Briefings in Bioinformatics, 2026, 27, bbag201
https://doi.org/10.1093/bib/bbag201
Published: 30 April 2026
Problem Solving Protocol

OXFORD

stGrads: decoding spatial gene expression gradients through proximity-driven analysis in complex tissues

Yifan Fu[1,2,3,‡], Fan Zhang[1,‡], Feifan Zhang[4], Taiping Zhang[2,*], Wei Chen[1,*]

[1]Key Laboratory of Biomechanics and Mechanobiology (Beihang University), Ministry of Education; Key Laboratory of Innovation and Transformation of Advanced Medical Devices, Ministry of Industry and Information Technology; National Medical Innovation Platform for Industry-Education Integration in Advanced Medical Devices (Interdiscipline of Medicine and Engineering); School of Engineering Medicine, Beihang University, No. 37 Xueyuan Road, Haidian District, 100191 Beijing, China
[2]General Surgery Department, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 1 Shuaifuyuan Road, Dongcheng District, 100730 Beijing, China
[3]4+4 Medical Doctor Program, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 5 Dongdansantiao, Dongcheng District, 100730 Beijing, China
[4]Lee Kong Chian School of Medicine, Nanyang Technological University, No. 50 Nanyang Avenue, 639798, Singapore

*Corresponding authors. Taiping Zhang, General Surgery Department, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No.1 Shuaifuyuan Road, Dongcheng District, 100730 Beijing, China. E-mail: tpingzhang@yahoo.com; Wei Chen, Beihang University, No. 37 Xueyuan Road, Haidian District, 100191 Beijing, China. E-mail: chenw123@buaa.edu.cn.
‡Yifan Fu and Fan Zhang contributed equally to this work (co-first authors).

Abstract

The advent of spatial transcriptomics has opened new avenues for exploring spatial heterogeneity in gene expression across tissues. However, effectively quantifying the influence of specific cell populations on their niche remains a key challenge. Here, we present stGrads (Spatial Transcriptomic Gradients), a computational framework designed for both spot-size and bin-size spatial data to characterize the spatial gradients induced by specific cell types at both gene expression and cell composition levels. By integrating spatial proximity modeling with attenuation-based signal propagation, stGrads computes multiple distance-dependent metrics, including expression gradients, compositional shifts, and interaction strengths, to reveal local patterns of cellular responses. Meanwhile, stGrads could identify the gradient-related genes for downstream work. We demonstrate the utility of stGrads on multiple spatial transcriptomic datasets, successfully identifying disease-associated spatial gradients. stGrads offers a generalizable and efficient tool for characterizing spatial interactions and functional responses in complex tissue environments.

Keywords spatial transcriptomics, spatial gradient modeling, tissue microenvironment, stGrads, gradient-related gene identification

Introduction

Spatial transcriptomics (ST) integrates spatial coordinate information with high-throughput sequencing technology to achieve precise localization and visualization of gene expression patterns within tissue contexts [1, 2]. This integration has greatly advanced our understanding of tissue microenvironments, physiological homeostasis, and pathological processes [3, 4]. With the ongoing development of platforms such as Visium, Xenium, and Stereo-seq [2, 5], the spatial resolution and throughput of ST data have significantly improved, enabling a transition from method development to widespread application in both basic and clinical research [6].

Tissue function fundamentally depends on the spatial organization of structural components and the regional specificity of cell populations [7]. Under various pathological conditions, such as cancer, inflammation, and fibrosis, the original tissue architecture is often destroyed or remodeled, resulting in abnormal cell aggregation and enhanced spatial heterogeneity [8]. In these tissues, areas rich in immune cells, necrotic areas, or hypoxic areas are often regarded as biologically meaningful spatial reference areas [9]. Due to their unique cell density, functional status, or pathological characteristics, these areas may have a perturbing effect on the surrounding spatial gene expression profile [10]. Identifying and modeling such spatial perturbations is crucial to revealing the mechanisms of local microen-

Received: December 17, 2025. **Revised:** March 9, 2026. **Accepted:** April 2, 2026

vironment regulation, tissue repair response, and intercellular communication [11].

Although existing tools such as SpatialDE and Giotto have been widely used to detect spatially variant genes, clustering, and visualization, they have significant limitations in handling spatial continuity and quantifying the effects of region-specific perturbations. While SpatialDE uses a Gaussian process framework to model spatial covariance, it is primarily used to identify spatial variation rather than analyze directional effects of user-defined reference regions [12]. Giotto provides spatial coherence detection through neighborhood-based enrichment analysis, but it relies heavily on expression binarization and cannot model how expression gradients evolve with increasing distance from a specific spatial domain [13]. These limitations are particularly evident when analyzing structurally disordered or highly heterogeneous tissues such as solid tumors, as traditional binary classification methods (tumor versus non-tumor) cannot capture the subtle biological gradients that underlie spatial regulation.

To address these gaps, we introduce stGrads, a simple, general, and automated framework for spatial gradient modeling. stGrads allows users to flexibly define spatial reference regions based on prior knowledge or automated annotations and to systematically construct gradient models describing how gene expression or cell-type proportions vary with distance from these regions. By integrating spatial distance decay models, gradient intensity visualization, and expression correlation analysis, stGrads provides a reusable and interpretable pipeline tailored for spatial transcriptomics. It is particularly suitable for studying spatial regulatory patterns and functional responses in complex tissue microenvironments.

Results

Overview of the stGrads framework

To investigate the spatial influence of defined tissue regions on their niche, we developed stGrads, a general framework for quantifying expression gradients based on spatial distance (Fig. 1 and Methods).

Specifically, stGrads follows four key steps and is applicable to different types of spatial transcriptomics data. For spot-based ST data such as 10× Visium, cell-type proportion matrices are used to visualize the distribution of specific cell types, and cell-enriched regions are selected as reference spots for downstream modeling. For bin-based high-resolution ST data (Stereo-seq and 10× Visium HD), biologically meaningful bins are identified through prior annotations or unsupervised clustering results. In either case, user-defined spatial reference spots (immune infiltrate regions, necrotic regions, or specific annotated cell clusters) are first selected, and the remaining locations are designated as query spots.

After defining reference and query spots, stGrads computes the nearest distance matrix by calculating the Euclidean distance from each query spot to the nearest reference spot in spatial coordinates. stGrads supports linear, logarithmic, and exponential decay models to represent different forms of spatial attenuation. Exponential decay reflects diffusion-like short-range effects observed in morphogen gradients such as Decapentaplegic and Wingless [14, 15]. Logarithmic decay captures more gradual long-range influence [16]. Linear decay provides a simple monotonic approximation and is recommended as a default starting point. Importantly, stGrads also allows fully user-defined decay functions, which can be passed through the self_decay argument, enabling customized distance-to-strength transformations for case-specific biological contexts.

In the final step, a variety of downstream analyses can be conducted to interrogate spatial gradients. These include visualizing spatial maps of distance and interaction strength, assessing correlations between gene expression and spatial proximity, and identifying distance-responsive genes (DRGs) that exhibit significant spatial variation. Through these steps, stGrads provides a flexible and interpretable solution that is compatible with both low-resolution and high-resolution spatial transcriptomics datasets, making it broadly applicable to a variety of tissue types and spatial scales.

Core principles of the stGrads framework visualized by spot-size data

The spot-size data (10× Visium) was first applied to visualize the core principles and general framework. One of the vital differences between low- and high-resolution is whether the single data point could reach single cell (or sub-cellular) level, which means for low-resolution data, we should first identify the composition of spots. While the high-resolution data would be input with bin/spot-based cellular type labels.

To quantify spatial regulatory effects originating from defined tissue regions, stGrads models gene expression and cell composition as continuous functions of spatial distance. The core computational principles of stGrads are shown in Fig. 2, comprising three conceptual layers: distance computation, interaction strength estimation, and expression-gradient analysis.

First, stGrads calculates the Euclidean distance between each query spot (a tissue spot or bin) and all reference spots. For each query spot, the shortest distance to any reference spot is retained as the nearest distance (Dis_n). This establishes a spatial proximity field radiating from the region of interest.

Second, to model the influence that reference spots may exert on their niche, stGrads estimates an interaction strength score for each query spot. This is achieved by summing contributions from nearby reference spots, weighted by a user-defined distance-decay model. stGrads supports three decay functions (linear, logarithmic, and exponential) to accommodate different biological assumptions about spatial signal propagation. Importantly, distant reference spots beyond a defined radius (D_{max}) are excluded from the calculation, allowing for local interaction modeling.

Third, the resulting distance and strength profiles are used to explore the relationship between spatial proximity and biological response. stGrads provides tools to correlate expression levels of individual genes with either distance or interaction strength, enabling the identification of distance-responsive genes. Expression trends can be visualized and statistically evaluated using Spearman correlation and Loess regression, revealing genes whose activity increases or decreases in relation to spatial location.

Through this distance-strength-expression modeling, stGrads provides a flexible and quantitative framework for capturing gradient-like regulation in tissue sections without requiring predefined structural boundaries. At the same time, it allows users to specify reference spots or cell populations, enabling targeted detection of genes influenced by particular cell types while maintaining flexibility in spatial modeling.

We compared differentially regulated genes in pancreatic cancer tissue identified by stGrads with those obtained using Giotto, SpatialDE, and HEARTSVG [17]. In addition to gene set compar-

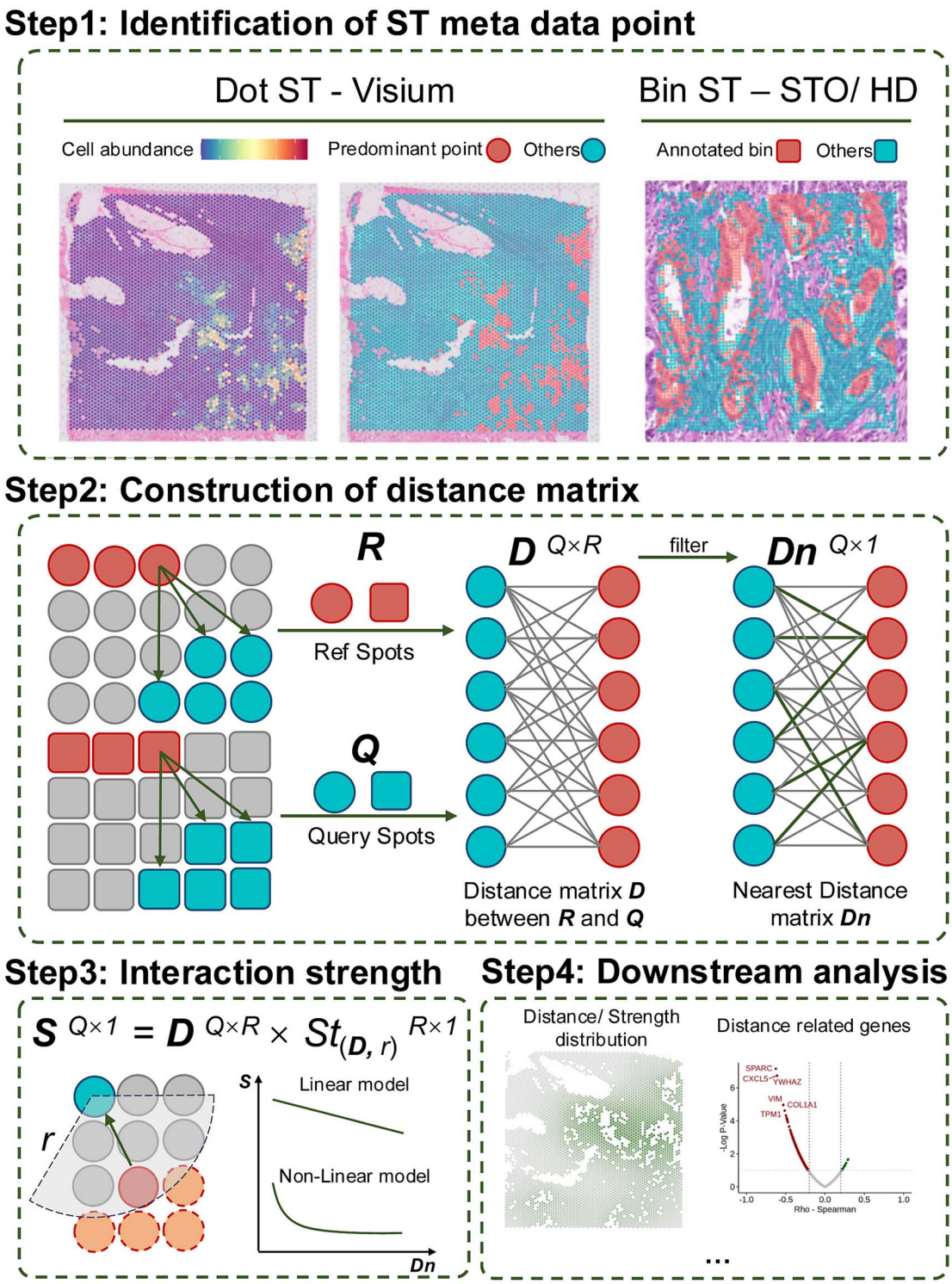

Figure 1 Schematic overview of the stGrads workflow. The framework consists of four steps: (1) selection of spatial reference spots based on biological annotations or marker expression; (2) computation of nearest distances between query and reference spots; (3) modeling of spatial influence as a distance-decay function; and (4) downstream analysis of gene expression patterns and cell-type proportions in relation to spatial proximity.

isons, we performed runtime and peak memory benchmarking to evaluate computational efficiency across these methods (Supplementary material Table S1, Fig. S1). The results show that stGrads is faster and more memory-efficient than the other frameworks, while providing the unique capability of defining query and reference cell populations for targeted spatial gradient modeling.

Although Giotto detected a larger number of genes, it lacks the ability to specify query and reference populations, limiting its capacity to model gradients originating from selected cell types. HEARTSVG and SpatialDE similarly operate on global or latent spatial signals without user-controlled gradient targeting. In contrast, stGrads allows the definition of spatial reference spots,

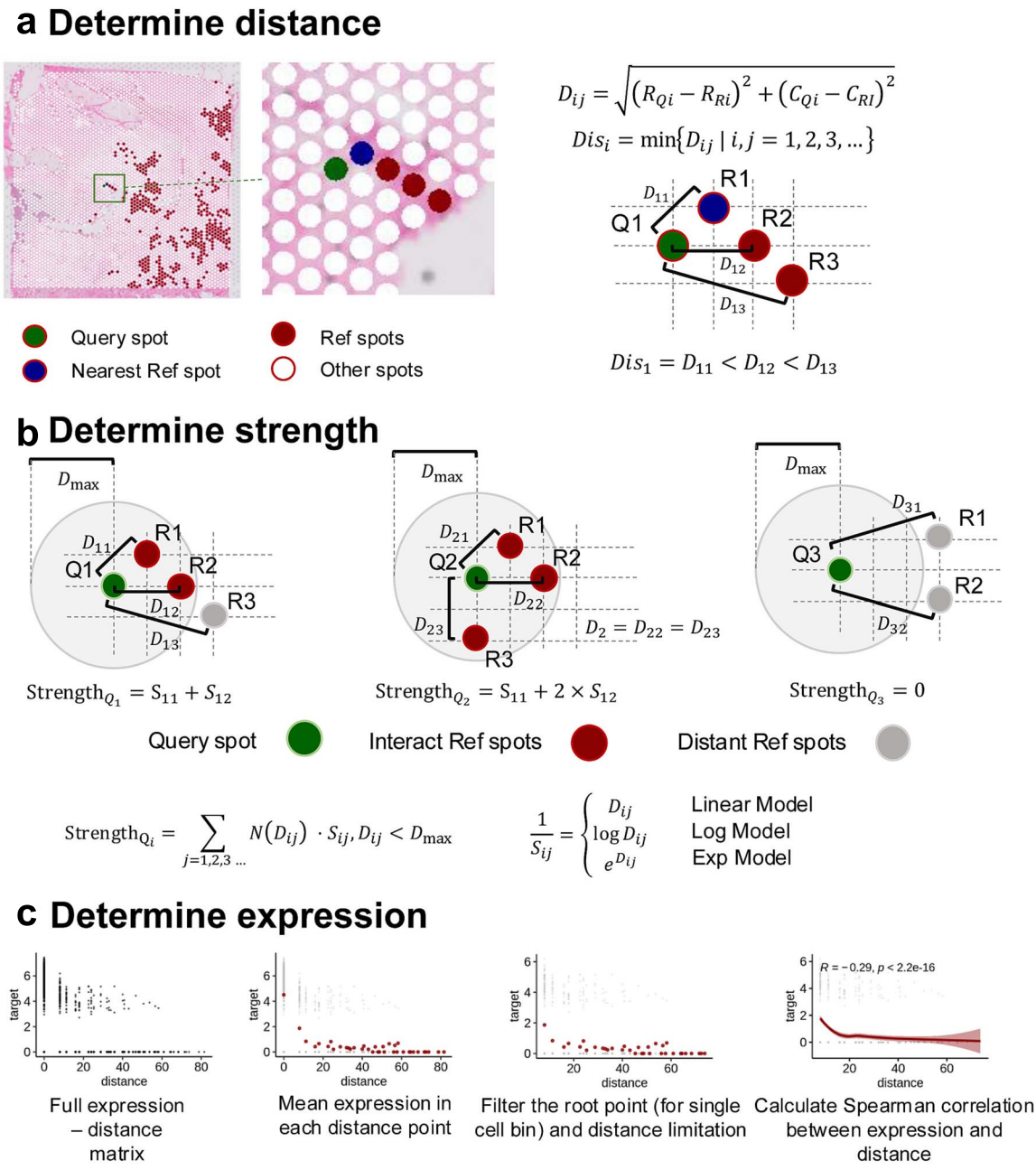

Figure 2 Core principles of the stGrads framework. (a) Euclidean distances between query and reference spots are computed, and the nearest distance is retained for each query. For each query spot i, the distance D_{ij} to the nearest reference region j is computed in Euclidean space. The strength S_{ij} is derived from D_{ij} using a user-defined decay function (linear, logarithmic, or exponential). $D_{\max}$ denotes the maximum radius within which reference regions are considered. (b) Interaction strength is estimated based on distance-weighted reference contributions, using a linear, logarithmic, or exponential decay function. (c) Gene expression or cell proportions are modeled against distance or strength, enabling identification of distance-responsive gradients.

enabling focused detection of genes whose expression is influenced by specific cell populations. A Venn diagram highlighted that stGrads captured unique genes not detected by Giotto or SpatialDE (Fig. S2A). Functional enrichment analysis of these stGrads-specific genes using FGSEA revealed enrichment in pathways associated with apoptosis, epithelial–mesenchymal transition, hypoxia, IL2–STAT5 signaling, inflammatory responses, and PI3K–AKT–mTOR signaling (Fig. S2B). Rather than yielding a broad set of candidate genes, stGrads highlighted a more focused repertoire associated with stress responses and microenvironmental adaptation, providing a precise scope of genes potentially regulated by spatial gradients.

Tumor ductal cells establish localized gradients in fibroblasts and macrophages

To investigate how tumor epithelial cells influence their surrounding microenvironment in pancreatic ductal adenocarcinoma (PDAC), we applied stGrads to simulate spatial gradients originating from Type 2 ductal cells (Duct 2 cells), a malignant epithelial subpopulation

characterized by proliferative and invasive features, elevated copy number variations (CNVs), and enrichment of proliferation-associated gene programs [18]. Duct 2 cells represent a transformed branch of the ductal epithelium and are spatially positioned at the tumor core. Ductal cells in pancreatic cancer exhibit heterogeneity, with at least one other major malignant subpopulation identified in our dataset, Type 1 ductal cells (Duct 1 cells), which display metabolic and epithelial programs. Using Duct 2 cells as spatial reference points, we examined their impact on adjacent fibroblasts and immune populations.

As shown in Fig. 3a, spatially resolved cell-type annotations reveal the distribution of Duct cell 1, Duct cell 2, macrophages, myofibroblastic cancer-associated fibroblasts (myCAFs), and T cells. We selected Duct 2 cells as the reference population to compute distance and signal strength toward surrounding cells. The distribution of query spots within a 100 μm radius from Duct 2 cells (Fig. 3b) suggests that myCAFs are more closely associated with Duct 2 cells compared to macrophages, as indicated by the density curve.

We next computed interaction strength using both linear and exponential decay models. In the Duct 2 to myCAFs (Fig. 3c), spatial signal strength is concentrated around Duct 2 clusters, suggesting a strong and localized interaction. In contrast, macrophages showed weaker and more dispersed signals (Fig. 3d), indicating a less structured spatial association. It is worth noting that relying solely on spatial distance may not fully capture the complexity of cellular crosstalk, particularly in microenvironments where immune cells interact via secreted cytokines or chemokines [19]. In such conditions, we propose that the interaction strength should be viewed as a combined manifestation of both cellular infiltration density and spatial proximity. By integrating decay functions with cell density data, our framework provides a more nuanced quantitative metric than simple distance-based clustering, allowing for a better representation of how cumulative paracrine signals shape the local tissue landscape.

To understand how gene expression changes with distance from tumor ductal cells, we examined spatial expression gradients. In the Duct 2 to myCAFs (Fig. 3e), multiple pro-fibrotic and extracellular matrix (ECM)-related genes, including *COL1A2*, and *TNFRSF12A*, showed significant negative correlations with distance, consistent with myCAFs activation near tumor cells. The volcano plot further illustrates the enrichment of these ECM genes. Previous studies have shown that myCAFs promote pancreatic cancer progression and fibrosis through ECM remodeling [20, 21], supporting the functional relevance of our observation.

Conversely, along the Duct 2 to macrophage axis (Fig. 3f), pro-inflammatory and pro-tumorigenic genes such as *COL1A1*, *SPARC*, *CXCL8*, and *CXCL5* showed higher expression in regions proximal to Duct 2 tumor cells, indicating a spatial association with macrophage-enriched zones. This pattern is consistent with prior evidence that *CXCL8* is expressed in areas enriched for tumor-associated macrophages [22]. In contrast, genes such as *APOL1* showed higher expression at increasing distances from the tumor, revealing a spatial expression gradient within the stromal microenvironment. Notably, APOL1 has been reported to promote pancreatic cancer cell proliferation and inhibit apoptosis via activation of the NOTCH1 signaling pathway [23], suggesting that its function may be modulated by spatial context. Evidence from pancreatic tissue indicates that ATG16L1, particularly its WD40 domain, is critical for non-canonical autophagy and cellular protection, highlighting its role in localized autophagic activity [24, 25]. Autophagy supports cell survival and metabolic adaptation, implying that the spatially patterned expression of ATG16L1 could indicate localized autophagic activity and immune modulation. While these observations are associative, they provide biological context for the detected gradients and warrant further functional validation.

These findings demonstrate that tumor ductal cells generate cell-type-specific, spatially resolved expression gradients, potentially influencing stromal activation and immune modulation within the tumor niche.

Spatial gradient analysis reveals lobule zonation in steady-state liver

To characterize the spatial organization of gene expression within liver lobules, we applied the stGrads framework to high-resolution Stereo-seq data from mouse liver under homeostatic conditions. Unlike layered structures such as the cerebral cortex, the liver displays a radial architecture centered on the central vein (CV), giving rise to functionally distinct regions: pericentral, midzonal, and periportal zones.

We divided the spatial landscape into nine concentric zonation layers (Fig. 4a), with Layer 1 corresponding to the pericentral region immediately surrounding the CV, Layers 3–6 representing the midzonal area, and Layers 8–9 enriched around the portal triads. This annotation is consistent with canonical models of liver zonation [26].

Taking Layer 1 as the spatial reference, we computed the Euclidean distance of each spatial bin to the Layer 1 (Fig. 4b). This continuous spatial metric captures the pericentral-to-periportal axis, enabling systematic quantification of radial gene expression gradients. Projection of all bins onto this axis revealed a progressive increase in distance from Layer 1 to peripheral layers, confirming radial zonation (Fig. 4c).

To identify genes exhibiting spatial gradients, we applied the stGrads framework to compute Spearman's correlation coefficients between gene expression levels and the Euclidean distance from Layer 1 across all spatial bins (Fig. 4d). This analysis systematically profiled gradient patterns for the entire transcriptome, revealing a spectrum of zonation-associated genes. Among them, *Ciart* showed a strong negative correlation with distance ($\rho = -0.80$, $P = 3.1 \times 10^{-4}$), peaking in the pericentral region (Layer 1) and decreasing toward the periportal area. Conversely, *Ly6d* exhibited a strong positive correlation ($\rho = 0.91$, $P < 2.2 \times 10^{-16}$), with expression increasing along the radial axis and reaching a maximum in midzonal and periportal regions (Fig. 4d).

To visually validate these patterns, we plotted the spatial expression maps of top-ranking genes, including *Ccl2* and *Ly6d* (Fig. 4e). These maps confirmed their opposing distribution patterns along the radial zonation axis, further supporting the gradient relationships detected by stGrads. This zonation pattern complements the mouse liver homeostatic atlas, although its relationship with metabolic and immune functions in the liver remains to be explored.

In summary, stGrads offers a robust computational framework for quantifying spatial gene expression gradients relative to histological landmarks such as the CV. Applied to the mouse liver, this approach reveals lobule-scale zonation patterns that reflect radially structured metabolic specialization.

Spatial gradient analysis validated colocalization of neutrophil and T cell in sweet syndrome

To further validate the feasibility of applying stGrads to real data and uncover new scientific questions, we utilized the previously

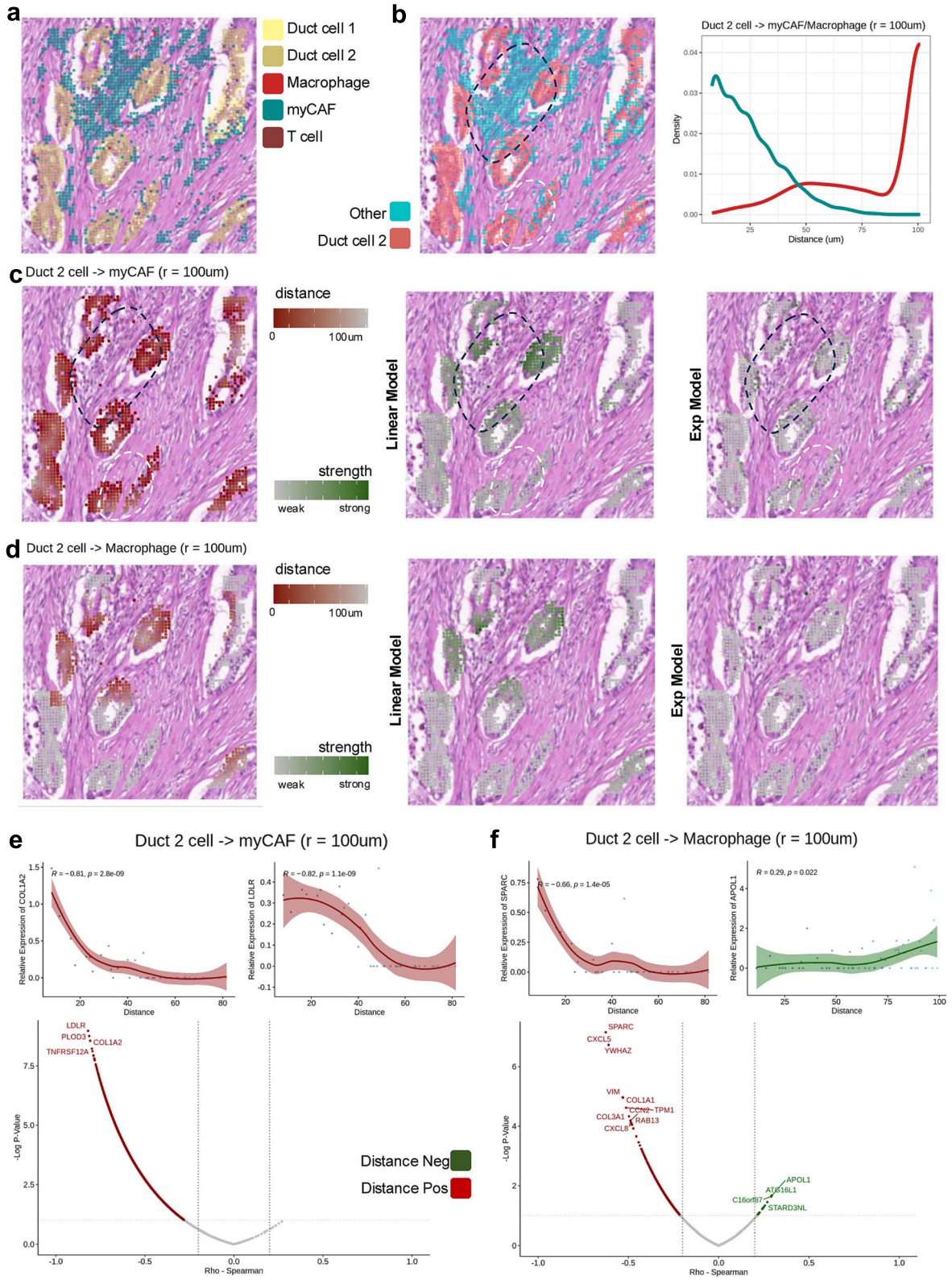

Figure 3 Spatial gradients from Duct 2 cells to myCAFs and macrophages. (a) Spatial map showing annotated cell types in a tumor section. Duct 2 cells were selected as reference cells for gradient simulation (n = 977 bins). (b) Query bins within 100 μm of Duct 2 cells and corresponding cell-type distributions; density plot shows distance distributions for myCAFs and macrophages. (c–d) Spatial maps showing distance (left) and signal strength (middle and right) from Duct 2 cells to myCAFs (c) or macrophages (d), under linear and exponential decay models. (e–f) Expression gradients from Duct 2 cells to myCAFs (e) or macrophages (f). Top: LOESS-smoothed expression versus distance plots for representative genes. Bottom: Volcano plots showing the Spearman correlation and statistical significance for all genes; distance-associated genes are labeled.

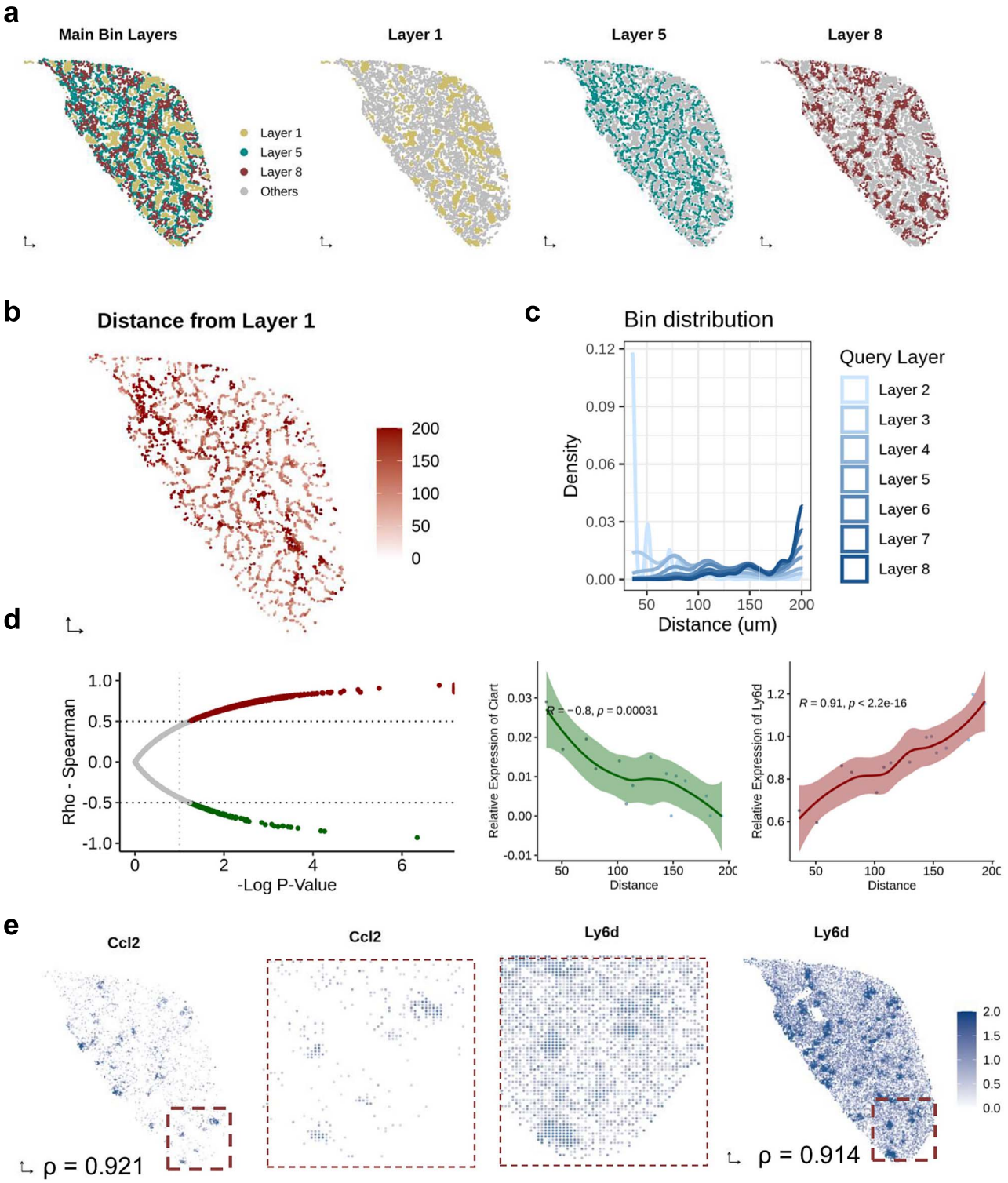

Figure 4 Spatial gradient analysis of gene expression in mouse liver tissue using stGrads. (a) Main bin classification and visualization of three representative layers (Layer 1, Layer 5, and Layer 8) defined by the spatial binning strategy (n = 19 405). (b) Spatial distribution of distances (in μm) from the reference Layer 1 across all bins. (c) Density distribution of bins at varying distances from Layer 1, stratified by query layers (Layer 2 to Layer 8), indicating their spatial positions relative to the reference. (d) Identification of genes significantly correlated with spatial distance from Layer 1. Left: Volcano plot showing Spearman correlation coefficients (ρ) versus $-$log10(*P*-value). Right: Representative distance-dependent expression trends for *Ciart* (negatively correlated) and *Ly6d* (positively correlated). Spearman correlation was calculated between gene expression and distance from Layer 1 across all bins (n = 19 405). (e) Spatial expression patterns of *Ccl2* and *Ly6d*. Insets show zoomed-in regions with notable expression enrichment. Spearman correlation coefficients (ρ) with distance from Layer 1 are shown at the bottom left of each gene panel (n = 19 405).

employed 10× Visium HD dataset of Sweet Syndrome [27]. After cellular deconvolution, we selected a 50 μm bin for analysis and mapped the distribution of neutrophils, keratinocytes, and T cells in Sweet syndrome skin tissues before and after Spesolimab treatment (Fig. 5a and b). Using the stGrads analytical framework, we identified spatially interacting regions. We observed significant co-localization of neutrophils and differentiated keratinocytes (dKC) before treatment, while after treatment, a substantial number of CD8 T cells co-localized with neutrophils, confirming the previously reported T cell delay phenomenon. To further substantiate these findings, we performed multiplex immunofluorescence staining on adjacent tissue sections, validating the spatial transcriptomic results after-treatment (Fig. 5c).

Discussion

In this study, we present stGrads, a general and flexible computational framework for modeling spatial gene expression gradients from spatial transcriptomics (ST) data. stGrads introduces several practical innovations: (1) a standardized workflow compatible with both spot-based and bin-based platforms, (2) an explicit separation of geometric distance and biologically motivated decay (interaction strength), (3) support for user-defined reference regions and fully customizable decay functions, and (4) a reproducible pipeline for identifying distance-responsive genes (DRGs) and spatially resolved cell composition gradients. By quantifying continuous, distance-dependent effects of biologically defined regions, stGrads provides a gradient-centric perspective that enables nuanced insights into how specific anatomical or pathological regions shape local gene expression and cell-state dynamics across tissues.

We demonstrate the versatility and robustness of stGrads through applications to both pathological (tumor and skin) and physiological (healthy liver) tissue contexts. In the tumor microenvironment, stGrads captured distinct spatial gradients originating from malignant ductal cells toward surrounding stromal and immune populations. By simulating interaction strengths between Duct 2 cells and neighboring myofibroblasts and macrophages, we uncovered localized expression of ECM remodeling genes such as *COL1A2* and *TNFRSF12A*, as well as pro-inflammatory mediators like *CXCL8* in macrophage-enriched zones. These findings highlight the value of spatial distance modeling in resolving niche-specific activation states that may be overlooked in global or binary comparisons. Moreover, the application of stGrads to Sweet Syndrome further substantiated its applicability to real datasets, as confirmed by multiplex immunohistochemistry.

Beyond pathological applications, we applied stGrads to high-resolution mouse liver Stereo-seq data and successfully recapitulated canonical radial zonation patterns of the hepatic lobule. By defining the pericentral zone (Layer 1) as the spatial reference, we constructed a continuous radial axis reflecting the pericentral-to-periportal organization of liver tissue. Genes such as *Ciart* and *Ly6d* exhibited monotonic expression gradients along this axis, validating the ability of stGrads to capture structured spatial gene regulation in steady-state conditions. Importantly, this zonation emerges without the need for explicit segmentation or prior knowledge of liver architecture, underscoring the generalizability of our framework.

A major strength of stGrads lies in its flexibility and modularity. It supports both spot-based (10× Visium HD) and bin-based (Stereo-seq) ST platforms, allows user-defined or unsupervised identification of reference regions, and provides multiple models for distance decay. These features make stGrads adaptable to a wide range of biological questions, including immune infiltration, tissue remodeling, developmental gradients, and environmental exposures. The integration of statistical correlation, visualization, and customizable parameters also facilitates intuitive interpretation and biological hypothesis generation.

Nevertheless, stGrads has limitations. The current implementation assumes that spatial proximity reflects biological influence, which may not account for long-range signaling events (endocrine regulation or neural projection). Additionally, the accuracy of gradient inference depends on the correct identification of spatial reference regions, which may be challenging in noisy or poorly annotated datasets. Variations in reference region boundaries or density may affect the smoothness and dynamic range of inferred gradients. In cases where reference regions are small or a substantial number of spots lie at the boundaries, we recommend expanding the selected regions or including multiple reference areas to achieve more robust results. Our framework is designed to provide a general and practical computational approach, without modeling spots outside the selected analysis regions. Future extensions of stGrads could incorporate ligand–receptor communication, cell–cell contact modeling, or machine learning–based spatial domain segmentation to enhance biological interpretability and robustness.

In conclusion, stGrads provides a quantitative, interpretable, and scalable framework for modeling spatial expression gradients in complex tissue environments. By capturing the continuous nature of spatial influence, it complements existing spatial analysis methods and enables the discovery of spatially responsive genes and localized biological processes. We anticipate that stGrads will serve as a useful tool for spatial biology research and contribute to a deeper understanding of tissue organization, disease mechanisms, and intercellular communication.

Methods

Spatial transcriptomics data acquisition

10× Visium spatial transcriptomics data used for the schematic overview of the stGrads workflow (Fig. 1) and the core principles of the stGrads framework (Fig. 2) were obtained from a previously published pancreatic cancer spatial transcriptomics study [28]. 10× Visium HD spatial transcriptomics data were generated from pancreatic cancer tissues by a commercial service provider using a customized protocol (Fig. 3). Stereo-seq data were downloaded from CNGBdb, based on a published mouse liver study [26] (Fig. 4). Raw sequencing data of human skin samples were obtained from the Genome Sequence Archive in the National Genomics Data Center, China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences [27] (Fig. 5). In addition, an in-house 10× Visium dataset, also derived from pancreatic cancer tissue, was included as a data source (Fig. S1). In addition, single-cell RNA-seq data were used to assist cell-type annotation [29].

Preprocessing and cell type annotation

stGrads directly accepts user-provided cell type annotation matrices, so no additional preprocessing of spatial transcriptomics data is required. Gene expression values are also determined by the user's input. For Visium data, we used default Cell2location deconvolution; for Visium HD data, RCTD outputs according to the Seurat workflow

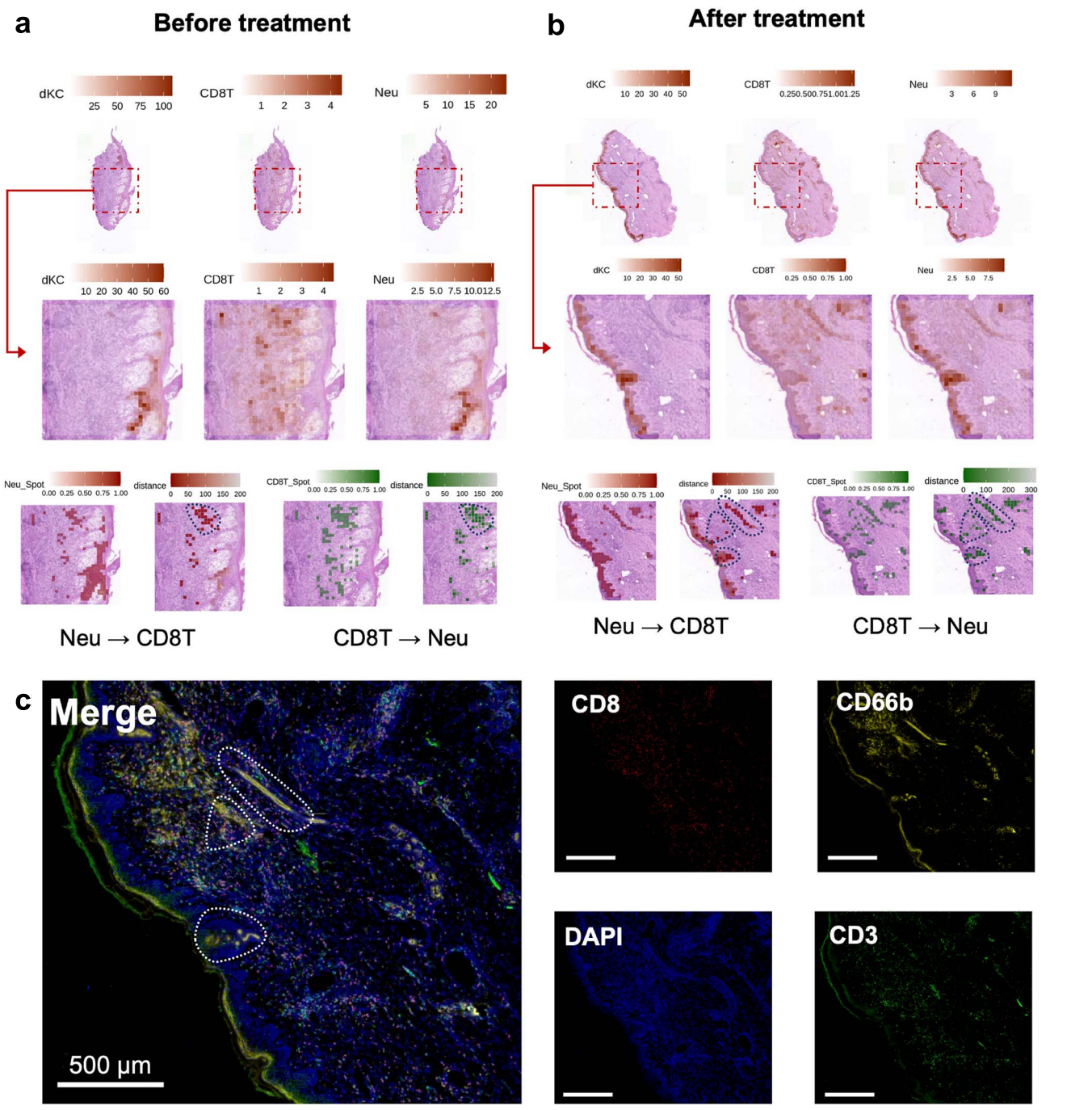

Figure 5 Spatial gradient analysis of gene expression in Sweet syndrome skin tissues using stGrads. (a–b) Main analysis pipeline of stGrads in Sweet syndrome skin tissues before (a) and after (b) Spesolimab treatment. (c) Multi-color immunohistochemical staining of neutrophil and T cell to validate the co-localization derived from stGrads. Scale bar: 500 μm.

were used; for STO data, preprocessed matrices were obtained from the BGI website [30–32].

stGrads framework implementation

The stGrads framework models how specific cell types exert spatially decaying effects on their surrounding tissue environment by integrating spatial distance and feature gradients in spatial transcriptomics data. The algorithm consists of three key computational steps: nearest distance calculation, signal decay modeling, and distance-related feature analysis (Fig. 2).

Nearest distance calculation

For each spatial unit (bin or spot), stGrads calculates the Euclidean distance between a query spot Q_i and each reference spot R_i containing the selected cell type:

$$D_{ij} = \sqrt{\left(R_{Q_i} - R_{R_j}\right)^2 + \left(C_{Q_i} - C_{R_j}\right)^2}$$

Here, R and C denote the row and column coordinates of the spots. The nearest distance to any reference spot is then defined as:

$$Dis_i = \min\{D_{ij}\}$$

This defines the proximity of each query spot to the nearest cell population of interest.

Signal strength computation

$D_{\max}$ is a user-defined parameter that specifies the maximum spatial radius within which reference regions are considered for gradient calculation.

To simulate the spatial influence exerted by nearby reference spots, stGrads aggregates distance-weighted interactions within a defined range D_{max} The strength for a query spot Q_i is calculated as:

$$\text{Strength}_{Q_i} = \sum\nolimits_{j=1,2,3\ldots} N(D_{ij}) \cdot S_{ij}, D_{ij} < D_{\max}$$

Here, S_{ij} denotes a weight assigned based on the biological relevance of each reference spot, typically set to 1 for standard references or 2 for highly confident or biologically important reference spots. $N(D_{ij})$represents the distance decay function that models how the spatial influence diminishes with increasing distance. stGrads supports three models:

$$\frac{1}{S_{ij}} = \begin{cases} D_{ij} & (Linear\ model) \\ \log D_{ij} & (Log\ model) \\ e^{D_{ij}} & (Exp\ model) \end{cases}$$

By default, an exponential decay function is used to compute spatial strength, reflecting diffusion-like biological signaling. Linear and logarithmic decay functions are additionally provided to capture alternative spatial assumptions. The framework is extensible, allowing users to implement custom decay functions when needed.

We emphasize that decay functions in stGrads are not meant to uniquely represent biological laws, but rather provide pragmatic parametric choices for capturing different spatial attenuation patterns.

Expression gradient analysis

Based on the computed spatial distance and interaction strength between query spots and reference regions, stGrads evaluates gradient-like changes in gene expression or cell-type proportions. For distance-based gradients, query spots are ordered by their nearest Euclidean distance to the reference region. Expression trends are visualized using LOESS (locally estimated scatterplot smoothing) smoothing, and the association between gene expression and distance is quantified using Spearman correlation, capturing monotonic spatial relationships without assuming linearity. For strength-based gradients, expression levels are compared against the aggregated interaction strength scores derived from distance-decay weighting. Because strength represents a continuous influence metric, stGrads typically applies linear regression and Pearson correlation to assess approximately linear relationships.

Visualization and statistical analysis

stGrads includes several visualization functions to facilitate intuitive interpretation of spatial gradients. PlotNearDis and PlotStrengthDis generate spatial heatmaps representing distance and interaction strength, respectively. PlotStrengthExpr and PlotDisExpr provide scatterplots linking gene expression to spatial metrics, overlaid with regression fits and correlation coefficients. For cell-type dynamics, PlotDisProp visualizes changes in cell-type proportions along spatial gradients. Finally, PlotDRGs summarizes distance-responsive genes via volcano plots, highlighting spatially regulated transcriptional programs.

Multiplex immunohistochemistry (mIHC)

Skin tissues from patients with Sweet syndrome underwent mIHC staining for CD66b, citrullinated histone H3 (citH3), CD3, and CD8, as well as hematoxylin and eosin (HE) staining. Standard operating procedures were followed for both mIHC and HE staining. Visualization of IHC signals was performed using the EnVision FLEX+ detection system (catalog K8009, Dako, Copenhagen, Denmark). The Think Color staining kit (catalog FH35050R-p, FreeThinking, Nanjing, China) was used for mIHC staining according to the manufacturer's protocol. Primary antibodies used for mIHC included anti-CD66b (1:40000 dilution, catalog ab207718, Abcam), anti-citH3 (1:2000 dilution, catalog ab5103, Abcam), anti-CD3 (1:2000 dilution, catalog ab16669, Abcam), and anti-CD8 (1:500 dilution, catalog 85336S, Cell Signaling Technology).

Key Points

- stGrads quantifies distance-dependent gene expression gradients in spatial transcriptomics data.
- The framework is compatible with both low- and high-resolution spatial platforms with flexible reference-region definition and customizable decay functions.
- stGrads enables identification of distance-responsive genes from defined tissue regions.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 32571524, 82141113, 82373116, and 82473086), the Beijing Natural Science Foundation (Grant No. 5232014), and the Key Research and Development Program of Ningxia Hui Autonomous Region (Grant No. 2022BEG02055). We also appreciate the support of the high-performance computing platform from center for bioinformatics, National Infrastructures for Translational Medicine, Institute of Clinical Medicine & Peking Union Medical College Hospital, CAMS&UMC.

Author contribution

Yifan Fu (Methodology, Data analysis, Investigation, Visualization), Fan Zhang (Investigation, Testing, Visualization, Writing—Original draft, Writing—Review & Editing), Feifan Zhang (Testing, Visualization), Taiping Zhang (Supervision, Writing—Review & Editing, Funding acquisition), and Wei Chen (Supervision, Writing—Review & Editing, Funding acquisition)

Supplementary material

Supplementary material is available at *Briefings in Bioinformatics* online.

Conflict of interest

We declare no conflicts of interest in regard to this manuscript.

Data availability

The datasets analyzed or generated during this study are available through public repositories or upon request. Raw sequencing data for human skin samples have been deposited in the GSA-Human under accession number PRJCA031241 (https://ngdc.cncb.ac.cn/gsa-human) [27]. Stereo-seq mouse liver data are accessible via CNGBdb (https://db.cngb.org/stomics/lista/) [26]. Other publicly available spatial and single-cell transcriptomics datasets were obtained from previously published studies [28, 29]. The 10× Visium HD pancreatic cancer datasets are available from the corresponding author upon reasonable request at chenw123@buaa.edu.cn.

Ethics statement

This study was approved by the Institutional Ethics Committee of Peking Union Medical College Hospital (Ethics code: I-24PJ1844) in accordance with the ethical guidelines of the Institutional Ethics Committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Code availability

The stGrads framework is implemented in R and is available upon request or at https://github.com/yifanfu01/stGrads, with full documentation and example datasets provided for reproducibility. The online user guide and tutorials can be accessed at https://stgrads.readthedocs.io

References

1. Ståhl PL, Salmén F, Vickovic S *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 2016;**353**:78–82. https://doi.org/10.1126/science.aaf2403
2. Janesick A, Shelansky R, Gottscho AD *et al.* High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat Commun* 2023;**14**:8353. https://doi.org/10.1038/s41467-023-43458-x
3. Shi Q, Chen YH, Li Y *et al.* Cross-tissue multicellular coordination and its rewiring in cancer. *Nature* 2025;**643**:529–38. https://doi.org/10.1038/s41586-025-09053-4
4. Elhanani O, Ben-Uri R, Keren L. Spatial profiling technologies illuminate the tumor microenvironment. *Cancer Cell* 2023;**41**:404–20. https://doi.org/10.1016/j.ccell.2023.01.010
5. Chen A, Liao S, Cheng MN *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 2022;**185**:1777–92.e21. https://doi.org/10.1016/j.cell.2022.04.003
6. Wang JJ, Ye F, Chai HX *et al.* Advances and applications in single-cell and spatial genomics. *Sci China Life Sci* 2025;**68**:1226–82. https://doi.org/10.1007/s11427-024-2770-x
7. Schrom EC, Mccaffrey EF, Sreejithkumar V *et al.* Spatial patterning analysis of cellular ensembles (SPACE) finds complex spatial organization at the cell and tissue levels. *Proc Natl Acad Sci USA* 2025;**122**:e2412146122. https://doi.org/10.1073/pnas.2412146122
8. Rao A, Barkley D, França GS *et al.* Exploring tissue architecture using spatial transcriptomics. *Nature* 2021;**596**:211–20. https://doi.org/10.1038/s41586-021-03634-9
9. Greenwald AC, Darnell NG, Hoefflin R *et al.* Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* 2024;**187**:2485–2501.e26. https://doi.org/10.1016/j.cell.2024.03.029
10. He SY, Jin YN, Nazaret A *et al.* Starfysh integrates spatial transcriptomic and histologic data to reveal heterogeneous tumor-immune hubs. *Nat Biotechnol* 2025;**43**:223–35. https://doi.org/10.1038/s41587-024-02173-8
11. Sun PS, Bush SJ, Wang SB *et al.* STMiner: gene-centric spatial transcriptomics for deciphering tumor tissues. *Cell Genom* 2025;**5**:100771. https://doi.org/10.1016/j.xgen.2025.100771
12. Svensson V, Teichmann SA, Stegle O. SpatialDE: identification of spatially variable genes. *Nat Methods* 2018;**15**:343–6. https://doi.org/10.1038/nmeth.4636
13. Dries R, Zhu Q, Dong R *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol* 2021;**22**:78. https://doi.org/10.1186/s13059-021-02286-2
14. Kicheva A, Pantazis P, Bollenbach T *et al.* Kinetics of morphogen gradient formation. *Science* 2007;**315**:521–5. https://doi.org/10.1126/science.1135774
15. Lander AD. Morpheus unbound: reimagining the morphogen gradient. *Cell* 2007;**128**:245–56. https://doi.org/10.1016/j.cell.2007.01.004
16. Wartlick O, Kicheva A, Gonzalez-Gaitan M. Morphogen gradient formation. *Cold Spring Harb Perspect Biol* 2009;**1**:a001255. https://doi.org/10.1101/cshperspect.a001255
17. Yuan X, Ma Y, Gao R *et al.* HEARTSVG: a fast and accurate method for identifying spatially variable genes in large-scale spatial transcriptomics. *Nat Commun* 2024;**15**:5700. https://doi.org/10.1038/s41467-024-49846-1
18. Peng JY, Sun BF, Chen CY *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res* 2019;**29**:725–38. https://doi.org/10.1038/s41422-019-0195-y
19. Scarpa A, Jung Y, Hamid A *et al.* Trained immunity induced by oxidized low-density lipoprotein is dependent on Glutaminolysis. *FASEB J* 2025;**39**:e70774. https://doi.org/10.1096/fj.202500802R
20. Li X, Zhou JB, Wang X *et al.* Pancreatic cancer and fibrosis: targeting metabolic reprogramming and crosstalk of cancer-associated fibroblasts in the tumor microenvironment. *Front Immunol* 2023;**14**:1152312. https://doi.org/10.3389/fimmu.2023.1152312
21. Huang HC, Brekken RA. Recent advances in understanding cancer-associated fibroblasts in pancreatic cancer. *Am J Physiol-Cell Ph* 2020;**319**:C233–43. https://doi.org/10.1152/ajpcell.00079.2020
22. Waugh DJJ, Wilson C. The Interleukin-8 pathway in cancer. *Clin Cancer Res* 2008;**14**:6735–41. https://doi.org/10.1158/1078-0432.Ccr-07-4843
23. Lin JW, Xu ZW, Xie JJ *et al.* Oncogene APOL1 promotes proliferation and inhibits apoptosis via activating NOTCH1 signaling pathway in pancreatic cancer. *Cell Death Dis* 2021;**12**:760. https://doi.org/10.1038/s41419-021-03985-1
24. Wei F, Qin N, Chen Y *et al.* ATG16L1 membrane recruitment in autophagy. *Crit Rev Biochem Mol Biol* 2025;**60**:107–22. https://doi.org/10.1080/10409238.2025.2521321
25. Chvanov M, Voronina S, Jefferson M *et al.* Deletion of the WD40 domain of ATG16L1 exacerbates acute pancreatitis, abolishes LAP-like non-canonical autophagy and slows trypsin degradation. *Autophagy* 2025;**21**:210–22. https://doi.org/10.1080/15548627.2024.2392478
26. Xu J, Guo P, Hao S *et al.* A spatiotemporal atlas of mouse liver homeostasis and regeneration. *Nat Genet* 2024;**56**:953–69. https://doi.org/10.1038/s41588-024-01709-7
27. Fu YF, Wu C, Pang ZY *et al.* Spatiotemporal transcriptomics deciphered neutrophil-activated interleukin-36 as driver and

effective therapeutic target in sweet syndrome. *Brit J Dermatol* 2026;**194**:311–25. https://doi.org/10.1093/bjd/ljaf396

28. Khaliq AM, Rajamohan M, Saeed O *et al.* Spatial transcriptomic analysis of primary and metastatic pancreatic cancers highlights tumor microenvironmental heterogeneity. *Nat Genet* 2024;**56**:2455–65. https://doi.org/10.1038/s41588-024-01914-4
29. Werba G, Weissinger D, Kawaler EA *et al.* Single-cell RNA sequencing reveals the effects of chemotherapy on human pancreatic adenocarcinoma and its tumor microenvironment. *Nat Commun* 2023;**14**:797. https://doi.org/10.1038/s41467-023-36296-4
30. Kleshchevnikov V, Shmatko A, Dann E *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat Biotechnol* 2022;**40**:661–71. https://doi.org/10.1038/s41587-021-01139-4
31. Cable DM, Murray E, Zou LLS *et al.* Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat Biotechnol* 2022;**40**:517–26. https://doi.org/10.1038/s41587-021-00830-w
32. Hao YH, Stuart T, Kowalski MH *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* 2024;**42**:293–304. https://doi.org/10.1038/s41587-023-01767-y

Briefings in Bioinformatics, 2026, **27**, bbag201
https://doi.org/10.1093/bib/bbag201
Problem Solving Protocol