nature biomedical engineering

Article

https://doi.org/10.1038/s41551-026-01754-z

HisToSpatialCNV: an interpretable deep learning method predicting spatial copy number variations from histopathology images

Received: 22 September 2025

Accepted: 12 June 2026

Published online: 24 August 2026

Check for updates

Tianao Chen[1], Thatchayut Unjitwattana[2], Xuhui Guo[3], Pooja Thakur[3], Shu Zhou[4], Zheng Jing[5], Yiwen Yang[4], Yuheng Du[4], Weiping Zou[6] & Lana X. Garmire[2,3,4]

Copy number variations (CNV) are key drivers of cancer progression, yet methods for predicting spatial CNVs directly from haematoxylin and eosin (H&E) images are currently lacking. We introduce HisToSpatialCNV, a multiscale deep-learning framework that integrates histopathological features from H&E-stained images to infer spatial CNV patterns. Combining interpretable feature extraction with graph neural networks and multihead self-attention, our approach captures both local and global tissue contexts. Applied to HER2+ breast cancer, skin cancer and brain cancer datasets, HisToSpatialCNV outperformed existing methods for spatial gene expression inference and showed strong concordance between predicted CNVs and gene expression. In addition, it enabled tumour subclone identification, phylogenetic reconstruction and detection of pathway alterations linked to tumour progression. HisToSpatialCNV generalized across Visium and Xenium spatial transcriptomics platforms, on HER2+ patients. Applying the HisToSpatialCNV HER2+ model to TCGA HER2+ histopathology data identified spatial-molecular subtypes associated with distinct survival outcomes in TCGA HER2+ patients. By connecting histopathology to spatial genomics, HisToSpatialCNV offers a powerful, cost-effective tool for studying intratumour heterogeneity using only routine pathology images.

Histology imaging is widely regarded as the gold standard for cancer identification and diagnosis[1]. Among the available techniques, haematoxylin and eosin (H&E) staining is the most widely used staining method for preparing histology images[2]. Compared to other biomarker modalities, such as radiological imaging (CT, MRI, PET) and genetic biomarkers (genetic mutations, RNA expression and protein markers), H&E-stained histological images are more cost-effective and provide direct visualization of tissue architecture and cellular morphology[2]. The cellular-level assessment enables us to understand tumour heterogeneity, identify invasive cancer cells and investigate the tumour microenvironment.

DNA copy number variation (CNV) plays a critical role in chromosomal rearrangements and genomic diseases[3]. CNVs are large-scale genomic changes, typically exceeding 1 kilobase, which manifest as submicroscopic deletions and duplications affecting a larger portion of the genome than single-nucleotide polymorphisms (SNPs). These

A full list of affiliations appears at the end of the paper. e-mail: lgarmire@uab.edu

variations are strongly associated with various cancers, including breast cancer[4] and colorectal cancer[5], gastric cancer[6], ovarian cancer[7], testicular germ cell tumours[8] and squamous cell carcinoma[9]. Beyond directly affecting large genomic regions, CNVs frequently encompass oncogenes and tumour suppressor genes, establishing a close link with key differential gene expression[10]. For example, copies of apoptosis-related genes are commonly lost during cancer development, whereas proliferation-related genes are typically retained[11]. CNVs also serve as crucial references for drug development, especially those therapies targeting cancer progression mechanisms, particularly when considering tumour heterogeneity[12] and subclonal evolution[13,14]. Conventionally, CNVs are detected using techniques such as fluorescence in situ hybridization[15], array-based comparative genomic hybridization[16] and single-nucleotide polymorphism arrays[16]. However, the high cost of these methods limits their widespread application in large-scale cancer studies. Computational methods, such as inferCNV[17], allow inference of CNVs from transcriptomic data without requiring additional CNV-detection experiments.

Recently, several methods have been proposed to infer CNVs from spatial transcriptomics data exclusively. For example, STARCH[18] leverages the spatial organization of cells within a tumour to infer copy number alterations, while CalicoST[19] uses a probabilistic model to infer allele-specific copy number aberrations and tumour phylogeography. However, these approaches do not utilize H&E images, which directly provide critical information on the tissue structure, cellular morphology and pathological features. We hypothesize that H&E images, with their detailed cellular morphology and complex visual features, hold immense predictive potential for spatial CNVs. Towards this, we propose HisToSpatialCNV, an innovative multiscale deep-learning framework that integrates H&E image features, spatial transcriptomics data and spatial context information to deliver accurate spatially informative CNV prediction with robust interpretability. Our approach also enables comprehensive downstream analyses, including tumour subclone identification, inference of evolutionary relationships and characterization of functional pathways. Based on pathology images alone, HisToSpatialCNV provides important insights into tumour heterogeneity and progression mechanisms that are informative to precision medicine. Overall, HisToSpatialCNV represents a powerful and cost-effective framework that leverages widely available histology images to fill in the void between pathology and spatial genetics (CNVs), enabling a deeper understanding of cancer biology and offering new opportunities for large-scale translational research.

Results

Overview of the HisToSpatialCNV method

The workflow of HisToSpatialCNV is illustrated in Fig. 1, comprising the HisToSpatialCNV model architecture and downstream analysis. HisToSpatialCNV directly predicts spatially resolved CNV states from H&E-stained histopathology images, rather than from the spatial transcriptomics data. Such inferred spatial CNVs can then be used to analyse tumour cell states, subclones and their pathophysical differences. As a result, the model can contribute to the comprehensive understanding of the underlying mechanisms of tumour progression, provide important information for tumour subtyping, help to identify biomarkers, and facilitate patient prognosis analysis and prediction.

In the training phase, HisToSpatialCNV is designed as a flexible framework to take paired H&E histopathology and spatial CNV information as the input. While experimental methods to measure spatial CNVs are being developed, the spatial CNV information can be broadly obtained using the inferCNV[17] method on spatial transcriptomics data currently. Histological images are then cropped into patches centred on the spot coordinates in the matched spatial transcriptomics (ST) data. These patches are processed using feature extraction tools (default: CellProfiler[20]) to extract interpretable and meaningful cellular and tissue characteristics, addressing a key limitation of black-box image-based methods. These extracted features are then represented by reduced-dimension embeddings using a 3-layer residual multilayer perceptron (ResMLP) with layer normalization and rectified linear unit (ReLU) activations, stabilizing gradient flow and mitigating vanishing gradients. The spot coordinates are used to create a spatial adjacency matrix that represents the graph structure based on Euclidean distances between spots. Following spatial CNV prediction with HisToSpatialCNV, a series of downstream analyses, including visualization, phylogenetic tree construction, pathway enrichment analyses and spatial domain identification are conducted to demonstrate its utilities.

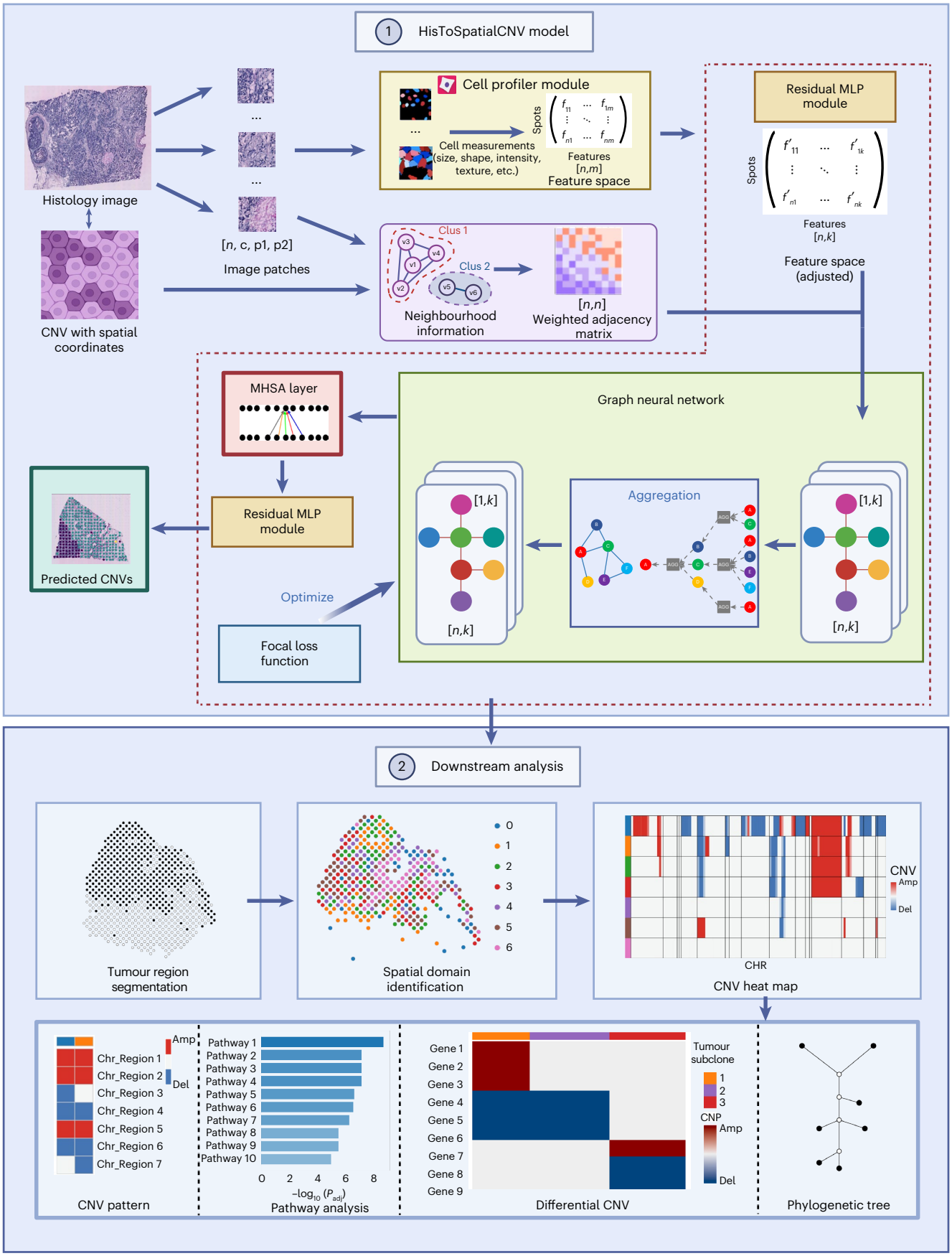
HisToSpatialCNV employs multiple analytical innovations. (1) Its sequential architecture progressively enriches representations from interpretable features to spatial context to global dependencies. The refined embeddings and spatial adjacency matrix feed into a sophisticated graph neural network (GNN) module that integrates local and higher-order structural information through stacked graph convolution layers. (2) It uses a jumping knowledge (JK) network[21] with a long short-term memory network (LSTM) aggregator[22], which enables the model to integrate intermediate representations from each GNN layer, capturing both local and extended spatial neighbourhoods. (3) It applies a multihead self-attention mechanism (MHSA)[23] to model global relationships between spatially distant regions, unlike existing methods that use transformers for initial feature extraction. (4) It uses a 3-layer residual MLP, incorporating LayerNorm (ResMLP) and ReLU activation, and outputs 3-category (amplification, deletion and no change) CNV predictions via a softmax layer. (5) It adopts focal loss[24] which prioritizes samples that are difficult to classify due to the imbalance issue among CNV categories.

Performance evaluation of HisToSpatialCNV

To quantitatively evaluate our model, we benchmarked on the following datasets: a HER2+ breast cancer dataset[25] with 26 tissue sections from the Visium platform, and a cutaneous squamous cell carcinoma (cSCC) dataset[26] with 9 tissue sections from the ST platform optimized for cSCC tissue (Supplementary Fig. 1a). Since models directly predicting spatial CNVs from H&E images are currently lacking, for comparison to HisToSpatialCNV we turned to methods that infer spatial gene expression data from H&E images and then estimated the spatial CNVs using inferCNV[17]. These spatial gene expression inference methods include the task-specific methods ST-Net[27], HisToGene[28], Hist2ST[29], TCGN[30], GHIST[31], iSCALE[32] and iSTAR[33]; the transformer-based contrastive learning method mclSTExp[34]; and the pretrained foundation model OmiCLIP[35]. As there are currently no experimental methods that directly measure spatial CNVs, we used the CNVs directly inferred from the paired spatial transcriptomics data coupled with the H&E images as the

Fig. 1 | Overview of the HisToSpatialCNV framework and analysis. The workflow consists of two main components: (1) the prediction model architecture and (2) downstream analysis. The HisToSpatialCNV model integrates histological image features and spatial CNVs with coordinates. Image features are processed through a 3-layer residual MLP to generate 1,024-dimensional embeddings, while spot coordinates create a spatial adjacency matrix. These inputs are fed into a graph neural network with multiple graph convolutional layers and a jumping knowledge network featuring an LSTM aggregator to capture multiscale features. A multihead self-attention mechanism then enriches node embeddings by modelling global dependencies. Finally, a 3-layer residual MLP with layer normalization and ReLU activation outputs 3-category CNV predictions via softmax, with focal loss addressing class imbalance. Downstream analyses include visualization, phylogenetic tree construction, spatial region identification using a GNN-based approach, differential CNV analysis, pathway enrichment analysis and validation on the TCGA breast cancer dataset. This comprehensive approach enables tumour subtyping, prognosis analysis and biomarker identification through the characterization of CNVs. Figure created in BioRender; Garmire, L. https://biorender.com/6vl3hj1 (2026).

https://doi.org/10.1038/s41551-026-01754-z



ground truth. For all methods in comparison, we used a leave-one-out cross-validation (LOOCV) evaluation on the same data split: in each iteration, all sections except one were used for training, and one hold-out section was reserved for testing (Supplementary Fig. 1b). All model architectures and hyperparameters were fixed before evaluation, and were evaluated on the same held-out test sections under identical data splits, ensuring a fair comparison.

As shown in Fig. 2a, HisToSpatialCNV achieved the highest area under the receiver operating characteristic curve (AUC) (0.825) across all samples in the HER2+ dataset, substantially higher than those of other approaches including ST-Net (AUC = 0.505), HisToGene (AUC = 0.512), Hist2ST (AUC = 0.554), TCGN (AUC = 0.536), mclSTExp (AUC = 0.655), OmiCLIP (AUC = 0.610), GHIST (AUC = 0.588), iSCALE (AUC = 0.557) and iSTAR (AUC = 0.505). Similarly, it also achieved significantly higher overall mean AUCs (AUC = 0.748) for cSCC data, compared with ST-Net (AUC = 0.500), HisToGene (AUC = 0.504), Hist2ST (AUC = 0.539), TCGN (AUC = 0.566), OmiCLIP (AUC = 0.559), mclSTExp (AUC = 0.561), GHIST (AUC = 0.549), iSCALE (AUC = 0.519) and iSTAR (AUC = 0.489). The superiority of HisToSpatialCNV is also demonstrated by the receiver operating characteristic (ROC) curves of these two datasets (Fig. 2b). We further evaluated the classification performance of the model across the 3 CNV states (deletion, neutral and amplification) in both datasets using confusion matrix analysis (Fig. 2c). For HER2+ samples, our model achieves high precision in identifying CNV deletions (79.4% accuracy), neutral regions (73.6% accuracy) and amplifications (74.0% accuracy). Similar performance is observed in the cSCC dataset, particularly in amplification states (65.4% accuracy). In addition, we examined the relationships between predicted CNV states from HisToSpatialCNV and the gene expression measured by the coupled ST data[36] (Fig. 2d). There is a clear positive correlation between the two data modalities in both datasets. The regions predicted as CNV+ (amplification) show significantly higher gene expression values, whereas CNV− (deletion) regions show lower expression.

To further evaluate the predicted CNVs, we analysed the alignment between the predicted and observed CNVs of representative genes from the HER2+ dataset (Fig. 2e). We selected CNVs from 4 genes: *CDH1* (ref. 37), *MIIP*[38], *FADS2* (ref. 39) and *SOX4* (refs. 40,41) with known relevance to cancer. HisToSpatialCNV efficiently detected CNVs, as shown by the Matthews correlation coefficient (MCC) values between the inferred spatial CNVs and the ground truths (0.621, 0.661, 0.758 and 0.606, respectively). In contrast, the other methods (Hist2ST, HisToGene and TCGN) failed to detect many or all of the CNVs. HisToSpatialCNV also accurately captured the spatial distribution patterns of key CNVs in the cSCC dataset, when compared to the ground truth derived from spatial transcriptomics (Supplementary Fig. 2). Together, these results highlight the robustness, accuracy and generalizability of HisToSpatialCNV in detecting spatial CNV patterns across different tumour types.

Ablation study reveals the importance of ResMLP layers

Given the complexity of HisToSpatialCNV, we conducted ablation studies to evaluate the contribution of various modules to model performance. Using results from HisToSpatialCNV as the positive control, we demonstrate the impacts of removing different architectural components leading to three HisToSpatialCNV variants (Fig. 3a): (1) a baseline model with both ResMLP and MHSA modules removed, (2) HisToSpatialCNV without the ResMLP module (HisToSpatialCNV−ResMLP) and (3) HisToSpatialCNV without the multihead self-attention module (HisToSpatialCNV−MHSA). As expected, the original HisToSpatialCNV with all modules achieved the highest average AUCs of ~0.825 and 0.748 for the HER2+ and cSCC datasets, respectively. The baseline models with both ResMLP and MHSA modules removed had the lowest accuracies. Notably, removing the ResMLP module resulted in a substantial decrease in average AUCs to ~0.693 and 0.641 for HER2+ and cSCC datasets, respectively. Removing the MHSA module reduced the accuracies to a lesser extent, to ~0.783 and 0.682 for HER2+ and cSCC datasets. These findings demonstrate that while both components contribute to model performance, the hybrid MLP module plays a more important role in capturing relevant spatial patterns.

In addition, we compared how training and prediction accuracy are affected by three distinct feature extraction approaches: CellProfiler[20], ResNet-50 (ref. 42) and the state-of-the-art foundation model UNI[43] (Fig. 3b). The training curves show that HisToSpatialCNV using the CellProfiler[20] module achieved the highest performance across both datasets, with AUCs of ~0.825 for HER2+ and 0.748 for cSCC. ResNet-50 showed comparable performance with AUCs of 0.822 and 0.744, respectively, while UNI exhibited slightly lower performance at 0.785 and 0.720. Notably, both ResNet-50 and UNI demonstrated substantially faster convergence compared with CellProfiler. ResNet-50 and UNI required only 22.5 and 24.2 min for HER2+, respectively, compared with 66.1 min for CellProfiler. Similarly, ResNet-50 and UNI converged in 40.3 and 42.2 min for cSCC, respectively, while CellProfiler required 76.9 min. Despite being slower, the CellProfiler[20] module offers the advantage of interpretability by extracting specific image features, which can be ranked by integrated gradient (IG)[44] values to elucidate feature importance in model predictions (Fig. 3c). For HER2+ tissues, area shape and granularity features show the highest contributions. Similarly, in cSCC tissues, area shape, granularity and texture difference variance features are most informative. The ResNet-50 module, on the other hand, generates high-dimensional, abstract feature vectors that are difficult to interpret.

To visualize how HisToSpatialCNV focuses on histologically relevant regions, we compared its spatial attention patterns in the H&E images with the annotations from pathologists (Fig. 3d). The corresponding attention heat maps reveal that our model learns to focus on regions rich in CNVs, as demonstrated by brighter colours which correspond to higher attention weights. These regions include tumours annotated by the experts and surrounding so-called 'adjacent normal regions' that already have abnormal CNVs—a phenomenon well observed before[13]. Detailed quantitative region-level analysis confirmed significantly elevated CNV burdens in tumour-associated regions versus non-tumour regions annotated by pathologists, with AUC = 0.68 and $P = 8.04 \times 10^{-6}$ in sample B1, and AUC = 0.80 and $P = 3.41 \times 10^{-17}$ in sample D1 (Supplementary Table 1). These results highlight that the model's predictions are based on meaningful tissue characteristics, thereby enhancing the interpretability and trustworthiness of the predicted CNVs to assist in clinical applications.

Downstream functional analysis of predicted spatial CNVs reveals subclones of HER2+ breast cancer

To demonstrate the utility of spatial CNVs predicted by HisToSpatialCNV, we conducted downstream subclone analysis and functional characterization using the B1 sample of the HER2+ cohort as an example (Fig. 4 and Supplementary Fig. 3). We classified the spots into tumour versus adjacent tissue labels, using the predicted spatial CNVs (Fig. 4a). Such CNV-guided annotations clearly distinguished tumour from adjacent normal regions. Focusing on the tumour compartment, we then applied a graph convolutional network-based spatial domain identification approach to classify tumour spots into 4 distinct subclones, 0–3 (Fig. 4b), where the optimal number of subclones was determined using Silhouette score (Supplementary Fig. 4a). Uniform manifold approximation and projection (UMAP) visualization of CNV profiles further confirms the clear separation among these subclones (Fig. 4c). Moreover, a chromosome-level CNV heat map of the 4 subclones reveals progressive genetic alterations among subclones (Fig. 4d). For example, subclone 3 exhibits unique deletion patterns in chromosome 1 but amplification patterns in chromosomes 2 and 10, while subclones 1 and 2 share similar patterns that differ from that of subclone 3. Consistent with this pattern, phylogenetic reconstruction based on CNV similarity

https://doi.org/10.1038/s41551-026-01754-z

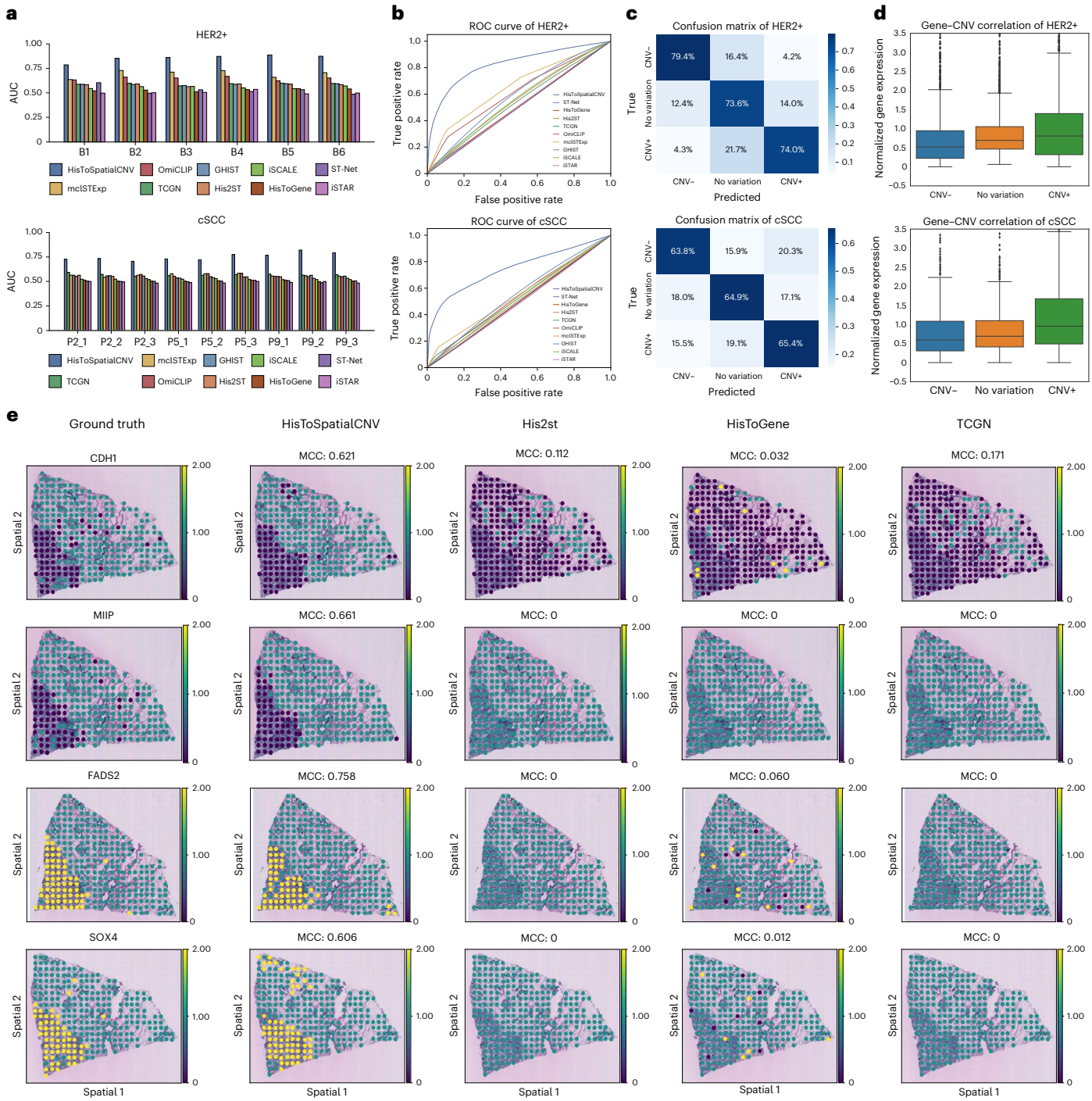


Fig. 2 | Performance evaluation of HisToSpatialCNV for spatial CNV prediction. **a**, Comparison of AUC values across individual tissue sections for HER2+ and cSCC datasets, showing consistently superior performance of HisToSpatialCNV over other opponent task-specific methods (ST-Net, Hist2ST, HisToGene, TCGN, GHIST, iSCALE and iSTAR), a transformer-based method (mclSTExp) and an FM method (OmiCLIP). **b**, ROC curves demonstrate the overall performance across all sections, with HisToSpatialCNV achieving significantly higher mean AUCs compared with baseline approaches. **c**, Confusion matrices illustrate the classification performance across 3 CNV states (deletion, neutral, amplification). **d**, Boxplots showing the positive correlation between predicted CNV states and normalized gene expression levels. The gene expression values were log transformed using $\log_{1p}$ transformation ($\log(1+x)$) to account for the typically skewed distribution of expression data. Each data point represents one gene. For the HER2+ panel, $n = 5$ patients covering 26 tissue sections; for the cSCC panel, $n = 3$ patients covering 9 tissue sections. Boxplots show the median (centre line), interquartile range (IQR, box bounds), whiskers extending to 1.5× IQR, and outliers as individual points. The compared groups correspond to inferred CNV states (loss/no variation/gain). **e**, Visualization of ground truth and predicted CNV patterns for 4 cancer-relevant genes (*CDH1*, *MIIP*, *FADS2* and *SOX4*) across representative tissue sections. Spot colours indicate CNV states (purple, deletion; green, neutral; yellow, amplification).

https://doi.org/10.1038/s41551-026-01754-z

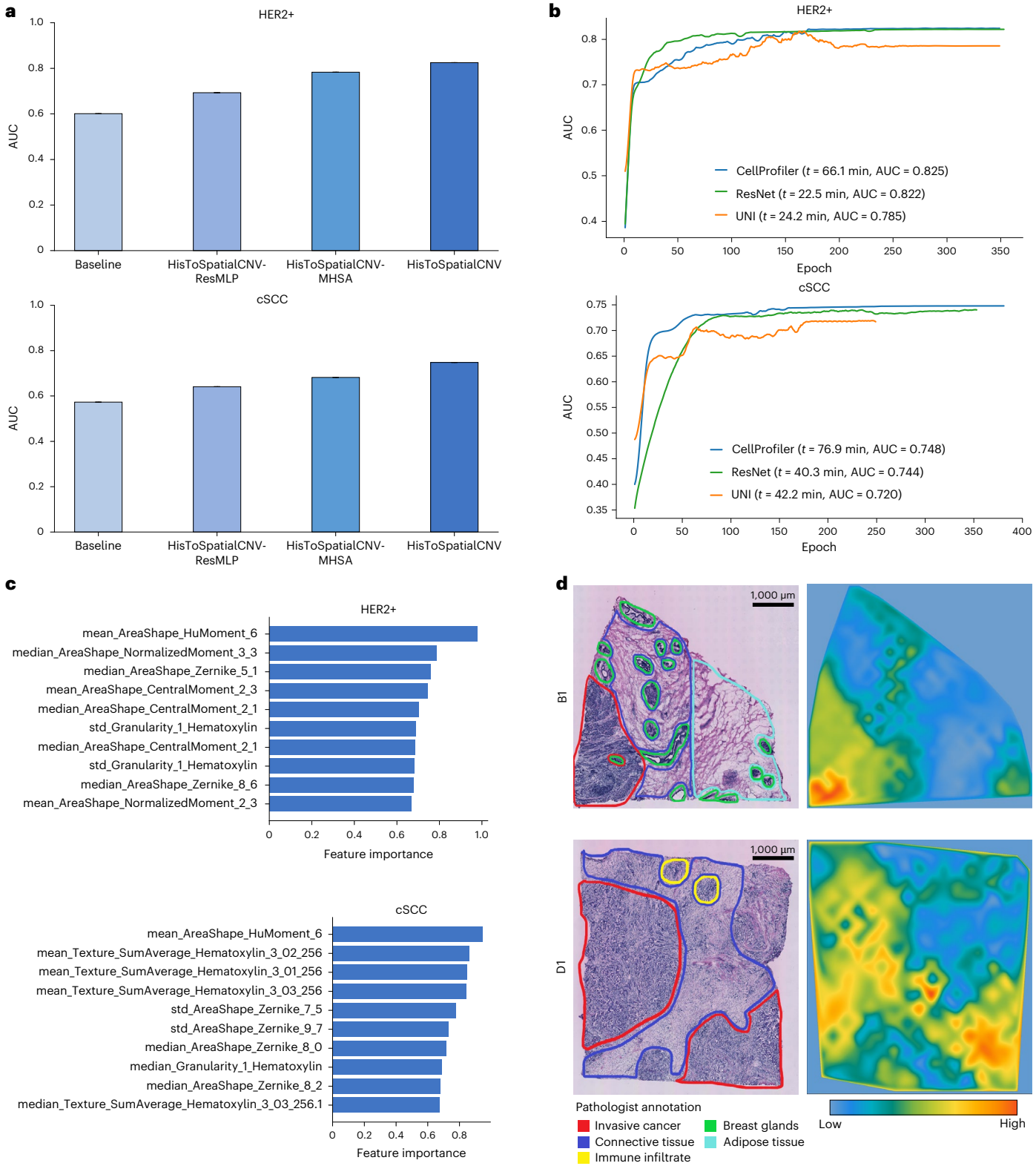


Fig. 3 | Ablation study and interpretability analysis of HisToSpatialCNV in HER2+ and cSCC datasets. a, Comparison of AUC performance across different model variants of HisToSpatialCNV: HisToSpatialCNV without multihead self-attention (HisToSpatialCNV−MHSA), HisToSpatialCNV without residual MLP (HisToSpatialCNV−ResMLP) and the baseline model in which both MHSA and ResMLP modules are removed. **b**, Training curves comparing CellProfiler, ResNet-50 and UNI feature extraction approaches. CellProfiler achieves marginally better final performance, but ResNet-50 and UNI demonstrate faster convergence. **c**, Shapley additive explanations (SHAP) values identifying the most contributive CellProfiler features for HER2+ and cSCC datasets. **d**, Comparison of a pathologist's tumour annotations (left) with model-derived CNV attention heat maps (right) for samples B1 and D1 from the HER2+ dataset. Red outlines indicate invasive tumour regions, with other significant structures marked in blue, green and yellow. Warmer colours in the attention maps (yellow to red) represent higher attention weights for CNV inference. Note that the high-attention CNV regions include the tumour and surrounding regions that already have CNVs, also known as field cancerization.

https://doi.org/10.1038/s41551-026-01754-z

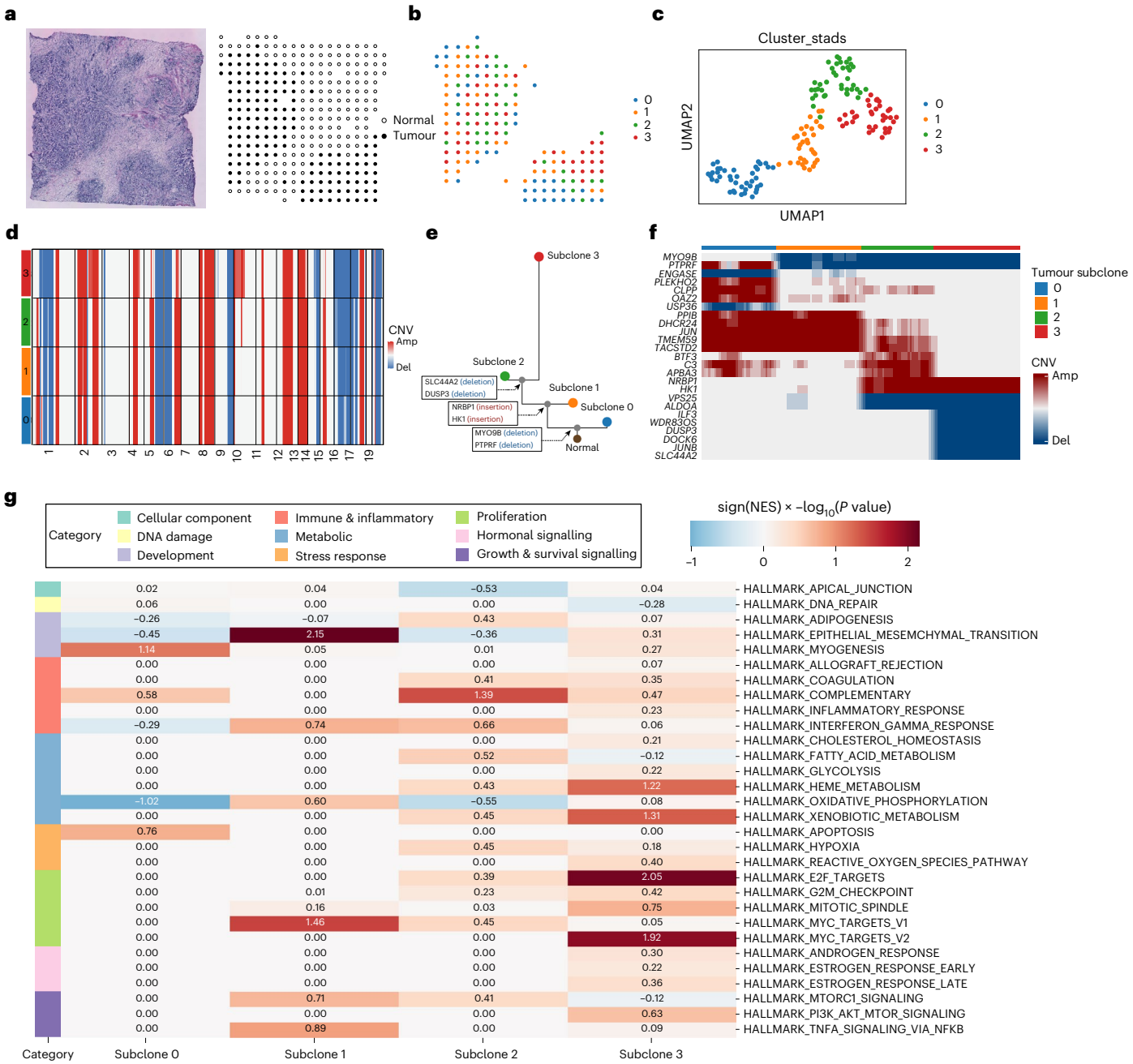


Fig. 4 | Spatial domain identification and functional characterization of tumour subclones in HER2+ dataset. a, H&E-stained tissue section with corresponding tumour segmentation. Black spots represent tumour, and white spots represent normal cells. **b**, Spatial distribution of 4 distinct subclones (labelled 0–3) identified using the GNN-based approach applied to predicted spatial CNV data. **c**, UMAP plot confirming the differences of the 4 identified subclones, with a clear separation between clusters. **d**, Heat map of CNVs mapped to the corresponding locations on the chromosomes between subclones. Red and blue colour indicate amplification and deletion, respectively. **e**, Phylogenetic tree of subclones in HER2+ breast cancer tissue. **f**, Gene-level differential analysis of significant CNVs in each subclone, revealing subclone-specific CNV alterations. **g**, MSigDB hallmark pathway enrichment analysis showing significantly altered pathways for each subclone. The pathways are categorized into 8 functional groups as labelled by different colours. NES and *P* values are shown on the figure. Schematics created in BioRender: **e**, Garmire, L. https://biorender.com/q2hxvdr (2026); **g**, Garmire, L. https://biorender.com/oqbbi9d (2026).

reveals a hierarchical evolutionary structure with increasing CNV burden from subclone 0 to subclone 3 (Fig. 4e and Supplementary Fig. 5a). Subclones 0, 1 and 2 form a closely related branch, whereas subclone 3 diverges with additional chromosomal events. Centroid-distance analyses jointly indicate that subclone 2 is a distinct genomic state rather than a mixture of subclones 1 and 3 (Supplementary Fig. 6). Many of the key CNVs that are characteristic of individual subclones have previously been implicated in tumour-associated processes, including deletions in *MYO9B*[45,46], *PTPRF*[47–49], *SLC44A2* (ref. 50) and *DUSP3* (refs. 51–53), and insertions in *NRBP1* (ref. 54) and *HK1* (ref. 55) (Fig. 4e and Supplementary Fig. 3b–j).

To fully investigate the genetic alterations across all four tumour subclones, we conducted gene-level differential spatial CNV analysis (Fig. 4f). Amplification of *JUN*, a proto-oncogene regulating cell proliferation and survival[56], is prominent in subclones 0 and 1, and partially retained in subclone 2. A similar pattern to *JUN* is observed in *TMEM59*

and *TACSTD2*, consistent with their previously reported roles in cancer progression[57,58]. In contrast, deletion of multiple genes, including *ILF3, WDR83OS, DUSP3, JUNB* and *SLC44A2*, is unique in subclone 3. Amplification of *BTF3*, which encodes a transcription factor implicated in cell cycle progression and apoptosis resistance[59], is prevalent in subclone 2 but not in subclones 1 and 3, supporting the idea that subclone 2 is not a simple mixture of subclones 1 and 3.

To examine the collective functional consequences of the differential CNVs shown in Fig. 4f, we next performed pathway enrichment analysis among the subclones (Fig. 4g). Subclone 0 showed relatively limited pathway perturbation, including enhanced myogenesis and apoptosis but suppression of both oxidative phosphorylation and epithelial–mesenchymal transition (EMT), reflecting a lineage commitment state under metabolic stress. In contrast, subclone 1 showed marked activation of EMT, MYC-target pathways and oxidative phosphorylation, suggesting active proliferation and transcriptional reprogramming associated with tumour plasticity and early aggressiveness[60,61]. Subclone 2 displayed broader but very distant functional enrichment from subclone 1, with activation of the innate immune complement cascade and increased fatty acid metabolism, adipogenesis and xenobiotic metabolism signatures, indicating tumour-promoting inflammation[62,63]. Subclone 3 demonstrated the most extensive pathway activation, including activation of MYC targets, E2F signalling and PI3K/AKT/mTOR pathways, as well as xenobiotic and haem metabolism pathways. Notably, many pathway-level activities in subclone 2 do not lie between those in subclones 1 and 3 (for example, EMT, innate immune activation and oxidative phosphorylation), further supporting the idea that subclone 2 represents a distinct evolutionary group rather than a simple mixture of subclones 1 and 3. Collectively, these patterns indicate progressive intensification of proliferative and metabolic programmes along the inferred evolutionary trajectory.

We applied the same analytical framework to the cSCC dataset (Supplementary Note, and Supplementary Figs. 7 and 8). In contrast to HER2+ breast cancer, the inferred cSCC subclones exhibited distinct pathway enrichment patterns characterized by UV response, DNA repair and inflammatory signalling pathways, consistent with the known biology of UV-driven skin cancers. Together, these findings highlight the utility of HisToSpatialCNV in enabling integrative analysis of spatial CNV architecture, subclonal evolution and associated functional pathway signatures within tumour tissues.

Application of the HisToSpatialCNV HER2+ model to the TCGA cohort identifies distinct prognostic subtypes

To demonstrate the predictive power of the HisToSpatialCNV HER2+ model, we directly applied it to predict spatial CNV patterns in all 92 HER2+ breast cancer samples from another study[64]. These samples are a subset from The Cancer Genome Atlas (TCGA), but have histopathology images manually annotated by pathologists for tumour and non-tumour regions[64] (Fig. 5a). Pathologist annotation is an important input for the inferCNV programme. Using 224 × 224-resolution windows from H&E-stained images, we created a total of 312,923 continuous pseudo-spots for this HER2+ cohort, allowing us to comprehensively characterize CNV spatial heterogeneity throughout the tumour microenvironment at an unprecedented resolution.

First, we evaluated the ability of the spatial CNVs (inferred by HisToSpatialCNV) to distinguish early- versus late-stage HER2+ breast cancers (Fig. 5b). Following the tumour-node-metastasis staging system[65], we classified stages 1 and 2 as early-stage disease and stages 3 and 4 as late-stage disease. The ROC curve demonstrates excellent predictive accuracy (AUC = 0.825) when using spatial CNVs inferred by the HisToSpatialCNV model as features. In contrast, using bulk-measured CNV data from the corresponding HER2+ samples yielded a substantially lower AUC of 0.594, highlighting the added value of spatial CNV information for disease staging.

Next, we utilized the inferred CNV data from HisToSpatialCNV and their corresponding spatial coordinates to generate spatially aware embeddings using a GNN-based approach (see Methods). We applied *K*-means clustering to these embeddings and identified 3 distinct spatial clusters among the 90 samples that passed quality filtering (see Methods) (Supplementary Fig. 9a). Using the proportion of each of the 3 spatial clusters as the metric to represent each patient, we clustered the HER2+ patients into 4 new patient subtypes using the elbow method (Supplementary Fig. 9b). Figure 5c shows the heat map of the spatial clusters and the patient subtypes. Patient subtype 1 is enriched with spatial cluster 2, patient subtype 2 is enriched with spatial cluster 0, patient subtypes 3 and 4 are enriched with spatial cluster 1 in different degrees. The subsequent Kaplan–Meier survival analysis shows that patient subtypes 1, 3 and 4 all exhibit an excellent prognosis with 100% survival probability over the follow-up period, but patient subtype 2 has the worst survival among the 4 patient subtypes (Fig. 5d).

To further understand the poor prognosis in patient subtype 2, we performed additional analyses on this patient subtype versus other subtypes. Through differential CNV analysis, we identified genes with significant CNVs that distinguish subtype 2 from others (Fig. 5e). Many genes known to be associated with tumour proliferation, metastasis and cancer progression are present in subtype 2, including *MKNK2* (ref. 66), *BSG*[67] and *BCL3* (ref. 68). The pathway enrichment analysis revealed several enriched cancer hallmark pathways (Fig. 5f), including the most significant mTORC1 signalling pathway, followed by the oxidative phosphorylation pathway, as well as other signalling pathways (for example, notch signalling), G2-M checkpoint regulation and lipid metabolism pathways (for example, cholesterol homeostasis). These findings suggest that dysregulation of multiple cancer-associated processes, particularly cell proliferation and lipid metabolism, characterizes this aggressive tumour subtype. Together, these results highlight the powerful clinical relevance of spatial CNV profiling using HisToSpatialCNV, demonstrating its ability to enhance patient stratification and prognostic assessment.

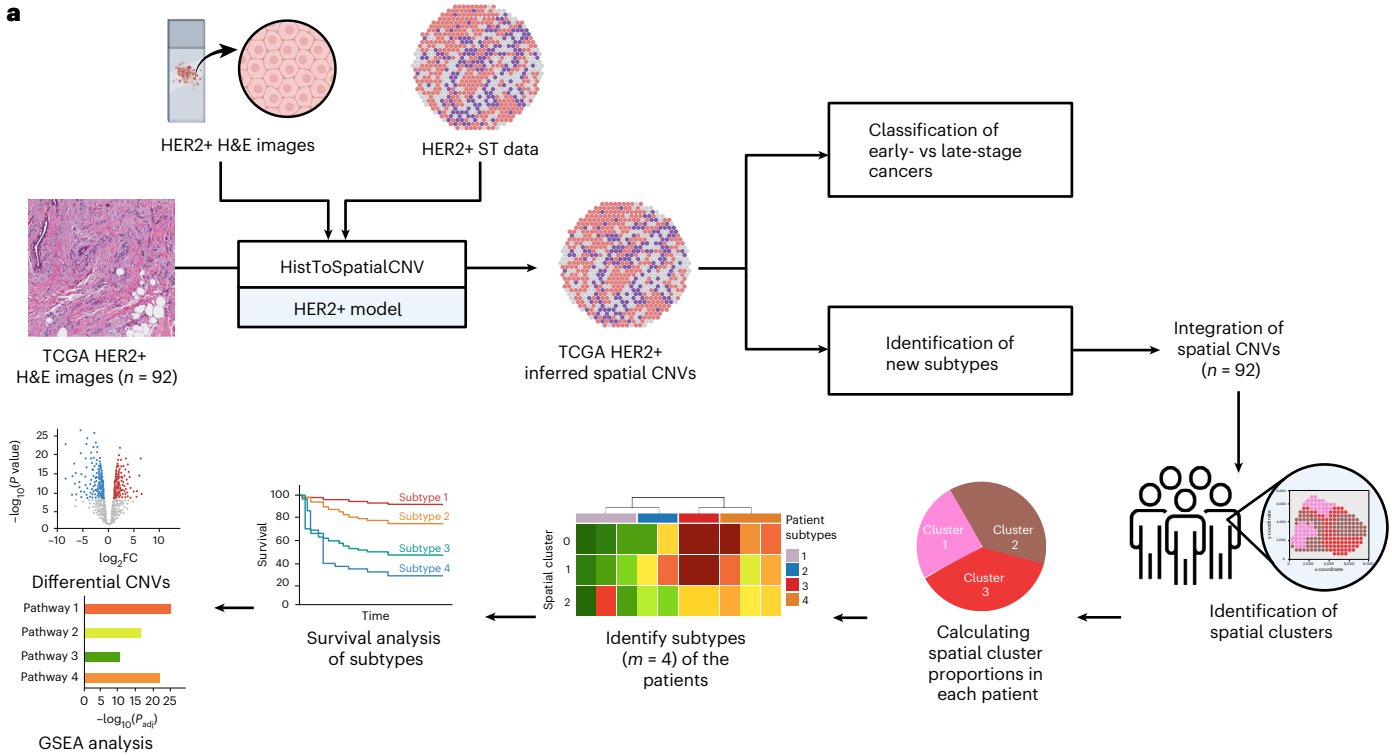
Generalization of HisToSpatialCNV models

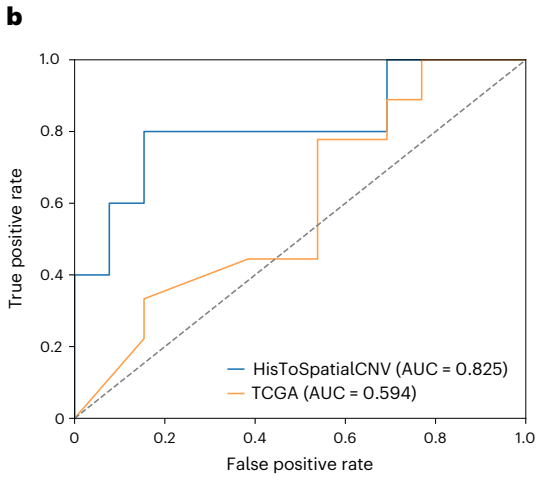
To evaluate within-dataset robustness and potential data leakage, we tested the HisToSpatialCNV model previously trained from the 26 H&E-stained sections of HER2+ breast cancers, on 6 additional HER2+ patient samples that were not included in the training set (Fig. 6a). The method maintained consistent performance with AUC values ranging from 0.715 to 0.747, demonstrating the within-dataset robustness of HisToSpatialCNV on new patient data. We further evaluated HisToSpatialCNV using the direct CNV measurements as the ground truth.

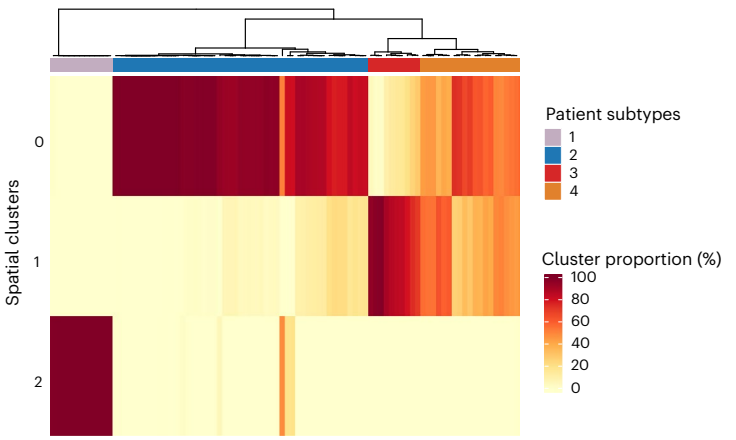
Fig. 5 | Spatial CNV inference on external HER2+ breast cancer cohort using the HisToSpatialCNV HER2+ model. a, Overall summary of spatial CNV inference, stage classification and prognostic stratification of another HER2+ breast cancer cohort ($n = 92$). **b**, ROC curve demonstrating the superior performance of HisToSpatialCNV compared with conventional classification analysis in distinguishing early-stage (stages 1 and 2) versus late-stage (stages 3 and 4) HER2+ cancers. **c**, Hierarchical clustering of the additional cohort of HER2+ breast cancer patients, based on the proportions of 3 spatial clusters identified from *K*-means clustering. The heat map reveals 4 distinct patient subtypes (determined using the elbow method). Colour intensity represents cluster proportion percentage from 0% (light yellow) to 100% (dark maroon). **d**, Kaplan–Meier survival analysis demonstrating significant differences in disease-free survival for subtype 2 compared to the other 3 subtypes. **e**, Differential CNV analysis comparing patient subtype 2 versus the other subtypes. The dot plot displays differentially altered genes, ranked by statistical significance ($-\log_{10}(P \text{ value})$). **f**, Hallmark pathway enrichment analysis of genes with significant CNVs in **e**. Schematic in **a** created in BioRender; Garmire, L. https://biorender.com/jv8nhn1 (2026).

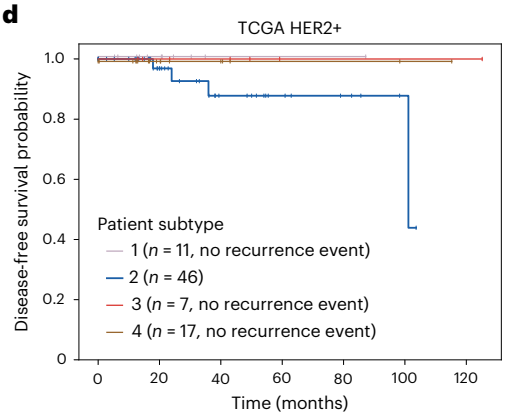
https://doi.org/10.1038/s41551-026-01754-z

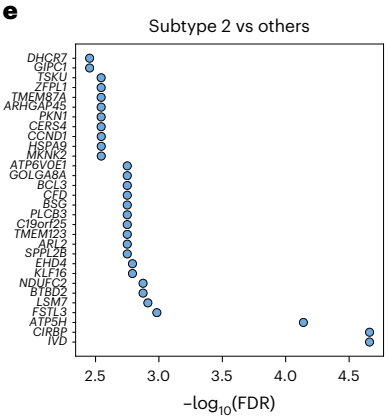
For this we applied the model to H&E images from 2 breast cancer patients in the Human Tumor Atlas Network (HTAN) WashU cohort[69]. These patients have coupled whole-exome sequencing (WES) data. Following the approach of ref. 19, we used the CNVs obtained from the WES data as the ground truth. As shown in Fig. 6b, HisToSpatialCNV achieved a macro-average AUC of 0.651 ± 0.019 (range 0.615–0.671) and a macro-average F1 score of 0.471 ± 0.071 among 9 H&E tissue sections. These consistent results confirm that HisToSpatialCNV can meaningfully predict spatial CNVs directly from H&E images, based on experimentally measured CNV ground truth.

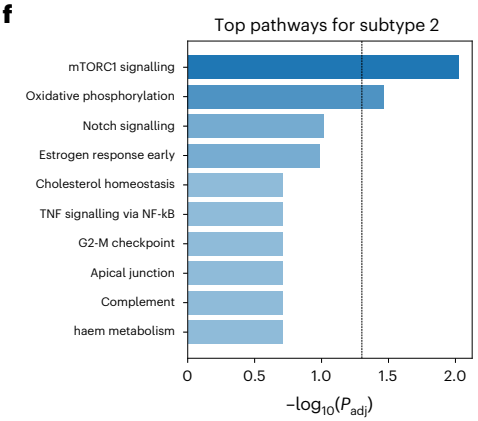












https://doi.org/10.1038/s41551-026-01754-z

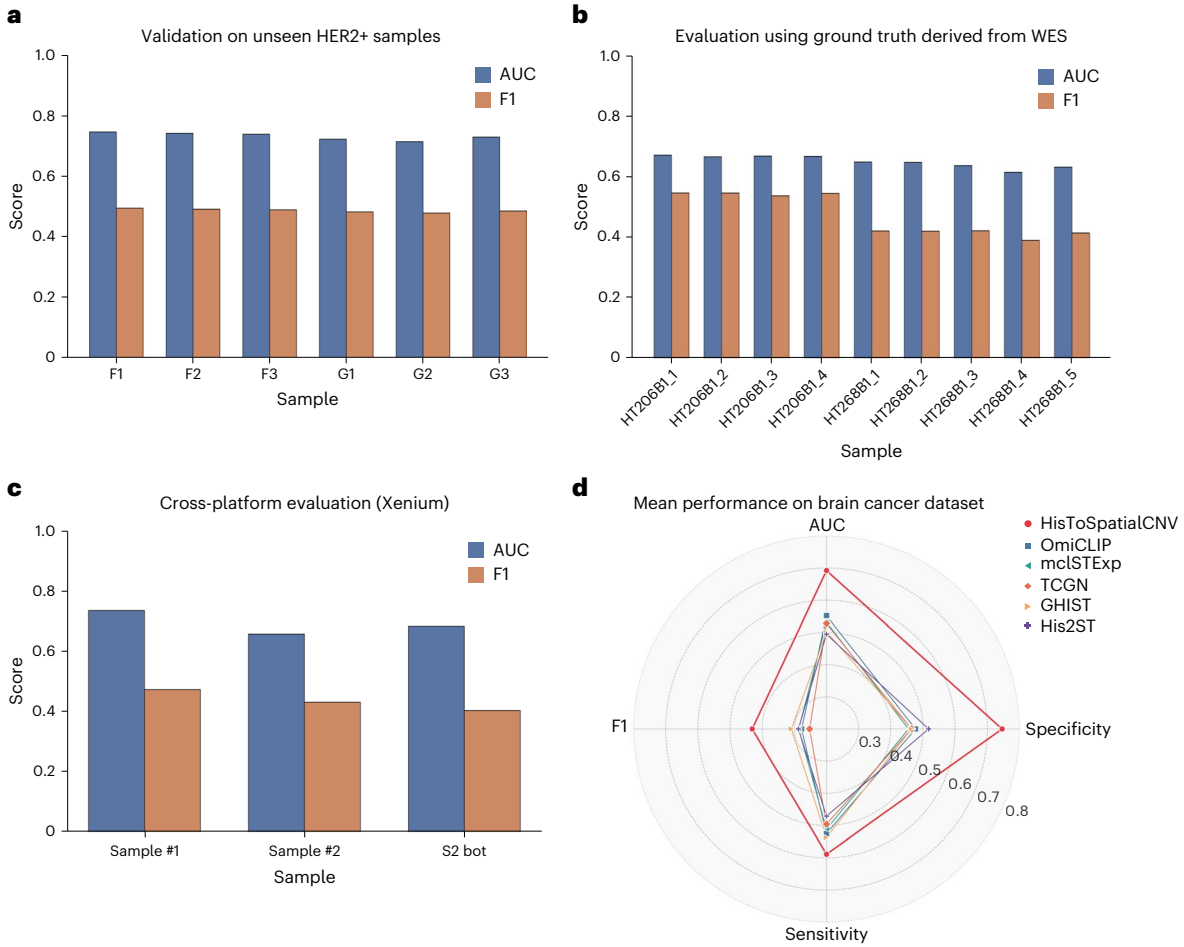


Fig. 6 | Generalization of the HisToSpatialCNV model across datasets and platforms. a, Validation results of the pretrained HisToSpatialCNV HER2+ model on 6 other HER2+ breast cancer samples (F1–F3 and G1–G3) in the same study[25], showing consistent predictability across samples. **b**, Evaluation of HisToSpatialCNV HER2+ model against whole-exome sequencing (WES)-derived CNV ground truth. The HisToSpatialCNV model was directly applied to H&E images from 2 HTAN WashU breast cancer patients that have 9 tissue sections. **c**, Cross-platform evaluation of the HisToSpatialCNV HER2+ model results on Xenium samples. The Visium-trained HisToSpatialCNV HER2+ model was applied to 3 HER2+ breast cancer Xenium samples, whose single-cell profiles were aggregated into pseudo-Visium spots to match the training resolution. **d**, Radar plot showing the performance metrics of the HisToSpatialCNV model and other methods on a brain cancer dataset[73]. The other methods are the top 5 baseline methods selected on the basis of their CNV prediction results on the HER2+ and cSCC datasets in Fig. 2a. The performance metrics include AUC, F1 score, sensitivity and specificity. Figure created in BioRender; Garmire, L. https://biorender.com/6vi8d1x (2026).

To test the cross-platform applicability of HisToSpatialCNV, we applied a model pretrained on HER2+ breast cancer data from the Visium platform (Fig. 2a) to predict high-tumour-content regions in 3 HER2+ breast cancer samples generated using the Xenium platform across two independent studies[70,71]. For compatibility, we aggregated Xenium single-cell profiles into pseudo-Visium spots to match the Visium-scale resolution following a previously described strategy[72]. HisToSpatialCNV achieved good separation between CNV classes, with macro-average AUCs of 0.734, 0.656 and 0.682 on the 3 samples, respectively (Fig. 6c).

Finally, we tested HisToSpatialCNV on brain cancer, which may have more complexity than breast cancer and cSCC tissues. We used a publicly available brain cancer Visium dataset[73] comprising 49 tissue sections from 7 patients. As shown in Fig. 6d, HisToSpatialCNV achieved the highest macro-average AUC of 0.692 ± 0.044 and a macro-average F1 score of 0.431 ± 0.033 across the 5-fold cross-validation, outperforming all other top 5 benchmarked methods in Fig. 2a previously: OmiCLIP[35], mclSTExp[34], GHIST[31], TCGN[30] and Hist2ST[29]. Thus, we conclude that HisToSpatialCNV can be generalized to more heterogeneous and morphologically complex tumour types such as brain cancer. Moreover, the attention map by HisToSpatialCNV highlights the region of high weights of CNVs, which covers the core of the tumour (with high tumour cellularity), and surrounding regions that already have somatic CNVs, a phenomenon well observed before[13] (Supplementary Fig. 10). The HisToSpatialCNV predictions show good concordance with pathologist annotations (ROC-AUC = 0.809, $P = 2.85 \times 10^{-131}$), confirming the value of spatially inferred CNVs.

Discussion

In this study, we introduce HisToSpatialCNV, a novel deep-learning model designed to infer spatial CNVs from histological images alone. HisToSpatialCNV uniquely fills a major void area relating histopathology and spatial CNVs. Unlike spatial transcriptomics, currently, experimental methods that enable unbiased, genome-wide measurement of spatial CNVs directly from tissues are not widely adopted yet. At the same time, other computational methods either predict spatial gene expression rather than spatial CNVs from histopathology, predict bulk CNVs without spatial resolution from histopathology, or predict CNVs from spatial transcriptomics rather than from histopathology. We exemplified the utilities of HisToSpatialCNV using datasets from HER2+ breast cancer and cutaneous squamous cell carcinoma (cSCC). In addition, HisToSpatialCNV facilitates a wide variety of functional

applications such as individualized subclone identification, phylogenetic tree inference and population-level new patient subtype identification. All of these highlight the potential clinical utility of HisToSpatialCNV-based spatial CNV prediction.

HisToSpatialCNV demonstrated high accuracy across both datasets after comprehensive benchmarking, even much better than the current popular foundation models (FMs) that claim to predict spatial gene expression well[35]. As a task-specific (spatial CNV) prediction tool, it is not surprising that HisToSpatialCNV outperforms other single-cell FMs for spatial gene expression. Similar observations were also shared by others, in that single-cell FMs currently do not necessarily outperform established task-specific methods across all downstream tasks[74–77]. HisToSpatialCNV owes its success to the underlying algorithm and multiple architecture designs. Technically, HisToSpatialCNV employs a sequential architecture that progressively processes representations from interpretable features to spatial context to global dependencies. The model incorporates a jumping knowledge network with LSTM aggregation within its GNN module, which integrates representations from multiple graph convolution layers to capture both local and extended spatial neighbourhoods. Unlike methods that apply transformers for initial feature extraction, HisToSpatialCNV positions the multihead self-attention mechanism after spatial structure has been established through the GNN, enabling more contextually informed modelling of long-range dependencies between spatially distant regions. This design maintains the benefits of global attention modelling while avoiding the computational overhead of full transformer encoders. Ablation studies confirmed that both the ResMLP and MHSA modules contribute significantly to prediction accuracy, with their removal resulting in a notable performance drop. Furthermore, CellProfiler[20]-based feature extraction achieved accuracy comparable to or exceeding those of ResNet-50 and the foundation model-based UNI method, reinforcing the value of interpretable, handcrafted features in enhancing both performance and explainability.

Some key strengths of HisToSpatialCNV lie in its interpretability and its ability to support rich downstream functional analyses. By incorporating CellProfiler[20]-derived features, such as cell shape, size, intensity and texture, the model offers biologically meaningful and human-interpretable metrics that go beyond black-box predictions. These interpretable features not only enhance model transparency but also enable detailed characterization of spatial CNV patterns. Moreover, the MHSA mechanism allows visualization of the most relevant regions for prediction tasks, with spatial patterns in the H&E images aligning with the tumour annotations from pathologists. This provides an opportunity to examine H&E images that may assist pathologists. In addition, HisToSpatialCNV facilitates downstream applications, including spatial clustering, subclone identification, phylogenetic inference, patient subtyping and prognosis prediction, which can better inform oncologists for personalized and precise treatment consideration.

Specifically, visualizations of spatial CNV distributions and their alignment with tissue states reinforce the biological relevance of the predictions. For instance, in the HER2+ model, key genes such as *CDH1* (ref. 37), *MIIP*[38], *FADS2* (ref. 39) and *SOX4* (refs. 40,41) (Fig. 2e), predicted with good accuracies of CNV alteration, are highlighted for their roles in breast cancer progression and metastasis. Downstream analyses of HisToSpatialCNV prediction, including subclone clustering, pathway enrichment and phylogenetic reconstruction, allow precise characterization of the tumour. For example, in HER2+ breast cancer, subclone 3 exhibits distinct CNV patterns (Fig. 4d) and a highly aggressive phenotype characterized by amplification of oncogenic regions and loss of tumour suppressors such as *DUSP3* (refs. 52,78) (Fig. 4f). Confirming this, subclone 3 also displays strong upregulation of MYC-target-related, E2F and PI3K/AKT/mTOR signaling pathways, suggesting high proliferative potential. In short, besides clinical applications, HisToSpatialCNV also offers opportunities to examine in-depth the biological and molecular-level dysfunctions among subclones, which is impossible to obtain using H&E images alone.

Some limitations of this study should be noted. First, due to the lack of experimentally measured spatial CNVs as the ground truth, our initial benchmarking relied on CNV states inferred from spatial transcriptomics data using inferCNV, which may introduce some biases. This is due to the fact that expression-based CNV inference methods such as inferCNV[17] estimate large-scale chromosomal alterations from coordinated gene expression shifts, rather than direct DNA measurements. Noise in gene expression (for example, low coverage), particularly in smaller CNV segments or regions with heterogeneous expression patterns, is likely to introduce inaccuracy in CNVs[79,80]. As a result, it is possible that HisToSpatialCNV learns patterns specific to inferCNV in addition to true CNV signals. To mitigate potential systematic biases, we used tumour versus non-tumour region labels as priors when applying inferCNV. As the histopathology heat maps highlighting regions enriched for informative CNVs showed good concordance with pathologist-annotated tumour core regions, spatial CNV estimates using HisToSpatialCNV does provide a certain level of confidence. To address the limitation of using inferred CNV from spatial transcriptomics data as the ground truth, we additionally validated HisToSpatialCNV against the CNV ground truth derived from whole-exome sequencing (WES) data of the HTAN cohort. Results demonstrated lower but still consistent performance (AUC = 0.65) compared with the CNVs inferred from spatial transcriptomics. However, it should be noted that WES provides bulk genomic measurements rather than spatially matched single-spot profiles, hence it may underestimate the true accuracy. Related to this issue, the other baseline methods benchmarked here, including TCGN, mclSTExp, OmiCLIP, GHIST, Hist2ST, iSCALE, HisToGene, ST-Net and iSTAR, follow a two-stage framework as they are not spatial CNV inference methods but spatial gene expression inference methods. Therefore, inaccuracy in spatial transcriptomics inference from these methods could propagate together with the inaccuracy in CNV inference from spatial transcriptomics data. In comparison, HisToSpatialCNV avoids the intermediate gene-expression reconstruction step, which contributes to improved performance in CNV inference.

Moreover, HisToSpatialCNV models are currently trained on spot-resolution Visium spatial transcriptomics data, given their predominant sizes and availability among current spatial transcriptomics platforms. However, real-world applications will inherently involve cross-dataset and cross-spatial transcriptomics technology settings. Although we showed preliminary applicability of the HisToSpatialCNV HER2+ model to Xenium datasets, broader testing on fine-grained applications (for example, phylogenetic reconstruction and pathway enrichment) are needed in the future. At subcellular resolution, increased technical variability and sparsity in gene expression measurements may require additional more-granularized model construction. Thus, refinement and harmonization with other spatial transcriptomics technologies (for example, Xenium, CosMx, MERFISH or Slide-seqV2) need to be carried out as a follow-up study, following a similar technique recently done by others[81]. Despite these limitations, we demonstrated that HisToSpatialCNV can predict spatial CNVs in new patient samples, identify subtypes in a new population cohort (for example, TCGA HER2+ patients), as well as reveal spatial CNVs in complex cancers, as exemplified by brain cancer.

In addition, HisToSpatialCNV is a task-specific model that trains and evaluates within a specific cancer type, rather than being a foundation model. The latter (for example, OmiCLIP) has been shown to have poorer comparative performance (Fig. 2a). As different cancer types harbour distinct CNV landscapes and histomorphological features, cross-cancer CNV prediction is not as clinically motivating, unless they share the same developmental origin (for example, gastric and colon cancers). Achieving broader cross-cancer generalizability would require a foundation model kind of paradigm, with large-scale pan-cancer pretraining. However, the size of paired

histology and spatial transcriptomics datasets across cancer types is currently still very modest. For example, HEST-1K[72] includes fewer than 350 cancer samples in total, with limited representation of various cancer types (often fewer than 10 samples or none). In addition, variability in the quality of paired H&E images may affect the performance of the foundation model. Nevertheless, it could be a future direction as sample sizes continue to grow and data quality continues to improve.

In conclusion, HisToSpatialCNV represents a substantial advancement in spatial CNV prediction, with high accuracy, easy interpretability and immediate clinical relevance. By leveraging graph neural networks and attention mechanisms, the model effectively captures both molecular and spatial patterns, providing a comprehensive framework for studying tumour evolution and heterogeneity. Applying HisToSpatialCNV to population-level H&E images of cancer samples will likely uncover new patient subtypes with distinct phenotypes.

Methods

Data collection

All datasets used in this study are publicly available and can be accessed through the links provided in the Data availability section.

Spatial transcriptomics datasets and preprocessing

To evaluate the performance of our model in predicting copy number variations, we utilized two human spatial transcriptomics datasets. These datasets include whole-slide images of H&E-stained tissue with spatial coordinate information and corresponding gene expression data. Specifically, we used the HER2+ breast tumour dataset[25] with 26 selected slides from 5 patients, and another cutaneous squamous cell carcinoma[26] dataset with 9 selected slides from 3 patients. We conducted preprocessing to filter out genes with low variability, and then intersected genes common across all samples within each dataset. As a result, 2,562 genes remained in the HER2+ dataset and 540 genes were kept in the cSCC dataset. Since pathologist-annotated CNV data are unavailable for these datasets, we inferred CNV status using InferCNV[17] as the 'ground truth'. Spots with gene counts below 5% were filtered out before analysis. The count and annotation files were input into the inferCNV tool to infer CNVs.

For the HER2+ and cSCC datasets, we extracted 224 × 224-pixel patches centred on the spatial coordinates of each spot. To ensure data quality, we applied two filtering criteria: removing spots containing fewer than 300 detected genes and eliminating genes that showed no mutations across all spots in the dataset. We then reordered the remaining genes according to their genomic coordinates for subsequent analysis.

inferCNV preprocessing and settings

For CNV inference, we used immune infiltrate as the reference set. The run() function was executed with the following parameter values: cut-off = 0.1, cluster_by_groups = TRUE, denoise = TRUE, HMM = TRUE, write_expr_matrix = TRUE and HMM_type = 'i3'.

Cellprofiler settings

CellProfiler[20] was used to extract features from H&E images. It begins with the 'UnmixColors' module, which separates histological stains for enhanced visualization of tissue components. The 'IdentifyPrimaryObjects' module identifies cellular structures, followed by the 'IdentifySecondaryObjects' module which detects cell bodies surrounding the identified primary objects. A comprehensive feature extraction pipeline is implemented with modules, including 'MeasureObjectSizeShape' for quantifying morphological characteristics, 'MeasureTexture' for analysing textural patterns within objects, 'MeasureColocalization' for evaluating spatial relationships between different markers, 'MeasureGranularity' for assessing granular patterns at different scales, and 'MeasureImageAreaOccupied' for calculating the fraction of image occupied by identified objects. This approach generated 2,952 interpretable features, providing insights into the biological and morphological characteristics for analysis tasks.

HisToSpatialCNV

The proposed model HisToSpatialCNV architecture, illustrated in Fig. 1, introduces a novel sequential refinement approach specifically designed for multiclass CNV prediction. Unlike existing methods that process raw image patches directly, our framework begins with biologically interpretable morphological features extracted from 224 × 224-pixel histopathology images using CellProfiler. This ensures that subsequent predictions remain grounded in meaningful cellular and tissue characteristics. These features are mapped into a 1,024-dimensional embedding space using a 3-layer residual multilayer perceptron with LayerNorm and ReLU activation. This step refines the raw image features into embeddings that preserve spatial context, making them suitable for graph-based learning. Simultaneously, CNV data along with spatial coordinates are used to construct a weighted adjacency matrix. The mapped image features and spot location information are then fused and passed into the GNN module, which employs jumping knowledge networks with LSTM aggregation to preserve intermediate representations from each layer. Node features are iteratively updated by aggregating information from neighbouring nodes via the adjacency matrix, with the LSTM dynamically learning to fuse multiscale spatial features. This process effectively captures both local neighbourhood information and higher-order spatial dependencies, encoding relational context into the node features. Following GNN processing, a multihead self-attention mechanism is integrated to capture global dependencies and complex node interactions. Unlike existing approaches that use transformers for initial feature extraction, this mechanism operates on graph-refined features, allowing each node to globally attend to others with spatial context and enabling adaptive weighting of distant relationships. Finally, the prediction head employs a normalized MLP to output CNV predictions across three categories for each spatial location.

Feature extraction module

Depending on the size and complexity of the dataset, HisToSpatialCNV provides three types of feature extraction module: the CellProfiler[20] module, the ResNet-50 (ref. 42) module and the UNI[43] module. The CellProfiler module is suitable for smaller datasets, while the ResNet-50 module is better suited for larger datasets, and the UNI module is optimized for medical domains.

Unlike deep-learning models such as ResNet-50, which require extensive datasets to generalize effectively, CellProfiler can extract meaningful features from smaller datasets. Its extracted features, such as size, shape, intensity and texture, have direct physical interpretations, making the results explicit and interpretable. In addition, CellProfiler employs deterministic algorithms, ensuring consistency and reproducibility across runs. Its transparency facilitates the identification of important features, helping to determine which features are closely linked to CNV events.

Residual multilayer perceptron module

We introduced a ResMLP module, which consists of a three-layer residual multilayer perceptron equipped with layer normalization (LayerNorm) and ReLU activation functions. Residual connections enable the network to learn identity mappings, effectively mitigating the vanishing gradient problem often encountered in deep neural networks and facilitating the training of deeper architectures. Compared to traditional MLPs, the residual connections in this module enhance its ability to learn rich feature representations while maintaining efficient training dynamics. These connections enable the network to learn modifications to the input features, rather than

completely relearning them. The formula for the residual connection is shown below:

$$\mathbf{h}_1 = \mathbf{x} + \mathrm{ReLU}\left(\mathrm{LayerNorm}\left(W_1\mathbf{x} + \mathbf{b}_1\right)\right) \quad (1)$$

$$\mathbf{h}_2 = \mathbf{h}_1 + \mathrm{ReLU}\left(\mathrm{LayerNorm}\left(W_2\mathbf{h}_1 + \mathbf{b}_2\right)\right) \quad (2)$$

$$\mathbf{z} = W_3\mathbf{h}_2 + \mathbf{b}_3 \quad (3)$$

where: $\mathbf{x} \in \mathbb{R}^d$ is the input feature vector, $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{R}^d$ are the hidden representations, $\mathbf{z} \in \mathbb{R}^{1024}$ is the output embedding, $W_1, W_2 \in \mathbb{R}^{d\times d}$ are weight matrices for the hidden layers, $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^d$ are bias vectors for the hidden layers, $W_3 \in \mathbb{R}^{1024\times d}$ and $\mathbf{b}_3 \in \mathbb{R}^{1024}$ are the weight matrix and bias vector for the output layer, $\mathrm{LayerNorm}(\cdot)$ denotes the layer normalization operation, and $\mathrm{ReLU}(\cdot)$ is the rectified linear unit activation function.

Graph neural network

To effectively capture CNV relationships between spots, we introduced a GNN architecture designed to model both the local and global structures of the graph through multilevel feature aggregation. The network uses image feature matrices as node attributes and the adjacency matrix to define the connectivity between nodes. By integrating flexible neighbourhood aggregation, deep stacking and adaptive layer-wise feature integration via a jumping knowledge network with an LSTM aggregator, the model learns complex relationships among nodes by fusing multilevel information. This approach enhances the representation capacity of image features embedded within the graph structure, enabling more accurate CNV predictions.

We first constructed a spatial adjacency matrix based on the Euclidean distances between nodes, represented by their 2D coordinates. This matrix defines connections between nodes within a specified distance or proximity, allowing the creation of a graph structure that reflects spatial relationships. More specifically, we first calculated the distances between one node and all other nodes, forming a distance matrix, which can be expressed as:

$$\mathrm{distMat}_{i,j} = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \quad (4)$$

where p_x, p_y and q_x, q_y are the x and y coordinates of nodes p and q, respectively. The adjacency matrix incorporates both k-nearest neighbours and distance thresholding. For each node, we first identified its k-nearest neighbours as:

$$\mathrm{res}[i] = \mathrm{argsort}\left(\mathrm{distMat}_{i,}\right)[:\, k+1] \quad (5)$$

where k represents selected nearest neighbours. In addition, we applied a distance threshold to ensure that connections are spatially meaningful:

$$\mathrm{Adj}_{i,j} = \begin{cases} 1, \text{if } \mathrm{distMat}_{i,j} \le 2.0 \text{ and } j \in \mathrm{res}[i] \\ 0, \text{otherwise} \end{cases} \quad (6)$$

This dual criterion ensures that connections in the graph represent meaningful spatial relationships, balancing both proximity and distance constraints. Once the graph structure is established through the adjacency matrix, we can leverage this spatial information within our graph neural network architecture. The constructed adjacency matrix encodes the neighbourhood relationships that will guide the message-passing process, allowing information to flow between spatially proximal nodes. The node feature matrix and adjacency matrix are used as inputs to the graph convolutional layer. This layer aggregates neighbourhood features using a mean pooling strategy, combining and transforming node features based on the structural encoding represented by the adjacency matrix. The operation can be effectively expressed as:

$$H_{\mathrm{neigh}}^{(l)} = \mathrm{Aggregate}\left(H^{(l-1)}, A\right) \quad (7)$$

$$H_{\mathrm{Combined}}^{(l)} = \mathrm{Combine}\left(H^{(l-1)}, H_{\mathrm{neigh}}^{(l)}\right) \quad (8)$$

$$H^{(l)} = \mathrm{ReLU}\left(H_{\mathrm{Combined}}^{(l)} W^{(l)^{\top}}\right) \quad (9)$$

where l denotes the number of layers, $A \in \mathbb{R}^{N\times N}$ represents the adjacency matrix, $H^{(l)} \in \mathbb{R}^{N\times E}$ is the node feature matrix at layer l, and $W^{(l)} \in \mathbb{R}^{E\times D}$ is the weight matrix for layer l. After processing through all layers, we collected the outputs from each layer to capture multiscale representations. For each node i, the sequence of its representations across layers was treated as a time series and passed through an LSTM network to learn how to combine these representations. The LSTM process can be expressed as:

$$S_i = \mathrm{LSTM}\left(\mathbf{h}_i^{(1)}, \mathbf{h}_i^{(2)}, \ldots, \mathbf{h}_i^{(L)}\right) \quad (10)$$

where $S_i \in \mathbb{R}^{L\times E}$ represents the hidden states of the LSTM for node i, and $\mathbf{h}_i^{(l)} \in \mathbb{R}^E$ is the feature vector of node i at layer l. Finally, the embedding for each node is obtained by averaging the hidden states of the LSTM across all layers:

$$\mathbf{h}_i^{\mathrm{final}} = \frac{1}{L}\sum_{l=1}^{L} S_i^{(l)} \quad (11)$$

where $S_i^{(l)} \in \mathbb{R}^E$ represents the hidden state for node i at the lth step of the LSTM processing.

Multihead self-attention mechanism

After the GNN aggregates information from immediate neighbours and updates node features to capture local structural information, we applied the MHSA mechanism. This mechanism enhances the node feature representations by capturing global dependencies across the graph. To fully leverage the complex relationships inherent in the graph structure, it is crucial to allow nodes to attend to all other nodes. The MHSA mechanism enables each node to weigh and aggregate information from every other node, capturing remote interactions and global context, as described by:

$$\mathrm{MHSA}(Q, K, V) = \mathrm{Concat}\left(\mathrm{head}_1, \ldots, \mathrm{head}_h\right) W^O \quad (12)$$

where each attention head is computed as:

$$\mathrm{head}_i = \mathrm{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right) \quad (13)$$

and the attention mechanism is defined as:

$$\mathrm{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right) = \mathrm{Softmax}\left(\frac{\left(QW_i^Q\right)\left(KW_i^K\right)^{\top}}{\sqrt{d_k}}\right) VW_i^V \quad (14)$$

where Q, K and V are the query, key and value matrices, respectively, and h represents the number of attention heads. W^O is the output projection matrix, $\mathrm{Concat}(\cdot)$ denotes the concatenation of the outputs from all h heads, W_i^Q, W_i^K, W_i^V are weight matrices, and d_k is the dimension of the key vectors.

Loss function and training strategy

In this study, we addressed the issue of imbalanced CNV data by employing the focal loss function[24], which is specifically designed to mitigate

class imbalance in classification tasks. CNV datasets often exhibit severe class imbalance, with certain CNV states being more prevalent than others. The focal loss function reduces the impact of class imbalance by emphasizing harder-to-classify samples. It is mathematically defined as:

$$\text{Focal Loss}\,(p_t) = -\alpha_t(1-p_t)^{\gamma}\log\,(p_t) \tag{15}$$

where p_t denotes the model's estimated probability for the true class label, α_t represents a weighting factor for the class t, and γ is the focusing parameter. In addition, we employed a LOOCV strategy to enhance the robustness and generalizability of our model. In LOOCV, one of the N slides is left out each time to serve as the testing dataset, while the remaining slides are used for training. The imputed CNV data are then compared to the original ST data. This method is particularly effective for small datasets.

Performance comparison with other models

Model performance was evaluated using a leave-one-out cross-validation strategy at the tissue-section level due to the limited number of available spatial transcriptomics samples. In each LOOCV iteration, one tissue section was held out as an independent test sample, while all the other sections were used for model training. Model architectures and hyperparameters were fixed before LOOCV iterations. Performance was primarily assessed using the AUC and the MCC. Metrics were computed for each held-out tissue section and subsequently aggregated across all LOOCV iterations to summarize overall model performance. Compared methods included ST-Net[27], HisToGene[28], Hist2ST[29], TCGN[30], mclSTExp[34], GHIST[31], iSCALE[32] and iSTAR[32], as well as the pretrained foundation model OmiCLIP[35]. For fair comparison, all task-specific baseline methods were evaluated under the same LOOCV data splits and evaluation protocol as those for HisToSpatialCNV. For the foundation model OmiCLIP[35], the default parameter setting was used. Since these methods use histopathology images to predict spatial gene expression rather than spatial CNV, their outputs were converted to CNV states using inferCNV, with identical preprocessing steps and parameter settings across all methods as well as those generating the CNV ground truth.

Application of HisToSpatialCNV to the TCGA cohort

To assess the applicability of HisToSpatialCNV beyond spatial transcriptomics datasets, we directly applied the model trained on the HER2+ Visium cohort[25] earlier to histopathology images from the TCGA cohort[82]. Specifically, we analysed 92 HER2+ breast cancer samples from a previously published TCGA subset in which whole-slide H&E images were manually annotated by pathologists to delineate tumour and non-tumour regions[64]. Whole-slide images were processed using the same preprocessing and patch extraction pipeline as used for spatial transcriptomics data. Briefly, H&E slides were divided into non-overlapping 224 × 224-pixel image patches, and features were extracted using the same feature extraction module as in the spatial datasets. These patches were treated as pseudo-spots to generate dense spatial CNV predictions across the tissue.

Predicted CNV states at each pseudo-spot were subsequently used for downstream analyses, including spatial embedding construction, patient stratification, pathway enrichment analysis and survival analysis, as described in the Results section. This strategy enabled the direct transfer of a HER2+-trained spatial CNV prediction model to a large independent cohort, facilitating exploratory analysis of spatial CNV patterns and their clinical relevance in breast cancer. Among the 92 TCGA HER2+ samples, 2 were removed by the single-patient cluster filter during K-means clustering and 9 lacked disease-free survival information, leaving 90 samples for spatial cluster analysis and 81 for Kaplan–Meier survival analysis.

Graph convolutional network (GCN)-based spatial domain identification

To identify spatial domains from predicted spatial CNV, we constructed a spatial graph for each slide using the Python package SpaGCN[83]. CNV data were first log normalized, followed by principal component analysis (PCA). The resulting PCA scores were used as graph node features, while spatial coordinates defined the adjacency matrix as graph edges. This graph was input into a GCN trained with a predefined number of clusters. A spatially informed embedding was extracted from the trained GCN and used for Leiden clustering[84], yielding spatial domain partitions.

Validation against CNV ground truth measured from whole-exome sequencing

To validate HisToSpatialCNV predictions against direct genomic measurements, we used the preprocessed data from 2 breast cancer patients (HT206B1 and HT268B1) in the HTAN (WashU cohort), shared by Ma et al. who developed CalicoST[19]. For these patients, whole-exome sequencing data from tumour and matched adjacent tissues were available, along with the H&E histopathology data. We used HATCHet2 (ref. 85) to infer clone-specific copy number states, which were discretized into 3 categories: deletion (total copy number < 1), neutral (total copy number = 1) and amplification (total copy number > 1). Since WES provides bulk-level CNV profiles without spatial localization, we applied K-means clustering to the predicted CNV profiles to delineate spatially coherent regions. We then used the Hungarian algorithm to optimally match predicted clusters to WES-derived clones, based on L2 distance, as done in ref. 19. Clusters corresponding to normal tissue were excluded from evaluation. Predictive performance was assessed using macro-average AUCs and F1 scores across the 3 CNV classes.

Functional downstream analysis

For downstream analysis of the predicted spatial CNV, we used the R package 'TUSCAN'[86] to perform tumour versus non-tumour spot identification. Spatial domain identification was conducted using a GNN-based approach. We constructed phylogenetic trees using the R package 'phangorn'[87] with the parsimony method to infer clonal relationships. Differential CNV analysis between subclones was performed using Fisher's exact test[88], with significant genes filtered using thresholds of $P < 0.05$ and odds ratio > 1.5 to identify cancer-associated genes with subclone-specific alterations. For pathway analysis, we conducted gene set enrichment analysis (GSEA) using the MSigDB Hallmark[89] gene set for both HER2+ breast cancer and cSCC datasets. Additional functional enrichment analysis for TCGA data integration was performed using the Python package 'erichr'[90].

Evaluation metrics

We employed the AUC to evaluate the model's ability to discriminate between CNV classes and adopted the MCC to evaluate the model's classification quality. AUC values range from 0 to 1, allowing evaluation of the model's performance across all classification thresholds. MCC values range from −1 to +1 and offer a correlation-based evaluation sensitive to all types of error, making them suitable for assessing overall classification quality. MCC can be expressed as follows:

$$\text{MCC} = \frac{c \times s - \sum_{k=1}^{K} p_k \times t_k}{\sqrt{\left(s^2 - \sum_{k=1}^{K} p_k^2\right) \times \left(s^2 - \sum_{k=1}^{K} t_k^2\right)}} \tag{16}$$

where K denotes number of classes, s represents the total number of samples, c represents the total number of correctly classified samples, p_k represents predicted samples for class k, and t_k represents actual samples for class k.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data used in this study are publicly available. The human HER2+ breast tumour spatial transcriptomics data, including images and spatial information used in this study, are available in GitHub at https://github.com/almaan/her2st. The human cutaneous squamous cell carcinoma 10x Visium data, including images and spatial information used in this study, are available from the Gene Expression Omnibus (GEO) under accession number GSE144240 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144240). The human glioblastoma Visium data are available at https://www.gbmspace.org/. The breast cancer whole-exome sequencing (WES) data from the HTAN cohort are available via Zenodo at https://doi.org/10.5281/zenodo.14175627 (ref. 91). The HER2+ breast cancer patient samples generated on the Xenium platform are available from GEO under accession number GSE243280 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE243280) and from the 10x Genomics Datasets portal (https://www.10xgenomics.com/datasets/xenium-ffpe-human-breast-biomarkers). The TCGA breast cancer cohort data used for validation can be accessed from the GDC Data Portal at https://portal.gdc.cancer.gov/projects/TCGA-BRCA. The inferred CNV data used as ground-truth input to train HisToSpatialCNV generated in this study are available via Zenodo at https://doi.org/10.5281/zenodo.20141481 (ref. 92). Source data are provided with this paper.

Code availability

The software used to generate the results reported in this study is the subject of a pending patent application (U.S. provisional patent application number 63/875) filed by The University of Michigan 3 September 2025 in the United States. The source code is not currently publicly available. Any inquiries expressing interest in this technology should be directed towards the University of Michigan's technology transfer office. The code can be available to academic researchers for non-commercial research purposes upon reasonable request.

References

1. Mills, S. *Histology for Pathologists* (Lippincott Williams & Wilkins, 2012).
2. Feldman, A. T. & Wolfe, D. Tissue processing and hematoxylin and eosin staining. *Methods Mol. Biol.* **1180**, 31–43 (2014).
3. Lupski, J. R. et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell* **66**, 219–232 (1991).
4. Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54 (2016).
5. Mamlouk, S. et al. DNA copy number changes define spatial patterns of heterogeneity in colorectal cancer. *Nat. Commun.* **8**, 14093 (2017).
6. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202–209 (2014).
7. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
8. Litchfield, K. et al. Whole-exome sequencing reveals the mutational spectrum of testicular germ cell tumours. *Nat. Commun.* **6**, 5973 (2015).
9. Inman, G. J. et al. The genomic landscape of cutaneous SCC reveals drivers and a novel azathioprine associated mutational signature. *Nat. Commun.* **9**, 3667 (2018).
10. Beroukhim, R. et al. The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905 (2010).
11. Davoli, T. et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**, 948–962 (2013).
12. Dentro, S. C. et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. *Cell* **184**, 2239–2254 (2021).
13. Erickson, A. et al. Spatially resolved clonal copy number alterations in benign and malignant tissue. *Nature* **608**, 360–367 (2022).
14. McGranahan, N. & Swanton, C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* **168**, 613–628 (2017).
15. Buysse, K. et al. Challenges for CNV interpretation in clinical molecular karyotyping: lessons learned from a 1001 sample experience. *Eur. J. Med. Genet.* **52**, 398–403 (2009).
16. Carter, N. P. Methods and strategies for analyzing copy number variation using DNA microarrays. *Nat. Genet.* **39**, S16–S21 (2007).
17. Tickle, T., Tirosh, I., Georgescu, C., Brown, M. & Haas, B. *inferCNV of the Trinity CTAT Project* (Klarman Cell Observatory, Broad Institute of MIT and Harvard, 2019).
18. Elyanow, R., Zeira, R., Land, M. & Raphael, B. J. STARCH: copy number and clone inference from spatial transcriptomics data. *Phys. Biol.* **18**, 035001 (2021).
19. Ma, C. et al. Inferring allele-specific copy number aberrations and tumor phylogeography from spatially resolved transcriptomics. *Nat. Methods* **21**, 2239–2247 (2024).
20. Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
21. Xu, K. et al. Representation learning on graphs with jumping knowledge networks. In *International Conference on Machine Learning* 5453–5462 (PMLR, 2018).
22. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
23. Vaswani, A. et al. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **30**, 5998–6008 (2017).
24. Lin, T.-Y., Goyal, P., Girshick, R., He, K. & Dollár, P. Focal loss for dense object detection. In *Proc. IEEE International Conference on Computer Vision* 2980–2988 (IEEE, 2017).
25. Andersson, A. et al. Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. *Nat. Commun.* **12**, 6012 (2021).
26. Ji, A. L. et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* **182**, 497–514 (2020).
27. He, B. et al. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nat. Biomed. Eng.* **4**, 827–834 (2020).
28. Pang, M., Su, K. & Li, M. Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. Preprint at *bioRxiv* https://doi.org/10.1101/2021.11.28.470212 (2021).
29. Zeng, Y. et al. Spatial transcriptomics prediction from histology jointly through transformer and graph neural networks. *Brief. Bioinform.* **23**, bbac297 (2022).
30. Xiao, X., Kong, Y., Li, R., Wang, Z. & Lu, H. Transformer with convolution and graph-node co-embedding: an accurate and interpretable vision backbone for predicting gene expressions from local histopathological image. *Med. Image Anal.* **91**, 103040 (2024).
31. Fu, X. et al. Spatial gene expression at single-cell resolution from histology using deep learning with GHIST. *Nat. Methods* **22**, 1900–1910 (2025).
32. Schroeder, A. et al. Scaling up spatial transcriptomics for large-sized tissues: uncovering cellular-level tissue architecture beyond conventional platforms with iSCALE. *Nat. Methods* **22**, 1911–1922 (2025).

33. Zhang, D. et al. Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. *Nat. Biotechnol.* **42**, 1372–1377 (2024).
34. Min, W., Shi, Z., Zhang, J., Wan, J. & Wang, C. Multimodal contrastive learning for spatial gene expression prediction using histology images. *Brief. Bioinform.* **25**, bbae551 (2024).
35. Chen, W. et al. A visual–omics foundation model to bridge histopathology with spatial transcriptomics. *Nat. Methods* **22**, 1568–1582 (2025).
36. Stranger, B. E. et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
37. Tang, X. et al. Mixed ductal–lobular carcinoma: an analysis of CDH1 DNA copy number variation and mutation. *Breast Cancer* **28**, 1318–1327 (2021).
38. Gao, Y. et al. MIIP functions as a novel ligand for ITGB3 to inhibit angiogenesis and tumorigenesis of triple-negative breast cancer. *Cell Death Dis.* **13**, 810 (2022).
39. Kothapalli, K. S., Park, H. G., Kothapalli, N. S. & Brenna, J. T. FADS2 function at the major cancer hotspot 11q13 locus alters fatty acid metabolism in cancer. *Prog. Lipid Res.* **92**, 101242 (2023).
40. Mehta, G. A. et al. Amplification of SOX4 promotes PI3K/Akt signaling in human breast cancer. *Breast Cancer Res. Treat.* **162**, 439–450 (2017).
41. Zhang, J. et al. SOX4 promotes the growth and metastasis of breast cancer. *Cancer Cell Int.* **20**, 468 (2020).
42. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 770–778 (IEEE, 2016).
43. Chen, R. J. et al. Towards a general-purpose foundation model for computational pathology. *Nat. Med.* **30**, 850–862 (2024).
44. Sundararajan, M., Taly, A. & Yan, Q. Axiomatic attribution for deep networks. In *International Conference on Machine Learning* 3319–3328 (PMLR, 2017).
45. Gao, S., Gao, X., Wang, N., Xie, Y. & Tang, Y. MYO9B deficiency promoted head and neck tumor progression through HIF1α/MYC/STAT2 signaling pathway. *Biochim. Biophys. Acta Gen. Subj.* **1869**, 130834 (2025).
46. Kong, R. et al. Myo9b is a key player in SLIT/ROBO-mediated lung tumor suppression. *J. Clin. Invest.* **125**, 4407–4420 (2015).
47. Tian, X., Yang, C., Yang, L., Sun, Q. & Liu, N. PTPRF as a novel tumor suppressor through deactivation of ERK1/2 signaling in gastric adenocarcinoma. *Onco Targets Ther.* **11**, 7795–7803 (2018).
48. Lv, Z., Wang, T., Cao, X., Sun, M. & Qu, Y. The role of receptor-type protein tyrosine phosphatases in cancer. *Precis. Med. Sci.* **12**, 48–57 (2023).
49. Du, W. W. et al. MicroRNA miR-24 enhances tumor invasion and metastasis by targeting PTPN9 and PTPRF to promote EGF signaling. *J. Cell Sci.* **126**, 1440–1453 (2013).
50. Yang, Y. et al. SLC44A2 negatively regulates mitochondrial fatty acid oxidation to suppress colorectal progression by blocking the MUL1-CPT2 interaction. *Cell Death Dis.* **16**, 468 (2025).
51. Patysheva, M. R., Prostakishina, E. A., Budnitskaya, A. A., Bragina, O. D. & Kzhyshkowska, J. G. Dual-specificity phosphatases in regulation of tumor-associated macrophage activity. *Int. J. Mol. Sci.* **24**, 17542 (2023).
52. Hao, L. & ElShamy, W. M. BRCA1-IRIS activates cyclin D1 expression in breast cancer cells by downregulating the JNK phosphatase DUSP3/VHR. *Int. J. Cancer* **121**, 39–46 (2007).
53. Chou, H.-C. et al. DUSP3 regulates phosphorylation-mediated degradation of occludin and is required for maintaining epithelial tight junction. *J. Biomed. Sci.* **29**, 40 (2022).
54. Yang, X. et al. The pseudokinase NRBP1 activates Rac1/Cdc42 via P-Rex1 to drive oncogenic signalling in triple-negative breast cancer. *Oncogene* **42**, 833–847 (2023).
55. Ma, X. et al. ErbB2-upregulated HK1 and HK2 promote breast cancer cell proliferation, migration and invasion. *Med. Oncol.* **40**, 154 (2023).
56. Jiao, X. et al. c-Jun induces mammary epithelial cellular invasion and breast cancer stem cell expansion. *J. Biol. Chem.* **285**, 8218–8226 (2010).
57. Zhang, X., Zhou, L., Wei, S. & Zhang, X. Comprehensive analysis of TMEM9 in human tumors. *Discov. Oncol.* **16**, 276 (2025).
58. Morgenstern-Kaplan, D. et al. Genomic, immunologic, and prognostic associations of TROP2 (TACSTD2) expression in solid tumors. *Oncologist* **29**, e1480–e1491 (2024).
59. Ding, J. et al. Inhibition of BTF3 sensitizes luminal breast cancer cells to PI3Kα inhibition through the transcriptional regulation of ERα. *Cancer Lett.* **440**, 54–63 (2019).
60. Buyuk, B., Jin, S. & Ye, K. Epithelial-to-mesenchymal transition signaling pathways responsible for breast cancer metastasis. *Cell. Mol. Bioeng.* **15**, 1–13 (2022).
61. Huang, F. et al. HER2/EGFR–AKT signaling switches TGFβ from inhibiting cell proliferation to promoting cell migration in breast cancer. *Cancer Res.* **78**, 6073–6085 (2018).
62. Shihab, I. et al. Understanding the role of innate immune cells and identifying genes in breast cancer microenvironment. *Cancers* **12**, 2226 (2020).
63. Danforth, D. N. The role of chronic inflammation in the development of breast cancer. *Cancers* **13**, 3918 (2021).
64. Farahmand, S. et al. HER2 and trastuzumab treatment response H&E slides with tumor ROI annotations. *Cancer Imaging Arch.* https://doi.org/10.7937/E65C-AM96 (2022).
65. Amin, M. B. et al. The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more 'personalized' approach to cancer staging. *CA Cancer J. Clin.* **67**, 93–99 (2017).
66. Lineham, E., Tizzard, G. J., Coles, S. J., Spencer, J. & Morley, S. J. Synergistic effects of inhibiting the MNK-eIF4E and PI3K/AKT/mTOR pathways on cell migration in MDA-MB-231 cells. *Oncotarget* **9**, 14148 (2018).
67. Fu, J. et al. Impact of BSG/CD147 gene expression on diagnostic, prognostic and therapeutic strategies towards malignant cancers and possible susceptibility to SARS-CoV-2. *Mol. Biol. Rep.* **50**, 2269–2281 (2023).
68. Carr, D. et al. Multiple roles for Bcl-3 in mammary gland branching, stromal collagen invasion, involution and tumor pathology. *Breast Cancer Res.* **24**, 40 (2022).
69. Mo, C.-K. et al. Tumour evolution and microenvironment interactions in 2D and 3D space. *Nature* **634**, 1178–1186 (2024).
70. Janesick, A. et al. High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat. Commun.* **14**, 8353 (2023).
71. Janesick, A., Kravitz, S. N., Stauffer, W., Valencia, M. & Taylor, S. E. Biomarker quantification in breast cancer using xenium in situ. Preprint at *bioRxiv* https://doi.org/10.64898/2025.12.08.692193 (2025).
72. Jaume, G. et al. Hest-1k: a dataset for spatial transcriptomics and histology image analysis. *Adv. Neural Inf. Process. Syst.* **37**, 53798–53833 (2024).
73. de Jong, G. et al. A spatiotemporal cancer cell trajectory underlies glioblastoma heterogeneity. Preprint at *bioRxiv* https://doi.org/10.1101/2025.05.13.653495 (2025).
74. Kedzierska, K. Z., Crawford, L., Amini, A. P. & Lu, A. X. Zero-shot evaluation reveals limitations of single-cell foundation models. *Genome Biol.* **26**, 101 (2025).
75. Roohani, Y., Huang, K. & Leskovec, J. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nat. Biotechnol.* **42**, 927–935 (2024).

76. Ahlmann-Eltze, C., Huber, W. & Anders, S. Deep-learning-based gene perturbation effect prediction does not yet outperform simple linear baselines. *Nat. Methods* **22**, 1657–1661 (2025).
77. Xie, F. et al. Overcoming barriers to the wide adoption of single-cell large language models in biomedical research. *Nat. Biotechnol.* **43**, 1758–1762 (2025).
78. Amand, M. et al. DUSP3/VHR is a pro-angiogenic atypical dual-specificity phosphatase. *Mol. Cancer* **13**, 108 (2014).
79. Schmid, K. T. et al. Benchmarking scRNA-seq copy number variation callers. *Nat. Commun.* **16**, 8777 (2025).
80. Bařinka, J. et al. RNAseqCNV: analysis of large-scale copy number variations from RNA-seq data. *Leukemia* **36**, 1492–1498 (2022).
81. Zhao, J. et al. Interpretable, flexible and spatially aware integration of multiple spatial transcriptomics datasets from diverse sources. *Nat. Genet.* **58**, 1138–1150 (2026).
82. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
83. Hu, J. et al. SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods* **18**, 1342–1351 (2021).
84. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
85. Myers, M. A. et al. HATCHet2: clone- and haplotype-specific copy number inference from bulk tumor sequencing data. *Genome Biol.* **25**, 130 (2024).
86. Zang, C., Guo, C. C., Wei, P. & Li, Z. TUSCAN: tumor segmentation and classification analysis in spatial transcriptomics. *PLoS Comput. Biol.* **22**, e1014058 (2026).
87. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011).
88. Fisher, R. A. On the interpretation of χ2 from contingency tables, and the calculation of P. *J. R. Stat. Soc.* **85**, 87–94 (1922).
89. Liberzon, A. et al. The molecular signatures database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
90. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
91. Ma, C. et al, Inferring allele-specific copy number aberrations and tumor phylogeography from spatially resolved transcriptomics (output data), *Zenodo*. https://doi.org/10.5281/zenodo.14175627 (2024).
92. Chen, T. et al, HisToSpatialCNV inferred CNV data, *Zenodo*. https://doi.org/10.5281/zenodo.20141481 (2026).

Acknowledgements

We thank H. Zhang and all members of the Garmire Group for helpful discussions; C. Ma from the University of Michigan for providing bulk whole-exome sequencing (WES) data used to derive the experimental CNV ground truth; and R. Fan from Yale University for valuable discussions regarding the validity of bulk WES data as the alternative ground truth. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

L.X.G. conceived this project and supervised the study. T.C. wrote the code of the model, carried out most of the analysis, and wrote the manuscript. T.U., X.G., P.T, S.Z., Y.Y., Y.D. and Z.J. assisted in the analysis. T.U., P.T. and X.G. also helped with manuscript writing and figure generation. W.Z. assisted in biological interpretation. All authors read and approved the final manuscript.

Funding

L.X.G. was supported by grants from NLM R01 LM012373 and NIH R03 OD039978. T.U. was supported by a fellowship from the Royal Thai government.

Competing interests

L.X.G. and some co-authors are inventors on a pending patent application related to the methods described in this manuscript.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41551-026-01754-z.

Correspondence and requests for materials should be addressed to Lana X. Garmire.

Peer review information *Nature Biomedical Engineering* thanks Andrew Erickson, Konstantin Khodosevich and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1]Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI, USA. [2]Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, USA. [3]Department of Biomedical Informatics and Data Science, University of Alabama at Birmingham, Birmingham, AL, USA. [4]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA. [5]Department of Applied Statistics, University of Michigan, Ann Arbor, MI, USA. [6]Departments of Surgery and Pathology, Center of Excellence for Cancer Immunology and Immunotherapy, University of Michigan Rogel Cancer Center, Ann Arbor, MI, USA. ✉e-mail: lgarmire@uab.edu

nature portfolio

Corresponding author(s): Lana X. Garmire

Last updated by author(s): May 11, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	CellProfiler v4.2.6 ST-Net (https://github.com/bryanhe/ST-Net) HisToGene (https://github.com/maxpmx/HisToGene) Hist2ST (https://github.com/biomed-AI/Hist2ST) TCGN (https://github.com/lugia-xiao/TCGN) mclSTExp (https://github.com/shizhiceng/mclSTExp) OmiCLIP / Loki (https://github.com/GuangyuWangLab2021/Loki) GHIST (https://github.com/SydneyBioX/GHIST) iSCALE (https://github.com/amesch441/iSCALE) iSTAR (https://github.com/daviddaiweizhang/istar)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

We analyzed the following datasets in this study: (1) The Human HER2-positive breast tumor spatial transcriptomics (ST) data (https://github.com/almaan/her2st). (2) The Human cutaneous squamous cell carcinoma 10x Visium data (GSE144240: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144240). (3) The TCGA breast cancer HER2-positive cohort data (https://portal.gdc.cancer.gov/projects/TCGA-BRCA). (4) The Human glioblastoma Visium data (https://www.gbmspace.org/). (5) Breast cancer whole-exome sequencing (WES) data of the HTAN cohort (https://zenodo.org/records/14175627). (6) HER2+ breast cancer patient samples done on Xenium platform (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE243280 and https://www.10xgenomics.com/datasets/xenium-ffpe-human-breast-biomarkers)

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	HisToSpatialCNV was evaluated across six independent datasets, including: (1) HER2-positive breast tumor spatial transcriptomics (ST) data (n = 32), (2) cutaneous squamous cell carcinoma spatial transcriptomics data (n = 9), (3) the TCGA breast cancer HER2-positive cohort (n = 92), (4) human glioblastoma Visium data (n = 45), (5) breast cancer whole-exome sequencing (WES) data from the HTAN cohort (n = 9), and (6) HER2+ breast cancer patient samples generated using the Xenium platform (n = 3).
Data exclusions	Spots with gene counts below 5% were excluded prior to analysis. Spots containing fewer than 300 detected genes and genes showing no mutations across all spots within a dataset were also removed during preprocessing. Among the 92 TCGA HER2+ samples, 2 samples were excluded during K-means clustering because they formed single-patient clusters, leaving 90 samples for spatial clustering analysis. An additional 9 samples lacking disease-free survival information were excluded from Kaplan–Meier analysis, leaving 81 samples for survival analysis.
Replication	Model performance was evaluated by leave-one-out cross-validation at the tissue-section level on the HER2+ and cSCC datasets. Model was validated on 6 held-out HER2+ samples, the TCGA HER2+ cohort, WES-derived CNV ground truth, Xenium data, and a brain cancer Visium dataset.
Randomization	Train/test splits were determined by leave-one-out cross-validation or 5-fold cross-validation at the tissue-section level.
Blinding	N/A. Blinding was not relevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A