobo Sarabia del Castillo[42], Ramona Schlenker[20], Petra C Schwalie[19], Severin Schwan[14], Tobias Schär[7], Gabriela Senti[32], Wenguang Shao[10], Franziska Singer[12,22], Sujana Sivapatham[40], Berend Snijder[5,22], Bettina Sobottka[36], Vipin T Sreedharan[12,22], Stefan Stark[9,22,31],

Daniel J Stekhoven[12,22], Tanmay Tanna[7,9], Alexandre PA Theocharides[35], Tinu M Thomas[9,22,31], Markus Tolnay[29], Vinko Tosevski[21], Nora C Toussaint[12,22], Mustafa A Tuncel[7,22], Marina Tusup[33], Audrey Van Drogen[10], Marcus Vetter[25], Tatjana Vlajnic[29], Sandra Weber[32], Walter P Weber[24], Rebekka Wegmann[5], Michael Weller[38], Fabian Wendt[10], Norbert Wey[36], Andreas Wicki[35,41], Mattheus HE Wildschut[5,35], Bernd Wollscheid[10], Shuqing Yu[12,22], Johanna Ziegler[33], Marc Zimmermann[9], Martin Zoche[36], Gregor Zuend[37]

[1]AI Center at ETH Zurich, Andreasstrasse 5, 8092 Zurich, Switzerland, [2]Cancer Research Center Cologne-Essen, University Hospital Cologne, Cologne, Germany, [3]Cantonal Hospital Baselland, Medical University Clinic, Rheinstrasse 26, 4410 Liestal, Switzerland, [4]Center for Integrated Oncology Aachen (CIO-A), Aachen, Germany, [5]ETH Zurich, Department of Biology, Institute of Molecular Systems Biology, Otto-Stern-Weg 3, 8093 Zurich, Switzerland, [6]ETH Zurich, Department of Biology, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland, [7]ETH Zurich, Department of Biosystems Science and Engineering, Mattenstrasse 26, 4058 Basel, Switzerland, [8]ETH Zurich, Department of Chemistry and Applied Biosciences, Vladimir-Prelog-Weg 1-5/10, 8093 Zurich, Switzerland, [9]ETH Zurich, Department of Computer Science, Institute of Machine Learning, Universitätstrasse 6, 8092 Zurich, Switzerland, [10]ETH Zurich, Department of Health Sciences and Technology, Otto-Stern-Weg 3, 8093 Zurich, Switzerland, [11]ETH Zurich, Institute of Molecular Health Sciences, Otto-Stern-Weg 7, 8093 Zurich, Switzerland, [12]ETH Zurich, NEXUS Personalized Health Technologies, Wagistrasse 18, 8952 Zurich, Switzerland, [13]F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland, [14]F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland, , [15]Joint Research Center Computational Biomedicine, University Hospital RWTH Aachen, Aachen, Germany, [16]Life Science Zurich Graduate School, Biomedicine PhD Program, Winterthurerstrasse 190, 8057 Zurich, Switzerland, [17]Max Planck ETH Center for Learning Systems, , [18]Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany, [19]Roche Pharmaceutical Research and Early Development, Roche Innovation Center Basel, Grenzacherstrasse 124, 4070 Basel, Switzerland, [20]Roche Pharmaceutical Research and Early Development, Roche Innovation Center Munich, Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany, [21]Roche Pharmaceutical Research and Early Development, Roche Innovation Center Zurich, Wagistrasse 10, 8952 Schlieren, Switzerland, [22]SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland, [23]University Hospital Basel and University of Basel, Department of Biomedicine, Hebelstrasse 20, 4031 Basel, Switzerland, [24]University Hospital Basel and University of Basel, Department of Surgery, Brustzentrum, Spitalstrasse 21, 4031 Basel, Switzerland, [25]University Hospital Basel, Brustzentrum & Tumorzentrum, Petersgraben 4, 4031 Basel, Switzerland, [26]University Hospital Basel, Brustzentrum, Spitalstrasse 21, 4031 Basel, Switzerland, [27]University Hospital Basel, Department of Information- and Communication Technology, Spitalstrasse 26, 4031 Basel, Switzerland, [28]University Hospital Basel, Gynecological Cancer Center, Spitalstrasse 21, 4031 Basel, Switzerland, [29]University Hospital Basel, Institute of Medical Genetics and Pathology, Schönbeinstrasse 40, 4031 Basel, Switzerland, [30]University Hospital Basel, Spitalstrasse 21/Petersgraben 4, 4031 Basel, Switzerland, [31]University Hospital Zurich, Biomedical Informatics, Schmelzbergstrasse 26, 8006 Zurich, Switzerland, [32]University Hospital Zurich, Clinical Trials Center, Rämistrasse 100, 8091 Zurich, Switzerland, [33]University Hospital Zurich, Department of Dermatology, Gloriastrasse 31, 8091 Zurich, Switzerland, [34]University Hospital Zurich, Department of Gynecology, Frauenklinikstrasse 10, 8091 Zurich, Switzerland, [35]University Hospital Zurich, Department of Medical Oncology and Hematology, Rämistrasse 100, 8091 Zurich, Switzerland, [36]University Hospital Zurich, Department of Pathology and Molecular Pathology, Schmelzbergstrasse 12, 8091 Zurich, Switzerland, [37]University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland, [38]University Hospital and University of Zurich, Department of Neurology, Frauenklinikstrasse 26, 8091 Zurich, Switzerland, [39]University of Bern, Department of BioMedical Research, Murtenstrasse 35, 3008 Bern, Switzerland, [40]University of Zurich, Department of Quantitative Biomedicine, Winterthurerstrasse 190, 8057 Zurich, Switzerland, [41]University of Zurich, Faculty of Medicine, Zurich, Switzerland, [42]University of Zurich, Institute of Molecular Life Sciences, Winterthurerstrasse 190, 8057 Zurich, Switzerland, [43]University of Zurich, Services and Support for Science IT, Winterthurerstrasse 190, 8057 Zurich, Switzerland, [44]University of Zurich, VP Medicine, Künstlergasse 15, 8001 Zurich, Switzerland

References

[1] Maynard, K.R., Collado-Torres, L., Weber, L.M., Uytingco, C., Barry, B.K., Williams, S.R., Catallini, J.L., Tran, M.N., Besich, Z., Tippani, M., *et al.*: Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. Nature neuroscience **24**(3), 425–436 (2021)

[2] Rao, A., Barkley, D., França, G.S., Yanai, I.: Exploring tissue architecture using spatial transcriptomics. Nature **596**(7871), 211–220 (2021)

[3] Baccin, C., Al-Sabah, J., Velten, L., Helbling, P.M., Grünschläger, F., Hernández-Malmierca, P., Nombela-Arrieta, C., Steinmetz, L.M., Trumpp, A., Haas, S.: Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. Nature cell biology **22**(1), 38–48 (2020)

[4] Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., *et al.*: Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science **353**(6294), 78–82 (2016)

[5] Nonchev, K., Andani, S., Ficek-Pascual, J., Nowak, M., Sobottka, B., Consortium, T.P., Koelzer, V.H., Raetsch, G.: Representation learning for multi-modal spatially resolved transcriptomics data. medRxiv, 2024–06 (2024)

[6] Yoosuf, N., Navarro, J.F., Salmén, F., Ståhl, P.L., Daub, C.O.: Identification and transfer of spatial transcriptomics signatures for cancer diagnosis. Breast Cancer Research **22**, 1–10 (2020)

[7] Longo, S.K., Guo, M.G., Ji, A.L., Khavari, P.A.: Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. Nature Reviews Genetics **22**(10), 627–644 (2021)

[8] Zhou, R., Yang, G., Zhang, Y., Wang, Y.: Spatial transcriptomics in development and disease. Molecular Biomedicine **4**(1), 32 (2023)

[9] Du, J., Yang, Y.-C., An, Z.-J., Zhang, M.-H., Fu, X.-H., Huang, Z.-F., Yuan, Y., Hou, J.: Advances in spatial transcriptomics and related data analysis strategies. Journal of Translational Medicine **21**(1), 330 (2023)

[10] Choe, K., Pak, U., Pang, Y., Hao, W., Yang, X.: Advances and challenges in spatial transcriptomics for developmental biology. Biomolecules **13**(1), 156 (2023)

[11] Sirinukunwattana, K., Domingo, E., Richman, S.D., Redmond, K.L., Blake, A., Verrill, C., Leedham, S.J., Chatzipli, A., Hardy, C., Whalley, C.M., *et al.*: Image-based consensus molecular subtype (imcms) classification of colorectal cancer using deep learning. Gut **70**(3), 544–554 (2021)

[12] Schmauch, B., Romagnoni, A., Pronier, E., Saillard, C., Maillé, P., Calderaro, J., Kamoun, A., Sefta, M., Toldo, S., Zaslavskiy, M., *et al.*: A deep learning model to predict rna-seq expression of tumours from whole slide images. Nature communications **11**(1), 3877 (2020)

[13] Alsaafin, A., Safarpoor, A., Sikaroudi, M., Hipp, J.D., Tizhoosh, H.: Learning to predict rna sequence expressions from whole slide images with applications for search and classification. Communications Biology **6**(1), 304 (2023)

[14] Jaume, G., Doucet, P., Song, A.H., Lu, M.Y., Almagro-Pérez, C., Wagner, S.J., Vaidya, A.J., Chen, R.J., Williamson, D.F., Kim, A., et al.: Hest-1k: A dataset for spatial transcriptomics and histology image analysis. arXiv preprint arXiv:2406.16192 (2024)

[15] Fan, Z., Chen, R., Chen, X.: Spatialdb: a database for spatially resolved transcriptomes. Nucleic acids research **48**(D1), 233–237 (2020)

[16] Xu, Z., Wang, W., Yang, T., Chen, J., Huang, Y., Gould, J., Du, W., Yang, F., Li, L., Lai, T., et al.: Stomicsdb: a database of spatial transcriptomic data. bioRxiv, 2022–03 (2022)

[17] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708 (2017)

[18] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255 (2009). Ieee

[19] He, B., Bergenstråhle, L., Stenbeck, L., Abid, A., Andersson, A., Borg, Å., Maaskola, J., Lundeberg, J., Zou, J.: Integrating spatial gene expression and breast tumour morphology via deep learning. Nature biomedical engineering **4**(8), 827–834 (2020)

[20] Chen, R.J., Ding, T., Lu, M.Y., Williamson, D.F., Jaume, G., Song, A.H., Chen, B., Zhang, A., Shao, D., Shaban, M., *et al.*: Towards a general-purpose foundation model for computational pathology. Nature Medicine **30**(3), 850–862 (2024)

[21] Filiot, A., Ghermi, R., Olivier, A., Jacob, P., Fidon, L., Mac Kain, A., Saillard, C., Schiratti, J.-B.: Scaling self-supervised learning for histopathology with masked image modeling. medRxiv, 2023–07 (2023)

[22] Xie, R., Pang, K., Chung, S., Perciani, C., MacParland, S., Wang, B., Bader, G.: Spatially resolved gene expression prediction from histology images via bi-modal contrastive learning. Advances in Neural Information Processing Systems **36** (2024)

[23] Pang, M., Su, K., Li, M.: Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. BioRxiv, 2021–11 (2021)

[24] Jia, Y., Liu, J., Chen, L., Zhao, T., Wang, Y.: Thitogene: a deep learning method for predicting spatial transcriptomics from histological images. Briefings in Bioinformatics **25**(1), 464 (2024)

[25] Zeng, Y., Wei, Z., Yu, W., Yin, R., Yuan, Y., Li, B., Tang, Z., Lu, Y., Yang, Y.: Spatial transcriptomics prediction from histology jointly through transformer and graph neural networks. Briefings in Bioinformatics **23**(5), 297 (2022)

[26] Zaheer, M., Kottur, S., Ravanbakhsh, S., Poczos, B., Salakhutdinov, R.R., Smola, A.J.: Deep sets. Advances in neural information processing systems **30** (2017)

[27] Saillard, C., Jenatton, R., Llinares-López, F., Mariet, Z., Cahané, D., Durand, E., Vert, J.-P.: H-optimus-0. https://github.com/bioptimus/releases/tree/main/models/h-optimus/v0

[28] Wolf, F.A., Angerer, P., Theis, F.J.: Scanpy: large-scale single-cell gene expression data analysis. Genome biology **19**, 1–5 (2018)

[29] Mohamed, A., Gonzalez, R.S., Lawson, D., Wang, J., Cohen, C.: Sox10 expression in malignant melanoma, carcinoma, and normal tissues. Applied Immunohistochemistry & Molecular Morphology **21**(6), 506–510 (2013)

[30] Graf, S.A., Busch, C., Bosserhoff, A.-K., Besch, R., Berking, C.: Sox10 promotes melanoma cell invasion by regulating melanoma inhibitory activity. Journal of Investigative Dermatology **134**(8), 2212–2220 (2014)

[31] Heumos, L., Schaar, A.C., Lance, C., Litinetskaya, A., Drost, F., Zappia, L., Lücken, M.D., Strobl, D.C., Henao, J., Curion, F., *et al.*: Best practices for single-cell analysis across modalities. Nature Reviews Genetics **24**(8), 550–572 (2023)

[32] Harpio, R., Einarsson, R.: S100 proteins as cancer biomarkers with focus on s100b in malignant melanoma. Clinical biochemistry **37**(7), 512–518 (2004)

[33] Gudbjartsson, D.F., Sulem, P., Stacey, S.N., Goldstein, A.M., Rafnar, T., Sigurgeirsson, B., Benediktsdottir, K.R., Thorisdottir, K., Ragnarsson, R., Sveinsdottir, S.G., *et al.*: Asip and tyr pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. Nature genetics **40**(7), 886–891 (2008)

[34] Price, M.A., Colvin Wanshura, L.E., Yang, J., Carlson, J., Xiang, B., Li, G., Ferrone, S., Dudek,

A.Z., Turley, E.A., McCarthy, J.B.: Cspg4, a potential therapeutic target, facilitates malignant progression of melanoma. Pigment cell & melanoma research **24**(6), 1148–1157 (2011)

[35] Wang, X., Wang, Y., Yu, L., Sakakura, K., Visus, C., Schwab, J., Ferrone, C., Favoino, E., Koya, Y., Campoli, M., *et al.*: Cspg4 in cancer: multiple roles. Current molecular medicine **10**(4), 419–429 (2010)

[36] Zhang, S., Chen, K., Liu, H., Jing, C., Zhang, X., Qu, C., Yu, S.: Pmel as a prognostic biomarker and negatively associated with immune infiltration in skin cutaneous melanoma (skcm). Journal of Immunotherapy **44**(6), 214–223 (2021)

[37] Li, W.-S., Chen, C.-I., Chen, H.-P., Liu, K.-W., Tsai, C.-J., Yang, C.-C.: Overexpression of mlph in rectal cancer patients correlates with a poorer response to preoperative chemoradiotherapy and reduced patient survival. Diagnostics **11**(11), 2132 (2021)

[38] Tokuo, H., Bhawan, J., Coluccio, L.M.: Myosin x is required for efficient melanoblast migration and melanoma initiation and metastasis. Scientific reports **8**(1), 10449 (2018)

[39] McKee, C.S., Hill, D.S., Redfern, C.P., Armstrong, J.L., Lovat, P.E.: Oncogenic braf signalling increases m cl-1 expression in cutaneous metastatic melanoma. Experimental dermatology **22**(11), 767–769 (2013)

[40] Senchenko, V.N., Krasnov, G.S., Dmitriev, A.A., Kudryavtseva, A.V., Anedchenko, E.A., Braga, E.A., Pronina, I.V., Kondratieva, T.T., Ivanov, S.V., Zabarovsky, E.R., *et al.*: Differential expression of chl1 gene during development of major human cancers. PloS one **6**(3), 15612 (2011)

[41] Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., *et al.*: The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature **483**(7391), 603–607 (2012)

[42] Vincent, K.M., Postovit, L.-M.: Investigating the utility of human melanoma cell lines as tumour models. Oncotarget **8**(6), 10498 (2017)

[43] Hui, Y.-H., Wood, K.V., McLaughlin, J.L.: Bullatencin, 4-deoxyasimicin, and the uvariamicins: Additional bioactive annonaceous acetogenins from annona bullata rich.(annonaceae). Natural Toxins **1**(1), 4–14 (1992)

[44] Beberok, A., Wrześniok, D., Szlachta, M., Rok, J., Rzepka, Z., Respondek, M., Buszman, E.: Lomefloxacin induces oxidative stress and apoptosis in colo829 melanoma cells. International journal of molecular sciences **18**(10), 2194 (2017)

[45] Chen, Y., Kramer, D.L., Diegelman, P., Vujcic, S., Porter, C.W.: Apoptotic signaling in polyamine analogue-treated sk-mel-28 human melanoma cells. Cancer research **61**(17), 6437–6444 (2001)

[46] Mazzio, E.A., Soliman, K.F.: Whole-transcriptomic profile of sk-mel-3 melanoma cells treated with the histone deacetylase inhibitor: Trichostatin a. Cancer Genomics & Proteomics **15**(5), 349–364 (2018)

[47] Badia-i-Mompel, P., Santiago, J.V., Braunger, J., Geiss, C., Dimitrov, D., Müller-Dott, S., Taus, P., Dugourd, A., Holland, C.H., Flores, R.O.R., Saez-Rodriguez, J.: decoupler: ensemble of computational methods to infer biological activities from omics data. Bioinformatics Advances (2022) https://doi.org/10.1093/bioadv/vbac016

[48] Inamdar, G.S., Madhunapantula, S.V., Robertson, G.P.: Targeting the mapk pathway in melanoma: why some approaches succeed and other fail. Biochemical pharmacology **80**(5), 624–637 (2010)

[49] Amaral, T., Sinnberg, T., Meier, F., Krepler, C., Levesque, M., Niessner, H., Garbe, C.: The mitogen-activated protein kinase pathway in melanoma part i–activation and primary resistance

mechanisms to braf inhibition. European journal of cancer **73**, 85–92 (2017)

[50] Shain, A.H., Yeh, I., Kovalyshyn, I., Sriharan, A., Talevich, E., Gagnon, A., Dummer, R., North, J., Pincus, L., Ruben, B., *et al.*: The genetic evolution of melanoma from precursor lesions. New England Journal of Medicine **373**(20), 1926–1936 (2015)

[51] Rademacher, A., Huseynov, A., Bortolomeazzi, M., Wille, S.J., Schumacher, S., Sant, P., Keitel, D., Okonechnikov, K., Ghasemi, D., Pajtler, K., et al.: Comparison of spatial transcriptomics technologies used for tumor cryosections. bioRxiv, 2024–04 (2024)

[52] Sato, Y., Silina, K., Broek, M., Hirahara, K., Yanagita, M.: The roles of tertiary lymphoid structures in chronic diseases. Nature Reviews Nephrology **19**(8), 525–537 (2023)

[53] Hou, Y., Qiao, S., Li, M., Han, X., Wei, X., Pang, Y., Mao, H.: The gene signature of tertiary lymphoid structures within ovarian cancer predicts the prognosis and immunotherapy benefit. Frontiers in Genetics **13**, 1090640 (2023)

[54] Rijthoven, M., Obahor, S., Pagliarulo, F., Broek, M., Schraml, P., Moch, H., Laak, J., Ciompi, F., Silina, K.: Multi-resolution deep learning characterizes tertiary lymphoid structures and their prognostic relevance in solid tumors. Communications Medicine **4**(1), 5 (2024)

[55] Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., Rajpoot, N.: Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Medical image analysis **58**, 101563 (2019)

[56] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. Neural computation **1**(4), 541–551 (1989)

[57] Diederik, P.K.: Adam: A method for stochastic optimization. (No Title) (2014)

[58] Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., *et al.*: Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic acids research **44**(W1), 90–97 (2016)

[59] Schubert, M., Klinger, B., Klünemann, M., Sieber, A., Uhlitz, F., Sauer, S., Garnett, M.J., Blüthgen, N., Saez-Rodriguez, J.: Perturbation-response genes reveal signaling footprints in cancer gene expression. Nature communications **9**(1), 20 (2018)

[60] Simon, N., Friedman, J.H., Hastie, T., Tibshirani, R.: Regularization paths for cox's proportional hazards model via coordinate descent. Journal of statistical software **39**, 1–13 (2011)

[61] Meylan, M., Petitprez, F., Becht, E., Bougoüin, A., Pupier, G., Calvez, A., Giglioli, I., Verkarre, V., Lacroix, G., Verneau, J., *et al.*: Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. Immunity **55**(3), 527–541 (2022)

[62] Valdeolivas, A., Amberg, B., Giroud, N., Richardson, M., Gálvez, E.J., Badillo, S., Julien-Laferrière, A., Turos, D., Voithenberg, L.V., Wells, I., et al.: Charting the heterogeneity of colorectal cancer consensus molecular subtypes using spatial transcriptomics. bioRxiv, 2023–01 (2023)