An image analysis pipeline to quantify the spatial distribution of cell markers in stroma-rich tumors

Antoine A. Ruzette[1], Nina Kozlova[2,3], Kayla A. Cruz[2,4], Taru Muranen[2,3] and Simon F. Nørrelykke[1]*

[1]Department of Systems Biology, Harvard Medical School, Boston, MA, United States, [2]Department of Medicine, Cancer Research Institute, Beth Israel Deaconess Medical Center, Boston, MA, United States, [3]Harvard Medical School, Boston, MA, United States, [4]Biological and Biomedical Sciences PhD Program, Harvard University, Boston, MA, United States

Aggressive cancers, such as pancreatic ductal adenocarcinoma (PDAC), are often characterized by a complex and desmoplastic tumor microenvironment, a stroma rich supportive connective tissue composed primarily of extracellular matrix (ECM) and non-cancerous cells. Desmoplasia, a dense deposition of stroma, is a major reason for therapy resistance, acting both as a physical barrier that interferes with drug penetration and as a supportive niche that protects cancer cells through diverse mechanisms. Precise understanding of spatial cell interactions in stroma-rich tumors is essential for optimizing therapeutic responses. It enables detailed mapping of stromal-tumor interfaces, comprehensive cell phenotyping, and insights into changes in tissue architecture, improving assessment of drug responses. Recent advances in multiplexed immunofluorescence imaging have enabled the acquisition of large batches of whole-slide tumor images, but scalable and reproducible methods to analyze the spatial distribution of cell states relative to stromal regions remain limited. To address this gap, we developed an open-source computational pipeline that integrates QuPath, StarDist, and custom Python scripts to quantify biomarker expression at a single- and sub-cellular resolution across entire tumor sections. Our workflow includes: (i) automated nuclei segmentation using StarDist, (ii) machine learning-based cell classification using multiplexed marker expression, (iii) modeling of stromal regions based on fibronectin staining, (iv) sensitivity analyses on classification thresholds to ensure robustness across heterogeneous datasets, and (v) distance-based quantification of the proximity of each cell to the stromal border. To improve consistency across slides with variable staining intensities, we introduce a statistical strategy that translates classification thresholds by propagating a chosen reference percentile across the distribution of marker-related cell measurement in each image. We apply this approach to quantify spatial patterns of distribution of the phosphorylated form of the N-Myc downregulated gene 1 (NDRG1), a novel DNA repair protein that conveys signals from the ECM to the nucleus to maintain replication

fork homeostasis, and a known cell proliferation marker Ki67 in fibronectin-defined stromal regions in PDAC xenografts. The pipeline is applicable for the analysis of markers of interest in stroma-rich tissues and is publicly available.

1 Introduction

Tumor microenvironment (TME) is a complex neighborhood that plays an active role in tumor progression, metastasis, immune evasion, and influences the efficacy of anti-cancer agents (Valkenburg, de Groot, and Pienta, 2018; Xu et al., 2022). TME consists of cancer cells, cancer-associated fibroblasts (CAFs), immune cells, endothelial cells, pericytes, neuronal cells, and adipocytes. Central to the TME is the stroma, a supportive framework composed primarily of connective tissue, extracellular matrix (ECM) proteins, blood vessels, fibroblasts, and immune cells. CAFs are known to secrete excessive amounts of ECM proteins in desmoplastic cancers, thus regulating the stromal density within a tumor. Desmoplasia is a hallmark of pancreatic ductal adenocarcinoma (PDAC), a malignancy marked by aggressive behavior, poor prognosis, and a dense, fibrotic stroma that limits drug penetration, fosters an immunosuppressive environment, and orchestrates signaling events leading to therapy resistance (Kozlova et al., 2020; Ho et al., 2020; Halbrook et al., 2023). Precise understanding of the spatial organization of the TME is essential for decoding the complexity of cell-stroma interactions, and for predicting treatment efficacies.

Each cell population within the TME can be defined by immunostaining with specific markers (Hu et al., 2023). While epithelial cells can be distinguished by positive staining of various cytokeratins (Karantza, 2011), the visualization of stromal compartment within a tumor can be done via Masson's Trichrome stain that stains for collagen and fibrin for brightfield imaging (Foot, 1933), or by immunofluorescent staining of matrix proteins such as collagens, fibronectin and laminins (Kozlova et al., 2020). Recent advances in multiplexed immunofluorescence imaging allow for high-resolution visualization of multiple biomarkers across entire tumor sections. In parallel, tools to analyze bioimages have emerged (Cimini et al., 2024). Several software platforms support general microscopy image analysis—including open-source tools such as QuPath (Bankhead et al., 2017), CellProfiler (Stirling et al., 2021b; Stirling et al., 2021a), Fiji (Schindelin et al., 2012; Schneider et al., 2012; Schindelin et al., 2015; Rueden et al., 2017), and Napari (Napari contributors, 2019), as well as commercial packages like HALO (Indica Labs, Albuquerque, NM, United States) and Visiopharm (Hørsholm, Denmark). While each offers strengths, most fall short in modeling spatial relationships between cell phenotypes and stromal structures across heterogeneous, multi-channel datasets. Commercial solutions provide turnkey pipelines but lack flexibility for custom analysis; open-source tools often require manual threshold tuning or lack built-in support for stromal modeling and spatial distance quantification. Moreover, ensuring reproducibility in the life sciences is crucial, and the price barrier of commercial solutions hampers accessibility and reproducibility.

Open-source workflows offer greater transparency but require careful attention to understand the sensitivity of methods used under image variability (Miura and Nørrelykke, 2021).

To address these challenges, we developed a robust, open-source pipeline for quantifying the spatial distribution of cell markers relative to stromal borders in tumor tissues. Built on QuPath (Bankhead et al., 2017), our workflow integrates nuclei segmentation with StarDist (Schmidt et al., 2018), machine learning-based cell classification at cellular and sub-cellular resolution, stromal region modeling, sensitivity analysis, and spatial distance measurements. We also introduce a statistical mapping method that translates intensity-based thresholds across heterogeneous images using percentile propagation in continuous distributions of cell measurements. The pipeline leverages Python-based tools for data aggregation and spatial visualization, facilitating seamless analysis of large, multi-image datasets.

Our previous study identified N-Myc downregulated gene 1 (NDRG1) as a novel DNA repair protein able to convey signals from the ECM to the nucleus to maintain replication fork homeostasis and mediate resistance to chemotherapies (Kozlova et al., 2025). In the present work, we apply our image analysis pipeline to multiplexed immunofluorescence images of chemotherapy-treated PDAC xenografts, focusing on the spatial distribution of phosphorylated NDRG1 (pNDRG1) and the proliferation marker Ki67 in cancer cells within stroma-rich tumor regions.

2 Methods

2.1 Xenograft studies, tissue processing and image acquisition

Pancreatic adenocarcinoma cell line AsPC was a kind gift from Dr. Nada Kalaany (Boston Children's Hospital, Boston, MA). All animal studies were performed according to protocols approved by the Institutional Animal Care and Use Committee at Beth Israel Deaconess Medical Center (BIDMC). Athymic male nude mice (NU/J) were purchased from Jackson labs (#002019). AsPC tumor cells (1×10^6) were injected subcutaneously in one flank per mouse in 1:1 mix of Matrigel and PBS when the mice were 8–12 weeks old. Tumor growth was monitored thrice weekly. Once the tumors reached 200 mm³ in volume, animals were treated with gemcitabine 25 mg/kg twice per week. The entire cohort of animals was sacrificed when the tumor size of at least one animal reached close to 1,000 mm³. Xenograft tumor tissues were processed, paraffin embedded and cut, deparaffinized in xylene and rehydrated in a descending ethanol series. Deparaffinized sections were subjected to antigen unmasking in SignalStain® Citrate Unmasking Solution

(CST #14746) followed by permeabilization in 1% Triton X-100, and blocking of nonspecific binding in TBST/5% normal goat serum solution. Sections were later stained with Pan-cytokeratin (AE1/AE3) Alexa Fluor™ 488 conjugated (53-9003-82, eBioscience), mouse anti-Fibronectin (ab6328-250, Abcam), and rabbit phospho-NDRG1 Thr346 (#5482S, CST) or Ki67 (D3B5) Rabbit mAb (Alexa Fluor® 647 Conjugate) (#12075, CST). Appropriate secondary antibodies were used (Goat anti-Mouse IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 568 A-11004 for Fibronectin, and Goat anti-Rabbit IgG (H + L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 647 #A-21245 for pNDRG1) and whole section images were acquired using Olympus BX-UCB whole slide scanner equipped with Olympus UPLSAPO 20x and VS-ASW v2.7 software, tile registration was performed using the scanner's proprietary software. Each image consisted of four fluorescence channels: DAPI (nuclei), FITC (pan-cytokeratin), TRITC (fibronectin), and CY5, labeling either phosphorylated NDRG1 (Datasets 1) or Ki67 (Datasets 2). Each field covered approximately 10,000 μm × 10,000 μm (30,000 × 30,000 pixels at 0.3215 μm per pixel) and contained roughly 100,000 cells per image. Datasets 1 (pNDRG1) contained a total of 496,936 cells across all images; Datasets 2 (Ki67) contained 454,556 cells. The methods described below are channel-agnostic and serve as a generalizable pipeline for multi-marker spatial analysis in tissue sections.

2.2 Modeling the stromal region and its border

The stroma in PDAC is composed of a complex mixture of cells and extracellular matrix (ECM) components, which makes delineation of precise stromal borders nontrivial. Fibronectin, a key ECM protein enriched in the stromal compartment, was used as a surrogate marker to define stromal regions. To reduce noise and imaging artifacts, fibronectin channel intensities were first smoothed using a Gaussian filter; the sigma parameter was tuned empirically based on expert input to balance edge preservation with noise reduction and then confirmed with a sensitivity analysis. A threshold-based pixel classifier was then applied to segment stromal versus non-stromal tissue. Pixels exceeding the fibronectin intensity threshold were classified as stromal. This threshold was selected by an experienced cancer biologist and later translated across the dataset using a percentile-based statistical propagation method. The border of the resulting binary mask was used as the stromal edge and served as the spatial reference for downstream analyses (Figure 1A).

2.3 Image analysis components

2.3.1 Nuclei detection and cell boundary approximation

Nuclei were segmented using the pre-trained StarDist model ("dsb2018_heavy_augment.pb", available at QuPath models repository: https://github.com/qupath/models/tree/main/stardist) via the StarDist extension in QuPath (v0.5). StarDist models nuclei as star-convex polygons and has demonstrated high accuracy across diverse nuclear morphologies (Schmidt et al., 2018). The output was a set of nuclear regions of interest (ROIs), each representing a single cell nucleus. To approximate whole-cell boundaries, QuPath's built-in cytoplasmic expansion algorithm was applied, radially expanding each nucleus by 5 μm. Cells were filtered based on nuclear area to exclude likely artifacts: nuclei smaller than the 5th percentile or larger than the 99th percentile were discarded, as they likely represented immune cells, debris, or fused nuclei caused by tissue processing artifacts (Figure 1B).

2.3.2 Distributions of cell intensity measurements

For each cell, fluorescence intensity features were extracted from the relevant subcellular compartments (nucleus, cytoplasm, whole cell) in QuPath. Signal intensity distributions were visualized and compared across images in the same batch using fixed bin widths. To reduce the influence of outliers, intensities were clipped at the 1st and 99th percentiles. No normalization was applied, preserving the raw distribution.

2.3.3 Cell classification

Cells were classified as marker-positive or -negative using either a supervised classifier or a percentile-based intensity thresholding approach (Figure 1B).

1. Supervised classification: A Random Forest classifier was trained within QuPath using a manually annotated subset of cells. The training set focused on intensity-based features relevant to the marker being evaluated (e.g., maximum nuclear intensity for Ki67 foci). Only intensity-related features were used to simplify model interpretation and reproducibility.
2. Percentile-based thresholding: A fixed intensity percentile was selected from the cell population within a reference image and applied across all images. This thresholding approach classified the top N% of cells by marker intensity as "positive", enabling harmonized comparisons across heterogeneous images.

Confusion matrices and derived standard metrics (agreement percentage) were computed to evaluate agreement between these two classification strategies. High agreement between the methods indicated classifier stability, while deviations flagged potential staining inconsistencies or classifier bias.

2.3.4 Percentile mapping of expert-chosen classification thresholds

To propagate expert-defined intensity thresholds across images with varying intensity distributions, we implemented a statistical mapping method based on percentile preservation. For each image, the distribution of intensity values for a given marker was modeled using the best-fit probability distribution (minimizing least-squares error). The expert-selected intensity threshold in a reference image was then mapped to the corresponding percentile, and this percentile was used to determine equivalent thresholds in other images. The algorithm is statistically described in Box 1. The resulting mapped thresholds were compared to machine learning-based classifications. Using one method's labels as "ground truth", confusion matrices were computed to evaluate accuracy and agreement. This percentile mapping approach enabled reproducible thresholding across large heterogeneous datasets while preserving expert intent.
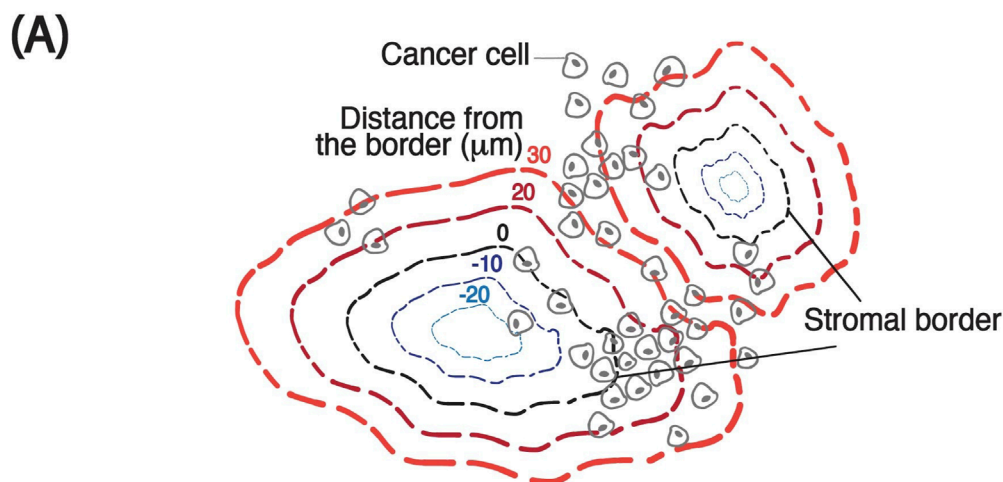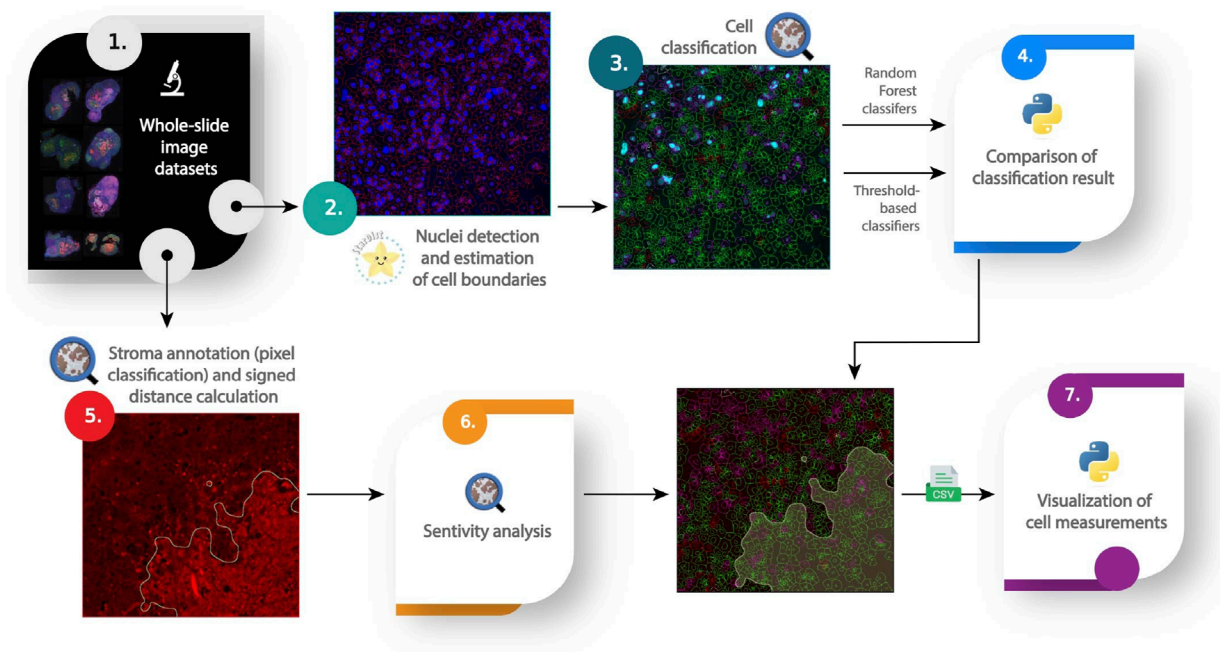
FIGURE 1
An image analysis pipeline for quantifying the spatial distribution of cellular markers in stroma-rich tumors [adapted from (Kozlova et al., 2025)]. **(A)** Graphical abstract of the spatial model for stroma-dense tissues. **(B)** Workflow for analyzing cellular and subcellular fluorescent markers relative to a modeled stromal border, including nuclei segmentation with StarDist, cell boundary estimation using QuPath, Random Forest models for cell classification, pixel threshold-based stroma annotation, 2D signed distance calculations for spatial analysis, and data visualization in Python.

2.3.5 2D signed distance between cells and the stromal border

We used the signed Euclidean distance between each cell's nuclear centroid and the nearest point on the stromal border to quantify spatial relationships between tumor cells and the stroma (Figure 1B). Positive distances indicate cells located outside, while negative distances represent cells within the fibronectin-defined stromal region. Cells located at the interface thus have distances close to 0 µm. Cells were binned by distance using 10 µm intervals, approximately corresponding to one cell diameter. These distance bins formed the basis for downstream spatial analyses and visualization of intensity over distance from the stroma.

2.3.6 Sensitivity analysis on intensity thresholds

To assess the robustness of the stromal mask and downstream spatial results, we conducted sensitivity analyses on two key parameters: the Gaussian smoothing sigma and the fibronectin intensity threshold. For each parameter combination, the stromal mask was recomputed, followed by recalculation of the signed distances and intensity-distance correlations.

BOX 1 Algorithm for statistically propagating classification thresholds across images based on the underlying distribution of cell measurements extracted from QuPath. It adjusts classification thresholds for each image by leveraging the statistical properties of cell marker intensity distributions. Given a reference threshold in a chosen image, the method translates this threshold to other images by aligning percentiles within the fitted distribution of cellular measurements. This approach accounts for variations in staining intensity and imaging conditions, ensuring consistency in classification across images.

Batch-wise statistical propagation of classification thresholds

Given a set of images $I_1, I_2, \ldots, I_N$ and for a specific channel c, we retrieve the histogram of each image's pixel intensity distributions using a fixed number of bins, N_b.

1. Histogram binning:

$$H_i(x) = \text{Histogram}(I_{i,c}, N_b)$$

Where $H_i(x)$ is the histogram of image I_i for channel c, and x the intensity levels.

2. Least-square fit of a pool of non-negative distributions: We fit a pool of non-negative distributions, D, on the intensity distributions:

$$D = \{f_1(x), f_2(x), \ldots, f_M(x)\}$$

Here, $f_i(x)$ represents the ith distribution function in the pool. The minimal pool of distributions contains common non-negative distributions: Log-normal, Wald, Burr, Beta and Gamma distributions. The best fit $f^*(x)$ is retrieved using the least-squares method:

$$f^*(x) = \arg\min_{f \in D} \sum_{i=1}^{N} \|H_i(x) - f(x)\|^2$$

3. Percentile mapping: Given a pixel intensity threshold t_i in distribution i, its percentile is mapped to another distribution j using their cumulative distribution functions (CDFs):

$$F_i(t_i) = F_j(t_j)$$

where $F_i(t)$ and $F_j(t)$ are the CDFs of distributions i and j, respectively.

4. Inverse probability calculation: The value of t_j is then calculated using the inverse cumulative distribution function of distribution j:

$$t_j = \left(F_j^{-1} \circ F_i\right)(t_i)$$

As a summary metric, we calculated the Pearson correlation coefficient between marker intensity I and distance to the stromal border d, separately for cells inside and outside the stroma. The correlation coefficient r was computed (Equation 1).

$$r = \frac{\sum_{i=1}^{n} \left(I_i - \bar{I}\right)\left(d_i - \bar{d}\right)}{\sqrt{\sum_{i=1}^{n} \left(I_i - \bar{I}\right)^2} \cdot \sqrt{\sum_{i=1}^{n} \left(d_i - \bar{d}\right)^2}} \tag{1}$$

where:

- I_i is the signal intensity for the ith cell
- d_i is the signed distance to the stromal border for the ith cell
- $\bar{I}$ is the mean signal intensity
- $\bar{d}$ is the mean signed distance

For each image and parameter setting, correlation coefficients were computed separately for regions before and after the zero-distance boundary. Because both correlation values were derived from the same image under identical conditions, the data were treated as paired measurements. Differences in correlation values across the two distance partitions (inside and outside the stroma) were assessed using a two-sided Wilcoxon signed-rank test, a non-parametric test for paired data. Mean correlation values and standard errors were estimated via 500 bootstrap iterations across images.

2.3.7 Visualization

Final cell-level data, including marker intensities, class labels, and signed distances, were exported from QuPath and analyzed in Python 3.11. Cells were grouped into 10 μm distance bins, and average marker intensity with bootstrapped standard error was computed for each bin. Profiles were generated for marker-positive and marker-negative populations separately. To assess differences in spatial patterns, bin-wise subtractions were performed between groups (e.g., Ki67-positive vs Ki67-negative cancer cells). The standard error of the difference in bin means was computed using error propagation (Equation 2).

$$SEM = \sqrt{\sum_{i=1}^{N} \frac{\sigma_i^2}{n_i}} \tag{2}$$

where:

- SEM is the propagated standard error
- σ_i is the standard deviation of signal intensities in a cell population, and n_i the number of cells in the population

Spatial distributions were visualized as line plots with error bands, highlighting marker-specific spatial gradients in relation to the stromal border.

2.3.8 Software implementation

All steps were performed on a workstation running Python 3.7 (Python Software Foundation, available at https://www.python.org) with standard scientific libraries (NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), Pandas (McKinney 2010; The Pandas Development Team, 2024), Fitter, Matplotlib (Hunter, 2007); a uv lock file is provided for dependencies). QuPath (v0.5+) was used for interactive segmentation, annotation, and feature export. Trained machine learning classifiers are available as .json files and can be loaded in QuPath to predict cell classes on new images or retrained on new annotations. The complete pipeline code is made open-source and is hosted on GitHub: https://github.com/HMS-IAC/stroma-spatial-analysis-web.

3 Results

In this work we present the pipeline 'in action' on multiplexed immunofluorescence images of PDAC xenografts treated with chemotherapy, based on a previous study of ours (Kozlova et al., 2025). This pipeline defines spatial distribution of cells expressing pan-cytokeratin (PDAC cells), phosphorylated NDRG1 (a stromal sensor), and Ki67 (a marker of cell proliferation) in relation to stromal border modeled by fibronectin staining (Figure 2). Following best practices for image publication (Schmied et al., 2024), Figure 2 is also provided in grayscale for visual comparison between channels (Supplementary Figures S1, and S2). Our analysis reveals the enrichment of the phospho-NDRG1- and Ki67-positive cells at the stromal border and their detailed spatial distribution based on signal intensity. The pipeline is scalable, robust to staining variability, and adaptable to other biomarker panels, offering a reproducible framework for spatial analysis of components of the tumor microenvironment in relation to stromal regions and their border.

3.1 Batch variability and statistical propagation of thresholds

To ensure the reliability of downstream analyses and aggregation across images, we first assessed the consistency of per-cell fluorescence intensity distributions within each image batch. While some variation is expected due to acquisition conditions and biological heterogeneity, we observed overall alignment in the profiles of the histograms and kernel density estimations (KDE) across both datasets (Supplementary Figure S3). KDEs revealed some variability between images in the Ki67 datasets, particularly in the pan-cytokeratin channel (Supplementary Figure S3B). In contrast, median nuclear intensity in the DAPI channel remained highly consistent across all five images, indicating uniform nuclear staining and robust segmentation quality in both datasets (Supplementary Figure S3). Next, we examined the best-fit probability distributions for cell measurements, used to map thresholds to a reference percentile across images (Figure 3). The pan-cytokeratin channel in the pNDRG1 dataset—used as a marker to distinguish cancer cells from non-cancer cells—was consistent across images (Figure 3A). In contrast, the pan-cytokeratin channel in the Ki67 dataset exhibited more substantial variability in median cytoplasmic signal intensity, especially in Image #2 and #5. This variability was reflected in the differing shapes of the probability

functions (PDF and CDF, Figure 3B). Although Images #2 and #5 deviated the most from the others, their values remained within acceptable bounds and did not compromise downstream analysis. Rather than excluding these images, we retained them to test the robustness of our threshold calibration strategy under conditions of variable staining. This variability offered an opportunity to assess how well our pipeline handles intensity heterogeneity without compromising classification consistency. Finally, the probability distribution of the maximum cellular pNDRG1 intensity and the maximum nuclear Ki67 intensity were both consistent across all images in their datasets (Figure 3).

We modeled the cell-level fluorescence histograms in each image using continuous, positive-valued probability distributions (Box 1). We evaluated multiple candidate distributions and found that the log-normal distribution consistently provided the best fit, as assessed by least-squares error. This distribution was especially well-suited to the skewed, elbow-shaped profiles commonly observed in fluorescence intensity histograms across both cytoplasmic and nuclear compartments. While fluorescence signals originate from discrete photons and are digitized into grey levels, the continuous log-normal distribution remains a good approximation at high photon counts—conditions under which the discreteness of the signal becomes negligible due to the Central Limit Theorem. Accordingly, we used log-normal fits for all channels in subsequent statistical threshold propagation.

Expert-defined thresholds were established in a single reference image for each marker (pan-cytokeratin, fibronectin, pNDRG1 or Ki67; excluding DAPI as a deep learning model was used for nuclei detection), and these values were mapped to the remaining images by preserving their percentile rank within the fitted log-normal distribution. This ensured that a consistent fraction of cells—corresponding to the top X% by intensity—was classified as marker-positive in each image, adapting the absolute threshold to the local intensity distribution. We applied this method to both phospho-NDRG1 and Ki67 image datasets. This percentile-based propagation approach offered advantages over using a fixed global threshold or a threshold specifically picked for each image. In practice, it improved the selection of marker-positive cells by statistically accounting for image-specific differences in staining intensity. It also provided a way to only have to manually select one threshold. We expect this method to be even more critical in studies with larger batch effects or multi-site imaging datasets, where inter-image variability is more pronounced.

3.2 Application case 1: revealing the spatial distribution between the phosphorylated form of NDRG1, a novel DNA repair protein required for stroma-induced chemoresistance, and stromal border

To demonstrate the utility of our pipeline in uncovering biologically meaningful spatial patterns, we analyzed the distribution of cells expressing high levels of phosphorylated NDRG1 (pNDRG1), a novel DNA repair protein implicated in chemoresistance and replication fork stability, in PDAC AsPC xenografts treated with gemcitabine, a standard-of-care chemotherapy used in the treatment of pancreatic cancer. Our previous work showed that phosphorylation of NDRG1 is regulated
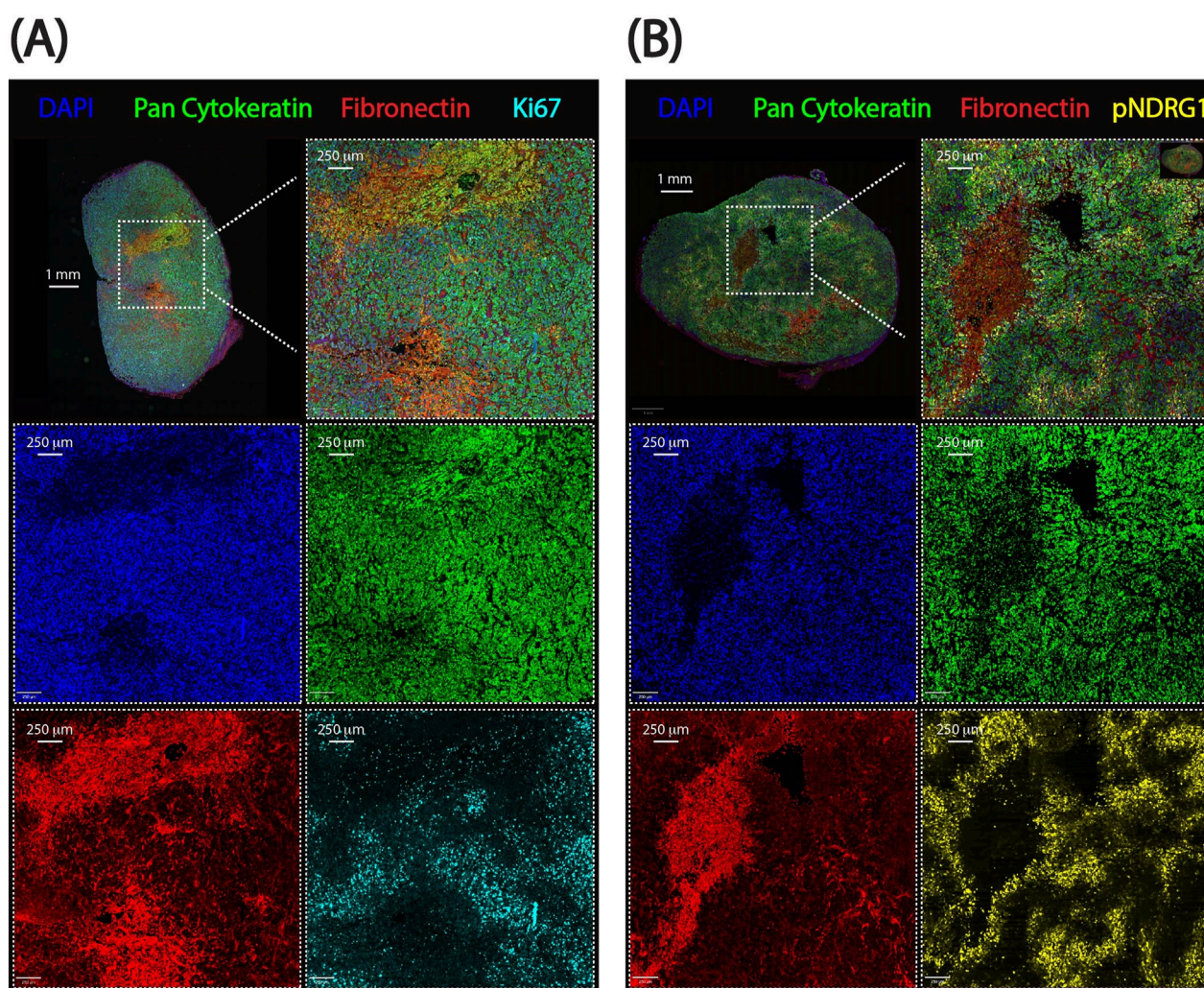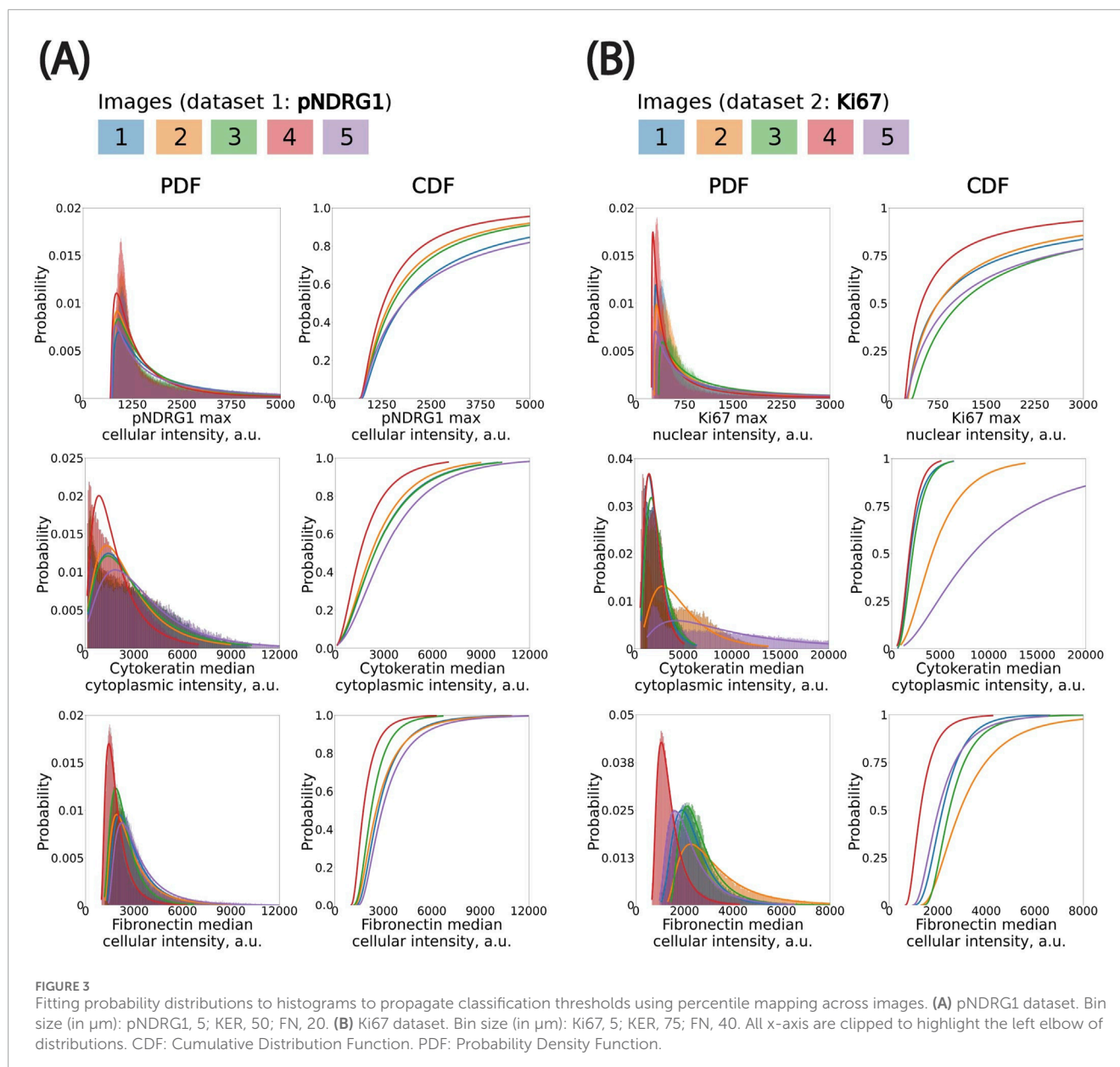
FIGURE 2
Representative whole-slide images from pancreatic ductal adenocarcinoma xenografts used for spatial analysis. **(A)** Ki67 dataset showing proliferating cancer cells. **(B)** pNDRG1 dataset highlighting phosphorylated NDRG1 signal intensity. Scalebars are displayed. DAPI: DAPI, Pan-cytokeratin: FITC, Fibronectin: TRITC, Ki67 or pNDRG1: CY5.

by ECM and adhesion receptors leading to the enrichment of pNDRG1-positive cells near tumor-stroma interfaces in PDAC SW1990 xenografts (Kozlova et al., 2025). Given that NDRG1 is localized both in the cytoplasm and the nucleus, we used maximum whole-cell intensity as the primary feature for classification. Additional features included median cytoplasmic pan-cytokeratin intensity (to select for cancer cells) and a median cellular fibronectin intensity (for stromal modeling). Fibronectin is a secreted matrix deposited outside of the cells; the median cellular intensity was only used as a proxy to visualize its distribution in a similar fashion (at the cell level) to the other markers (Figure 4A).

We trained a Random Forest classifier (QuPath, built-in) on a balanced set of 318 manually labeled cells spanning four classes: pan-cytokeratin-positive, pan-cytokeratin-negative, pNDRG1-positive, and pNDRG1-negative (Figure 4B). In parallel, we designed a threshold-based classifier using percentile-propagated cutoffs from a reference image (Image #1), as described in Box 1. This method was used to propagate classification thresholds from all channels except DAPI, that is pan-cytokeratin, fibronectin and pNDRG1. Agreement

between the machine learning and threshold-based methods was high (87.7%, Figure 4C, Supplementary Figure S4A), confirming the consistency of classification across heterogeneous images. Full classification parameters and translated thresholds are provided in the Supplementary Material (Supplementary Tables S1–S3). When comparing the percentile-propagation method to a fixed global threshold, we observed an average increase in classification agreement of 1.36% (±6.88% SD, Supplementary Table S4). These findings highlight that the necessity of accounting for intensity distribution variations is highly context dependent. In some cases, a single-threshold approach can approximate the discriminative power of a trained classifier, whereas in others, statistical propagation may be more appropriate. As a matter of illustration, when applying the set of thresholds of image #4 to the rest of the batch, the agreement drops to 78.7% (Supplementary Table S4). In other words, using statistical propagation averages out the variations observed when using a single set of thresholds for all images–providing a more robust way of picking a single set of thresholds for a whole batch.

FIGURE 3
Fitting probability distributions to histograms to propagate classification thresholds using percentile mapping across images. **(A)** pNDRG1 dataset. Bin size (in μm): pNDRG1, 5; KER, 50; FN, 20. **(B)** Ki67 dataset. Bin size (in μm): Ki67, 5; KER, 75; FN, 40. All x-axis are clipped to highlight the left elbow of distributions. CDF: Cumulative Distribution Function. PDF: Probability Density Function.

Spatial analysis applied to AsPC xenograft images revealed that pNDRG1-positive cells were enriched near stromal borders. Visual inspection of whole-slide images showed areas of high pNDRG1 intensity in tumor regions adjacent to stroma. Both classification methods consistently quantified this pattern. Two key trends emerged (Figure 4D, Supplementary Figure S4B).

1. pNDRG1 intensity peaked at the stromal borders of gemcitabine-treated tumors.
2. The number of pNDRG1-positive cells within the stomal regions is much lower compared to the one outside the stromal compartment, as can be assessed by the number of cells in each bin.

These findings are consistent with a general observation, that pancreatic stroma presents as a desmoplastic matrix limiting cell

infiltration. The cells adjacent to the stromal border respond to mechanical sensing of the ECM which is reflected in the phosphorylation of the NDRG1. These findings also emphasize the value of incorporating intensity calibration into cross-image analyses.

3.3 Application case 2: the case of Ki67 revealing the spatial distribution between cell proliferation and stromal border

We next applied the pipeline to Ki67, a canonical nuclear marker of cellular proliferation, to investigate whether spatial gradients of proliferative activity exist in relation to the stromal border. Since Ki67 localization is strictly nuclear, we used maximum nuclear intensity as the key classification feature (Figure 5A).
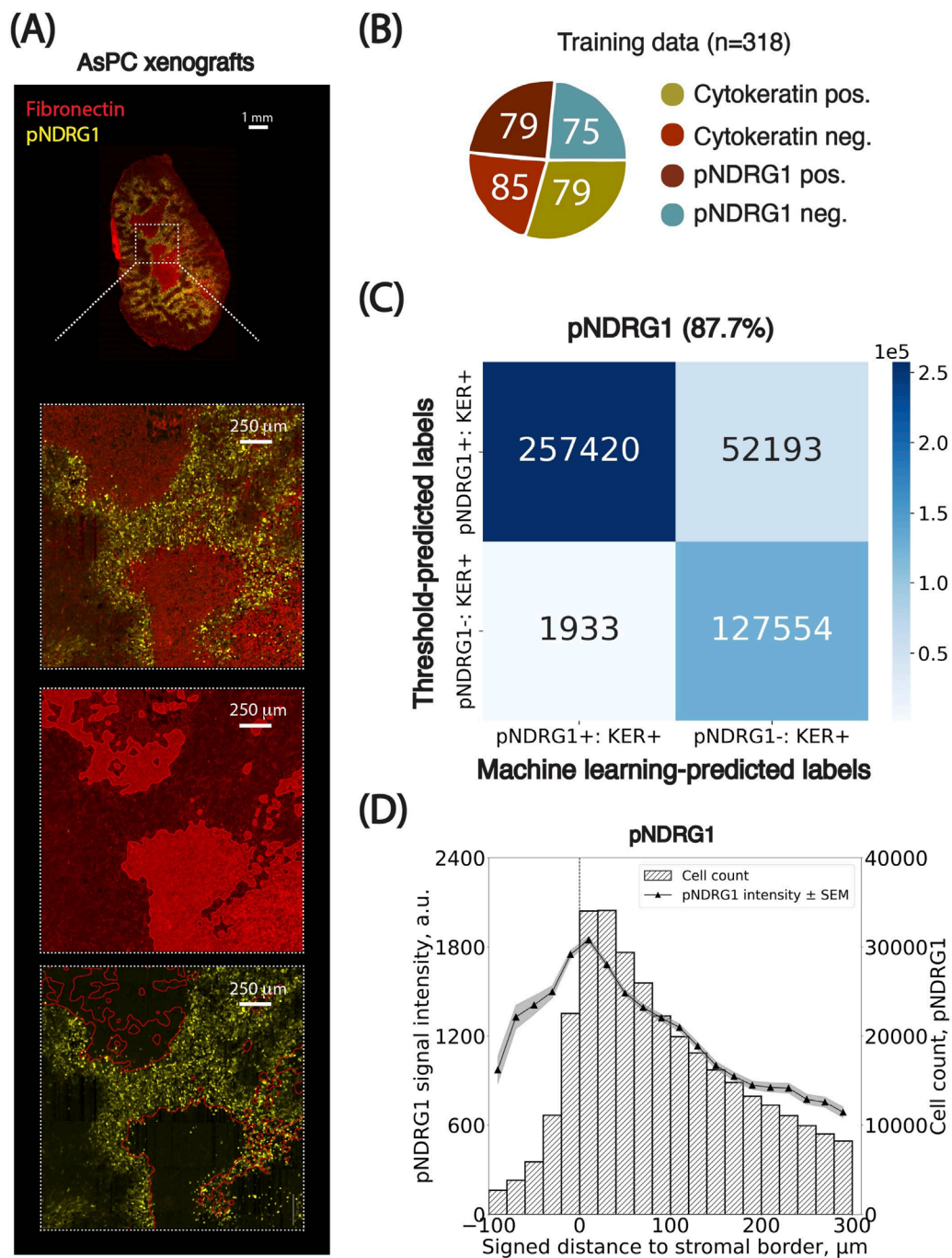
FIGURE 4
Spatial distributions of pNDRG1-positive cancer cells relative to the stromal border. **(A)** Representative images of the PDAC AsPC xenograft sections stained for Fibronectin (red) and phospho-NDRG1 (yellow) used for spatial analysis. **(B)** Breakdown of the training set curated for cells in pNDRG1 images. Numbers inserted in charts indicate the number of training examples per class. **(C)** Confusion matrix comparing threshold-based and machine learning-based classification of cells based on cellular marker signal intensity. Percentages indicate the agreement between classification results from the two methods. KER: Pan-cytokeratin. **(D)** Spatial analysis of pNDRG1 signal intensity in cancer cells as a function of their signed distance to the closest stromal border (x-axis, in µm). Cells positive for pNDRG1 were grouped into bins based on their distance from the stromal border. The left y-axis shows the average maximum cellular signal intensity for pNDRG1-positive cells (black dots) per bin, with standard error of the mean (SEM) indicated as grey overlays. The right y-axis represents the number of cells per bin. Negative distances correspond to cells within stromal regions, while positive distances indicate cells outside these regions. Bin size: 10 µm. The x-axis is clipped between −100 and 300 µm around the stromal border. Results from the cell classification using adaptive thresholds were used.
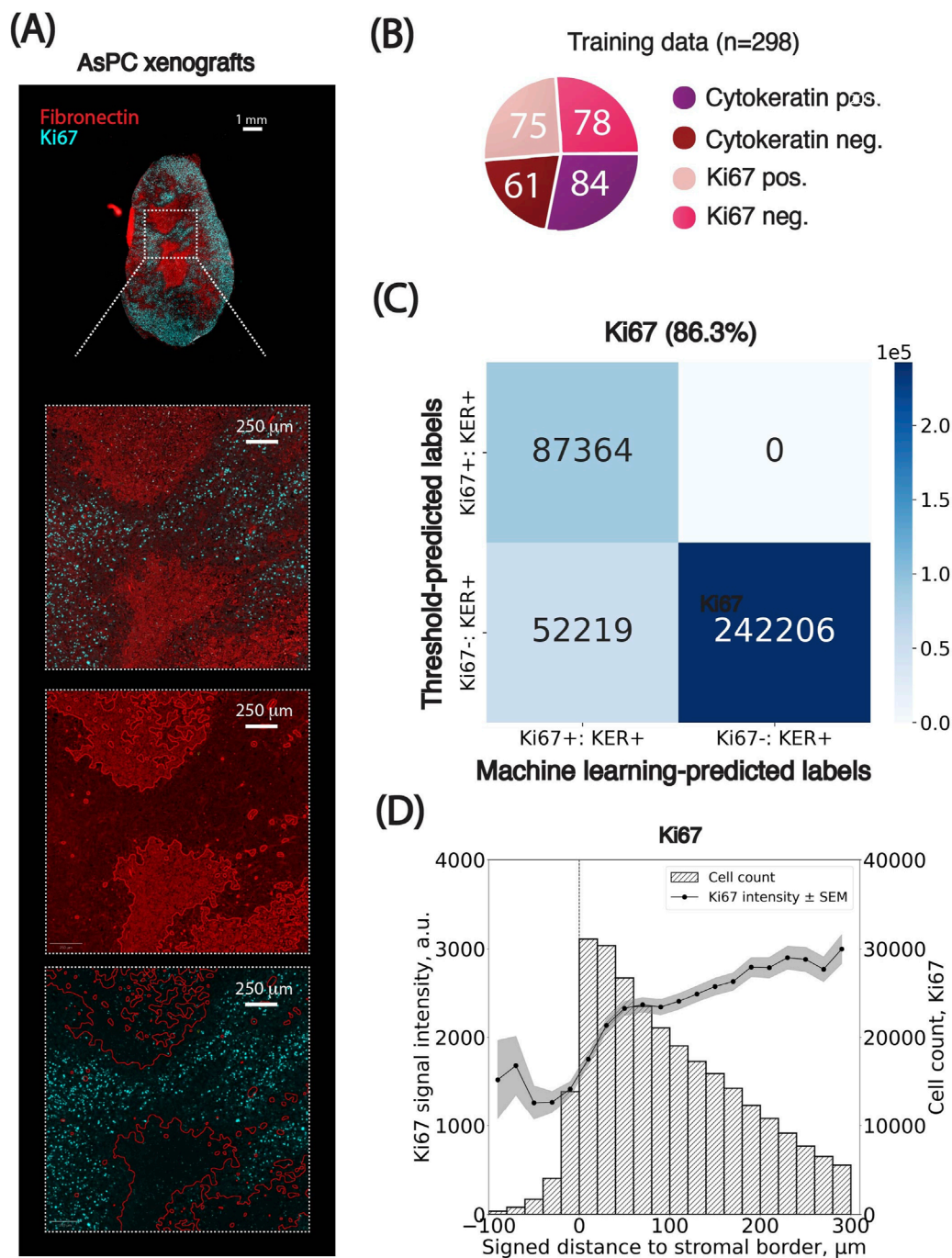
FIGURE 5
Spatial distributions of Ki67-positive cancer cells relative to the stromal border. **(A)** Representative images of the PDAC AsPC xenograft sections stained for Fibronectin (red) and Ki67 (cyan) used for spatial analysis. **(B)** Breakdown of the training set curated for cells in Ki67 images. Numbers inserted in charts indicate the number of training examples per class. **(C)** Confusion matrix comparing threshold-based and machine learning-based classification of cells based on cellular marker signal intensity. Percentages indicate the agreement between classification results from the two methods. KER: Pan-cytokeratin. **(D)** Spatial analysis of Ki67 signal intensity in cancer cells as a function of their signed distance to the closest stromal border (x-axis, in μm). Cells positive for Ki67 were grouped into bins based on their distance from the stromal border. The left y-axis shows the average maximum cellular signal intensity for Ki67-positive cells (black dots) per bin, with standard error of the mean (SEM) indicated as grey overlays. The right y-axis represents the number of cells per bin. Negative distances correspond to cells within stromal regions, while positive distances indicate cells outside these regions. Bin size: 10 μm. The x-axis is clipped between −100 and 300 μm around the stromal border. Results from the cell classification using adaptive thresholds were used.

A Random Forest classifier was trained on 298 manually annotated cells across four classes (pan-cytokeratin-positive, pan-cytokeratin-negative, Ki67-positive, and Ki67-negative; Figure 5B). We also defined percentile-based thresholds from Image #1 and propagated them across all images using our statistical method (Box 1) for pan-cytokeratin, fibronectin and Ki67 channels. Both classification approaches produced similar results, with an average agreement of 86.3% (Figure 5C, Supplementary Figure S5A). Full classification parameters and translated thresholds are provided in the Supplementary Material (Supplementary Tables S5–S7). However, unlike pNDRG1, Ki67 classification proved more sensitive to intensity heterogeneity: using a single global threshold led to an average agreement drop of 2.9% (±10.1% SD, Supplementary Table S8), highlighting the necessity of threshold calibration when analyzing markers with higher inter-image variability.

Spatial mapping of cells both positive for Ki67 and Pan-cytokeratin revealed a clear trend: the number of actively proliferating cancer cells was lowest within stromal regions and progressively increased with distance from the stroma-tumor interface, plateauing at approximately 300 μm (Figure 5D, Supplementary Figure S5B). In the following sections, cells referred to as Ki67+ are both Ki67- and Pan-cytokeratin-positive (selected for cancer cells). Two observations support this.

1. Ki67 signal intensity in increased monotonically from 0 to ~300 μm away from the stromal border, after which it stabilized.
2. Similarly, the number of both Ki67-positive cells within the stomal regions was much lower compared to the one outside the stromal compartment, assessed by the number of cells in each bin.

These findings support the notion that dense stroma creates a mechanical and a biochemical barrier that manifests in diminished exposure of cancer cells to chemotherapy, thus having a major impact on cancer cell proliferation. They also demonstrate the importance of adapting classification strategies to accommodate image-specific intensity profiles.

3.4 Sensitivity analysis of thresholds on the spatial distribution

We conducted a sensitivity analysis to assess the robustness of spatial trends with respect to two key parameters involved in stromal mask generation: (1) the standard deviation (σ) of the Gaussian kernel applied to the fibronectin channel and (2) the pixel intensity threshold used to distinguish stromal from non-stromal regions. We performed a grid search over both parameters and, for each combination, recalculated the stromal mask, cell-to-stroma distances, and the Pearson correlation between marker intensity and stromal proximity. Across this parameter space, spatial trends remained consistent for both markers (Figure 6). For pNDRG1, the correlation between intensity and distance ranged from −0.01 to 0.03 for cells inside the stroma and from −0.16 to −0.15 for cells outside the stroma. For Ki67, the correlation ranged from 0.04 to 0.07 inside the stroma and from 0.01 to 0.03 outside the stroma. These small fluctuations indicate that the observed spatial patterns

are robust to reasonable variations in smoothing and thresholding. Taken together, the results confirm that the observed spatial patterns are not artifacts of specific parameter choices. This step also led us to empirically select σ = 15 for pNDRG1 and σ = 10 for Ki67, based on the parameter set that maximized smoothness of the stromal border while preserving local spatial structure. These values produced segmentation masks that were both biologically plausible and numerically stable across images.
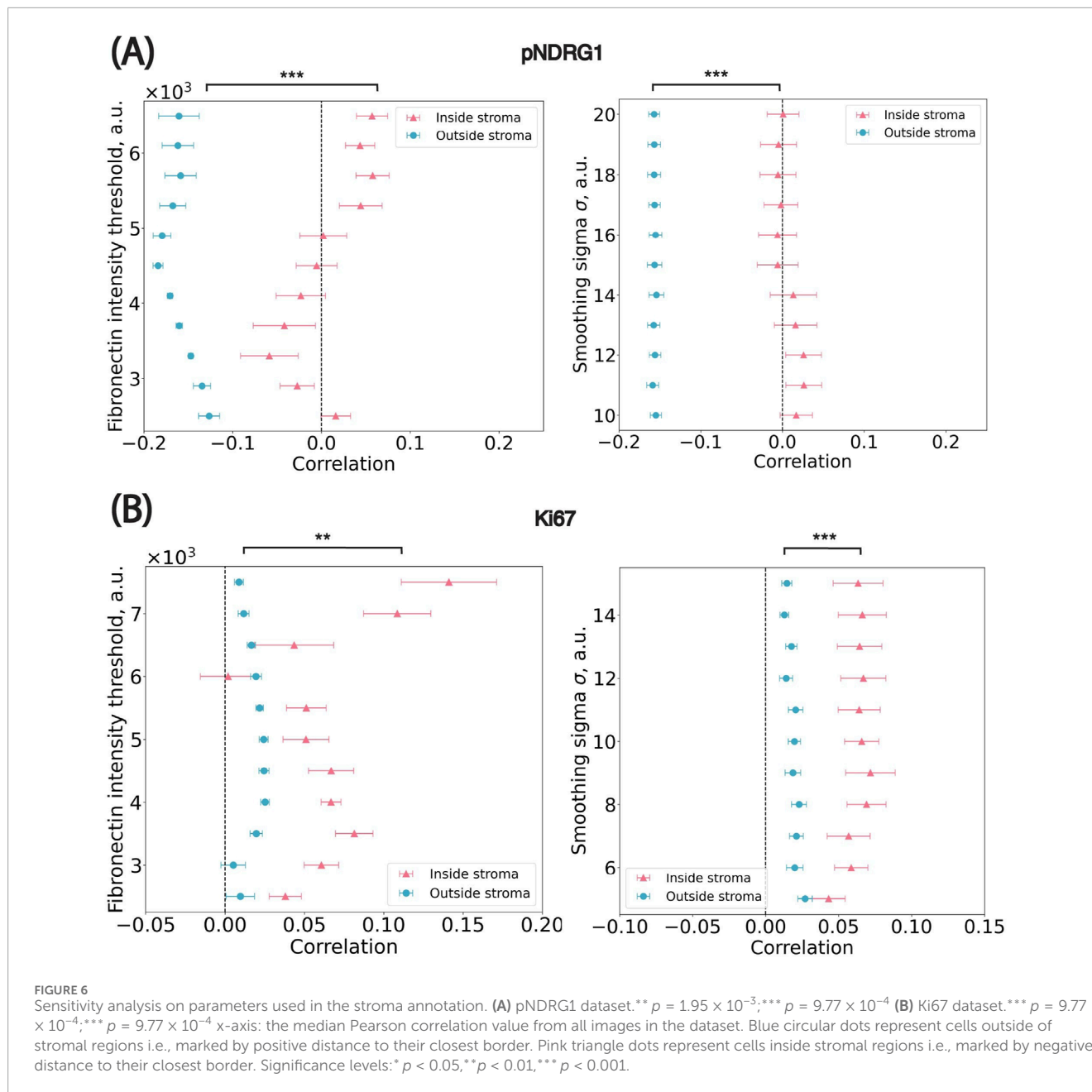
4 Discussion

In this study we present an open-source image analysis pipeline for the robust and reproducible quantification of spatial biomarker distributions in stroma-rich tumors. By integrating QuPath (Bankhead et al., 2017), StarDist (Schmidt et al., 2018), and Python-based statistical analyses, our framework enables the extraction of biologically relevant spatial patterns from heterogeneous immunofluorescence imaging datasets. Notably, we found that the fibronectin-defined stromal border consistently aligned with spatial changes in biomarker expression. Specifically, in cancer cells positive for the phospho-NDRG1, the signal intensity peaked at the tumor-stroma interface and started to rapidly decline in the cells located farther from the stromal border (Figure 4). As for the cells positive for Ki67, the intensity was low inside the stroma and increased progressively with greater distance from the border until plateauing at approximately 300 μm mark (Figure 5).

These patterns reveal a spatially constrained organization of ECM sensing and proliferative response programs shaped by the surrounding stroma. While dense stromal regions influence the activation of key intracellular signaling pathways and restrict drug availability, other components of the TME orchestrate a wide range of immune, metabolic, and transcriptional changes that contribute to the complexity of stroma-rich cancers. Further biological and biochemical studies are needed to uncover the underlying cues driving the distinct spatial distribution of these programs.

4.1 Scalability and generalizability

Our pipeline is designed to process gigapixel-scale whole-slide images with batch-level reproducibility. While classification via QuPath's machine learning tools still requires manual annotation, we show that statistical propagation of thresholds from a reference image can approximate the accuracy of trained classifiers. This approach significantly improves scalability for large datasets by reducing the need for manual curation, while also compensating for staining variability across slides. One key take home message is that the framework is not limited to pancreatic cancer mouse xenografts, but can be applied for the analysis of various tissues with high stromal content (e.g., breast, ovarian cancer tissues, fibrotic tissues, bone marrow, or adipose tissue) (Friščić and Hoffmann, 2022). Stromal regions can be delineated using any appropriate marker (e.g., collagens, laminins), making the pipeline adaptable to a wide range of tumor types and other dense tissues. Similarly, classification can be tailored to different biomarkers and experimental goals using either supervised or threshold-based methods, depending on the availability of labeled training data. This flexibility makes

FIGURE 6
Sensitivity analysis on parameters used in the stroma annotation. **(A)** pNDRG1 dataset.** $p = 1.95 \times 10^{-3}$;*** $p = 9.77 \times 10^{-4}$ **(B)** Ki67 dataset.*** $p = 9.77$ $\times 10^{-4}$;*** $p = 9.77 \times 10^{-4}$ x-axis: the median Pearson correlation value from all images in the dataset. Blue circular dots represent cells outside of stromal regions i.e., marked by positive distance to their closest border. Pink triangle dots represent cells inside stromal regions i.e., marked by negative distance to their closest border. Significance levels:* $p < 0.05$,$^{**}p < 0.01$,$^{***}p < 0.001$.

the pipeline applicable across diverse contexts even beyond cancer biology, where spatial phenotyping is relevant.

4.2 Limitations and future work

Despite its strengths, the pipeline has several limitations. First, stromal region delineation is currently based on intensity thresholding, which, although validated via sensitivity analysis, remains susceptible to staining variability and human bias. Incorporating texture features or adopting deep learning-based segmentation methods may further improve robustness by being even more agnostic for human-chosen parameters. Second, while the statistical propagation method mitigates batch effects, it still

depends on the initial selection of a reference threshold. We show that this dependency is dataset-specific, and recommend examining cell-level intensity distributions before choosing between machine learning or percentile-based classification strategies. Then, we acknowledge that a fully manual expert annotation comparison (e.g., from a pathologist or cancer biologist) would be an interesting future benchmark. While outside the scope of the present study, it would be useful to further assess the pipeline's biological and clinical validity. Additionally, this work focuses exclusively on 2D spatial measurements. Although informative, 2D projections may obscure spatial interactions that unfold in three dimensions. Future extensions of the pipeline could integrate volumetric imaging techniques and 3D segmentation to enable more comprehensive modeling of tumor-stroma interactions in space.

4.3 Conclusion

This study demonstrates a reproducible and scalable pipeline for analyzing the spatial organization of biomarkers in stroma-rich tissues using images acquired from the chemotherapy-treated pancreatic cancer xenografts as an example. By leveraging user-friendly open-source software and emphasizing interpretable statistical modeling, the approach lowers the barrier to quantitative spatial analysis in experimental pathology. As large-scale multiplexed imaging becomes increasingly common, tools like this will be essential for integrating spatial context into the study of tumor biology and therapy response.

Summary paragraph

Our study introduces a scalable and reproducible image analysis pipeline that quantifies spatial biomarker distributions relative to the stroma in tumor tissues using open-source tools. By modeling cell-level intensity distributions and calibrating classification thresholds across heterogeneous images, we uncover spatially organized patterns of stroma sensing, DNA damage, and proliferative response in pancreatic tumors. This approach enables robust, quantitative analysis of tumor-stroma interactions and is readily adaptable to other tumor types and biomarker panels, providing a valuable resource for spatial pathology and tumor microenvironment research.

Data availability statement

The datasets (Dataset 1 and Dataset 2) presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://doi.org/10.5281/zenodo.15297453.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used. The animal study was approved by Institutional Animal Care and Use Committee at BIDMC. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

AR: Software, Writing – original draft, Visualization, Writing – review and editing, Investigation, Formal Analysis, Methodology, Conceptualization. NK: Data curation, Methodology, Writing – review and editing, Conceptualization, Investigation, Writing – original draft. KC: Writing – review and editing, Data curation, Writing – original draft, Conceptualization. TM: Funding acquisition, Conceptualization, Writing – original draft, Writing – review and editing, Supervision. SN: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – original draft.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbinf.2025.1619790/full#supplementary-material

References

Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., et al. (2017). QuPath: open source software for digital pathology image analysis. *Sci. Rep.* 7 (1), 16878. doi:10.1038/s41598-017-17204-5

Cimini, B. A., Bankhead, P., D'Antuono, R., Fazeli, E., Fernandez-Rodriguez, J., Fuster-Barceló, C., et al. (2024). The crucial role of bioimage analysts in scientific research and publication. *J. Cell. Sci.* 137 (20), jcs262322. doi:10.1242/jcs.262322

Friščić, J., and Hoffmann, M. H. (2022). Stromal cell regulation of inflammatory responses. *Curr. Opin. Immunol.* 74 (February), 92–99. doi:10.1016/j.coi.2021.10.006

Halbrook, C. J., Lyssiotis, C. A., Pasca di Magliano, M., and Maitra, A. (2023). Pancreatic cancer: advances and challenges. *Cell.* 186 (8), 1729–1754. doi:10.1016/j.cell.2023.02.014

Harris, C. R., Jarrod Millman, K., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature* 585 (7825), 357–362. doi:10.1038/s41586-020-2649-2

Ho, W. J., Jaffee, E. M., and Zheng, L. (2020). The tumour microenvironment in pancreatic cancer - clinical challenges and opportunities. *Nat. Rev. Clin. Oncol.* 17 (9), 527–540. doi:10.1038/s41571-020-0363-5

Hu, C., Li, T., Xu, Y., Zhang, X., Li, F., Bai, J., et al. (2023). CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on ScRNA-seq Data. *Nucleic Acids Res.* 51 (D1), D870–D876. doi:10.1093/nar/gkac947

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. and Eng.* 9 (3), 90–95. doi:10.1109/mcse.2007.55

Karantza, V. (2011). Keratins in health and cancer: more than mere epithelial cell markers. *Oncogene* 30 (2), 127–138. doi:10.1038/onc.2010.456

Kozlova, N., Cruz, K. A., Doh, H. M., Ruzette, A. A., Willis, N. A., Hong, Su M., et al. (2025). A novel DNA repair protein, N-myc downstream regulated gene 1 (NDRG1), links stromal tumour microenvironment to chemoresistance. *bioRxiv*. New York, NY: Cold Spring Harbor. doi:10.1101/2025.01.22.634323

Kozlova, N., Grossman, J. E., Iwanicki, M. P., and Muranen, T. (2020). The interplay of the extracellular matrix and stromal cells as a drug target in stroma-rich cancers. *Trends Pharmacol. Sci.* 41 (3), 183–198. doi:10.1016/j.tips.2020.01.001

Foot, N. C. (1933). The Masson Trichrome Staining Methods in Routine Laboratory Use. *Stain Technology* 83, 101–110. doi:10.3109/10520293309116112

McKinney, W. (2010). Data structures for statistical computing in Python. In: Proceedings of the Python in science conference. Austin, TX: SciPy. p. 56–61.

Miura, K., and Nørrelykke, S. F. (2021). Reproducible image handling and analysis. *EMBO J.* 40 (3), e105889. doi:10.15252/embj.2020105889

Napari Contributors (2019). *napari: a multi-dimensional image viewer for python*. Geneva, Switzerland: Zenodo. doi:10.5281/zenodo.3555620

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., et al. (2017). ImageJ2: ImageJ for the next generation of scientific image Data. *BMC Bioinforma.* 18 (1), 529. doi:10.1186/s12859-017-1934-z

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9 (7), 676–682. doi:10.1038/nmeth.2019

Schindelin, J., Rueden, C. T., Hiner, M. C., and Eliceiri, K. W. (2015). The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol. Reproduction Dev.* 82 (7–8), 518–529. doi:10.1002/mrd.22489

Schmidt, U., Weigert, M., Coleman, B., and Myers, G. (2018). Cell detection with star-convex polygons. In: Medical image computing and computer assisted intervention – MICCAI 2018. Cham: Springer International Publishing. p. 265–273.

Schmied, C., Nelson, M. S., Avilov, S., Bakker, G.-J., Bertocchi, C., Bischof, J., et al. (2024). Community-developed checklists for publishing images and image analyses. *Nat. Methods* 21 (2), 170–181. doi:10.1038/s41592-023-01987-9

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH image to ImageJ: 25 Years of image analysis. *Nat. Methods* 9 (7), 671–675. doi:10.1038/nmeth.2089

Stirling, D. R., Carpenter, A. E., and Cimini, B. A. (2021a). CellProfiler analyst 3.0: accessible Data exploration and machine learning for image analysis. *Bioinforma. Oxf. Engl.* 37 (21), 3992–3994. doi:10.1093/bioinformatics/btab634

Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., Carpenter, A. E., Cimini, B. A., and Goodman, A. (2021b). CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinforma.* 22 (1), 433. doi:10.1186/s12859-021-04344-9

The Pandas Development Team (2024). *Pandas-Dev/Pandas: Pandas*. Geneva, Switzerland: Zenodo. doi:10.5281/ZENODO.3509134

Valkenburg, K. C., de Groot, A. E., and Pienta, K. J. (2018). Targeting the tumour stroma to improve cancer therapy. *Nat. Rev. Clin. Oncol.* 15 (6), 366–381. doi:10.1038/s41571-018-0007-1

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* 17 (3), 261–272. doi:10.1038/s41592-019-0686-2

Xu, M., Zhang, T., Xia, R., Wei, Y., and Wei, X. (2022). Targeting the tumor stroma for cancer therapy. *Mol. Cancer* 21 (1), 208. doi:10.1186/s12943-022-01670-1