Differential colocalization of cell-type pairs in single-cell spatial omics with DC-SPOMIC

Jacob Chang[1], Almudena Espín Pérez[1], Perla Molina[1], Lu Tian[1,2], Sylvia Plevritis[1,3,*]

[1]Department of Biomedical Data Science, Stanford University, CA, USA
[2]Department of Statistics, Stanford University, CA, USA
[3]Department of Radiology, Stanford University, CA, USA

*Corresponding author: plevriti@stanford.edu

Abstract

Advancements in single-cell spatial omics technologies allow unprecedented exploration of tissue microenvironments and the discovery of tissue-specific and cell type-specific, spatially-resolved molecular patterns associated with development, pathogenesis, and drug response. However, the process of pooling spatial estimates and findings across samples while hypothesis testing for meaningful differences between groups poses challenges under limited sample size and when the degree of heterogeneity within and between samples is high. To address this challenge, we introduce DC-SPOMIC, a tool for identifying differentially colocalized cell type pairs between groups of spatial omics samples. DC-SPOMIC uses local spatial statistics of each cell in a sample to calculate colocalization between pairs of cell types and spatial bootstrapping to estimate the degree of intra-sample variation of the colocalization estimate. The estimated intra-sample variances per sample are used in a random-effects meta-analysis model, which pools cell type pair colocalization estimates across samples while estimating the overall inter-sample heterogeneity. The random-effects output provides a pooled estimate of colocalization and its precision, which can be used for downstream analysis such as hypothesis testing of differential colocalization between groups of spatial omic samples. We demonstrate the robustness of colocalization estimation and pooling across samples using DC-SPOMIC on simulated data from spatial point process models. When used to analyze distinct histological groups of colorectal cancer samples generated by CODEX, DC-SPOMIC identifies seven statistically significant pairwise spatial colocalization differences amongst immune cell populations. By quantifying both intra- and inter- sample variability, DC-SPOMIC offers potential to uncover significant spatial interactions in complex tissues and across heterogeneous datasets.

Introduction

In recent years, single-cell spatial omics technologies have transformed tissue biology by enabling spatially resolved, cell-specific molecular profiling [1]. These platforms simultaneously capture the molecular profiles and physical locations of thousands of cells in situ, making it possible to quantify tissue architecture, investigate cell–cell interactions, and identify spatial biomarkers associated with development, disease progression, and therapeutic response [2, 3, 4, 5]. Their adoption is rapidly expanding across research fields, from developmental biology to oncology, promising new insights into how spatial cellular organization underlies biological function and clinical outcomes.

A central question in spatial omics analysis is how cell types are organized relative to one another. A widely used measure is colocalization, which has diverse mathematical definitions but generally quantifies the spatial proximity or co-occurrence of distinct pairs of cell types. Changes in colocalization can reveal biologically meaningful insights in tissue architecture — for example, increased interaction between cytotoxic T cells and tumor cells in responders to immunotherapy, or altered proximity of stromal and epithelial cells in early disease progression [6, 7]. By summarizing the complex spatial arrangements of cell types within a tissue into interpretable metrics, colocalization analysis offers a powerful lens for linking tissue microenvironments to functional properties and clinical outcomes.

Despite the influx of rich, spatial single-cell data, the robust statistical analysis of spatial omics data remains challenging. Data are inherently heterogeneous, with variation arising from both biological factors (e.g., patient-specific microenvironments, regional tissue differences) and technical factors (e.g., sample processing, image acquisition, post-processing). These data exhibit a nested structure in which spatial and molecular measurements from thousands of cells exhibit their own variability within a sample, while samples themselves vary within clinical groups and across patients. Most existing tools either ignore this dependency structure or apply methods designed for independent observations, leading to underestimated uncertainty and inflated significance [8, 9, 10, 11, 12]. Consequently, there is a pressing need for statistical frameworks that explicitly model multi-level variability and enable rigorous, reproducible detection of spatial patterns that generalize across patients.

To address this need, we developed DC-SPOMIC, an R package and computational framework for robust differential colocalization analysis in spatial omics data. We designed a two-stage approach that first estimates intra-sample variability in cell-type pair colocalization using spatial bootstrapping, and then quantifies inter-sample heterogeneity using random-effects meta-analysis. By explicitly modeling hierarchical variability, DC-SPOMIC enables statistically principled pooling of evidence across samples and accurate detection of colocalization differences between conditions and tissues.

This paper presents the DC-SPOMIC framework, detailing its methodological components and

demonstrating its application to both simulated datasets and multiplexed immunofluorescence tissue microarrays from colorectal cancer. DC-SPOMIC is broadly applicable to spatially-indexed point pattern data, supports diverse colocalization metrics, operates across single-cell spatial omics platforms, and facilitates network-based analyses to uncover higher-order spatial patterns.

Results

Overview of DC-SPOMIC framework

Input data

DC-SPOMIC operates on spatially-indexed point pattern data such as single-cell spatial omics datasets where each cell has been assigned a cell type label (Fig. 1A). Each sample is represented as a table with one row per cell, including sample identity, centroid coordinates (x, y), and cell type annotation. To support group-wise comparisons, the dataset also contains sample-metadata specifying a categorical group assignment (e.g., treatment vs. control, wild-type vs. mutant, pre- vs. post-treatment). This simple and standardized format enables DC-SPOMIC to process inputs from diverse spatial omics platforms.

Colocalization metrics for spatial feature characterization

A common objective of spatial omics analysis is to quantify how different cell types are arranged relative to one another within tissue. Colocalization, which measures the spatial proximity between cells of the same type (homotypic) or between different types (heterotypic), provides a powerful summary of these relationships. Such measures can reveal biologically meaningful organization, for example, immune cells clustering with cancer cells in drug-responsive tumors, or separately in nonresponsive tumors. DC-SPOMIC is designed to work flexibly with a variety of colocalization statistics, including user-defined metrics and those established in the literature. In the present study, and by default in the `dcspomic` R package, we use Ripley's homogeneous cross L-function with translation edge correction, as implemented in `spatstat` [13]. This statistic is related to the cross K-function, and is widely used in spatial point pattern analysis because it accounts for boundary effects, stabilizes variance for improved interpretability, and can be interpreted as a scale-dependent measure of spatial association (Supplementary Methods). The cross K-function (Equation 1) iterates through $i' \in I$ and $j' \in J$, the sets of all cells of types i and j, respectively. The cross K-function counts the number of type j cells within a distance r of each type i cell, normalized by the population sizes of each cell type category N_i and N_j, and the total tissue area W with $e(i', j')(r)$ providing the necessary edge correction for tissue shape and boundaries (Supplementary Methods):

$$K_{(i,j)}(r) = \frac{|W|}{N_i N_j} \sum_{i' \in I} \sum_{j' \in J} 1 \left[d_{(i',j')} \leq r \right] e_{(i',j')}(r) \tag{1}$$

The cross L-function (Equation 2) transforms the cross K-function based on assumptions of the distributions of cells under complete spatial randomness (Supplementary Methods, Fig. B.1).

$$L_{(i,j)}(r) = \sqrt{\frac{K_{(i,j)}(r)}{\pi}} \tag{2}$$

Cross L values above r indicate greater-than-expected colocalization and values below r indicate spatial avoidance. Notably, because $L_{(i,j)}(r) \neq L_{(j,i)}(r)$, there are m^2 ordered colocalization pairs for m distinct cell types in a sample. For example, the colocalization of cell types i and j is not necessarily equivalent to the colocalization of cell types j and i.

Spatial bootstrapping to estimate intra-sample variation of colocalization

While each spatial omics sample can contain thousands of cells, the number of samples is often limited by high assay costs and the challenges of acquiring and handling tissue. Typical analyses rely solely on a single point estimate of colocalization per sample, overlooking intra-sample variability of colocalization estimates and thus underestimating uncertainty. To address the need for more robust estimation of colocalization, DC-SPOMIC uses the Loh spatial bootstrap [13, 14, 15], which is designed for data exhibiting spatial correlation. For a given cell type pair, the method first calculates a local statistic for each cell, representing its contribution to the global colocalization measure. These local statistics are then resampled with replacement, averaged to form a new global estimate represented by the mean, and this process is repeated many times for a given cell type pair to generate bootstrap sampling distributions for that cell type pair. The variance of this distribution provides a direct estimate of intra-sample uncertainty of colocalization for each cell type pair (Fig. 1B(i–ii)). In other words, through many bootstrap iterations a sampling distribution is generated and the variance of this distribution is taken as an estimate for the uncertainty and intra-sample variance of the colocalization statistic.

Pooling colocalization estimates using random-effects meta-analysis

In order to generalize colocalization analysis across patients or experimental replicates, it is essential to account for both intra- and inter-sample variability. DC-SPOMIC achieves this by pooling the colocalization estimates of a cell type pair across samples within a larger group. The pooling is performed through random-effects meta-analysis models, which can be viewed as a two-level hierarchical framework (Fig. 1B(iii)). At the first level, the observed colocalization $\tilde{\theta}_{p,k}$ between a pair of cell types p for sample k, is modeled as normally distributed around a true but unknown value $\theta_{p,k}$, with sample-specific variance $\sigma^2_{p,k}$ estimated from the spatial bootstrap:

$$\tilde{\theta}_{p,k} \sim \mathcal{N}\left(\theta_{p,k}, \sigma^2_{p,k}\right) \tag{3}$$

At the second level, the true values $\theta_{p,k}$ for samples within a group are themselves assumed to follow a normal distribution with group mean μ_p and inter-sample variance τ_p^2:

$$\theta_{p,k} \sim \mathcal{N}\left(\mu_p, \tau_p^2\right) \tag{4}$$

This yields the marginal model for each sample k

$$\tilde{\theta}_{p,k} \sim \mathcal{N}\left(\mu_p, \sigma_{p,k}^2 + \tau_p^2\right) \tag{5}$$

where μ_p is the mean colocalization for cell type pair p across samples within a group, and τ_p^2 captures the inter-sample heterogeneity in colocalization between a pair of cell types.

In practice, the parameters of these distributions are unknown and must be estimated. To compute these estimates and thereby account for intra- and inter-sample variability, DC-SPOMIC uses spatial bootstrapping to estimate within-sample variance of colocalization and random-effects meta-analysis to pool colocalization estimates across samples in a group (Fig. 1B(iii)). To do this, the observed colocalization and intra-sample variance estimates of a cell type pair p in a given sample k from spatial bootstrapping are given as inputs to random-effects meta-analysis $(\hat{\theta}_{p,k}, \hat{\sigma}_{p,k}^2)$. DC-SPOMIC estimates μ_p and τ_p^2 using the Sidik-Jonkman estimator by default [16, 17, 18], although other estimators in `metafor` are supported [19]. The Sidik-Jonkman estimator uses each $\hat{\sigma}_{p,k}^2$ to estimate the residual variance associated with inter-sample heterogeneity, providing our $\hat{\tau}_p^2$ estimate. The random-effects meta-analysis model then uses all the estimates to compute group pooled estimates of colocalization mean $\hat{\mu}_p$. This combined information then provides estimated standard error of the group-pooled mean μ_p:

$$SE(\hat{\mu}_p) = \sqrt{\frac{1}{\sum_{k=1}^{K}(\hat{\tau}_p^2 + \hat{\sigma}_{p,k}^2)^{-1}}} \tag{6}$$

which incorporates both sources of variation in the observed colocalization values. This model fitting procedure is performed separately for each group and cell type pair $p \in \mathcal{P}$, where $\mathcal{P}$ is the set of cell type pairs.

Testing between-group differential colocalization

To identify differentially colocalized cell type pairs between groups of samples, DC-SPOMIC performs separate random-effects meta-analyses for each cell type pair p across all the samples within a group, producing group-specific estimates $\hat{\mu}_p^{(1)}$ and $\hat{\mu}_p^{(2)}$ in group 1 and group 2, respectively. These are compared using a Wald test,

$$Z = \frac{\hat{\mu}_p^{(1)} - \hat{\mu}_p^{(2)}}{\sqrt{SE(\hat{\mu}_p^{(1)})^2 + SE(\hat{\mu}_p^{(2)})^2}} \tag{7}$$

with two-sided p-values derived from the standard normal distribution. Multiple testing correction is applied using the false discovery rate (FDR), and cell type pairs with $p_{\text{adjusted}} < 0.05$ are reported as significantly differentially colocalized (Fig. 1B(iv)).

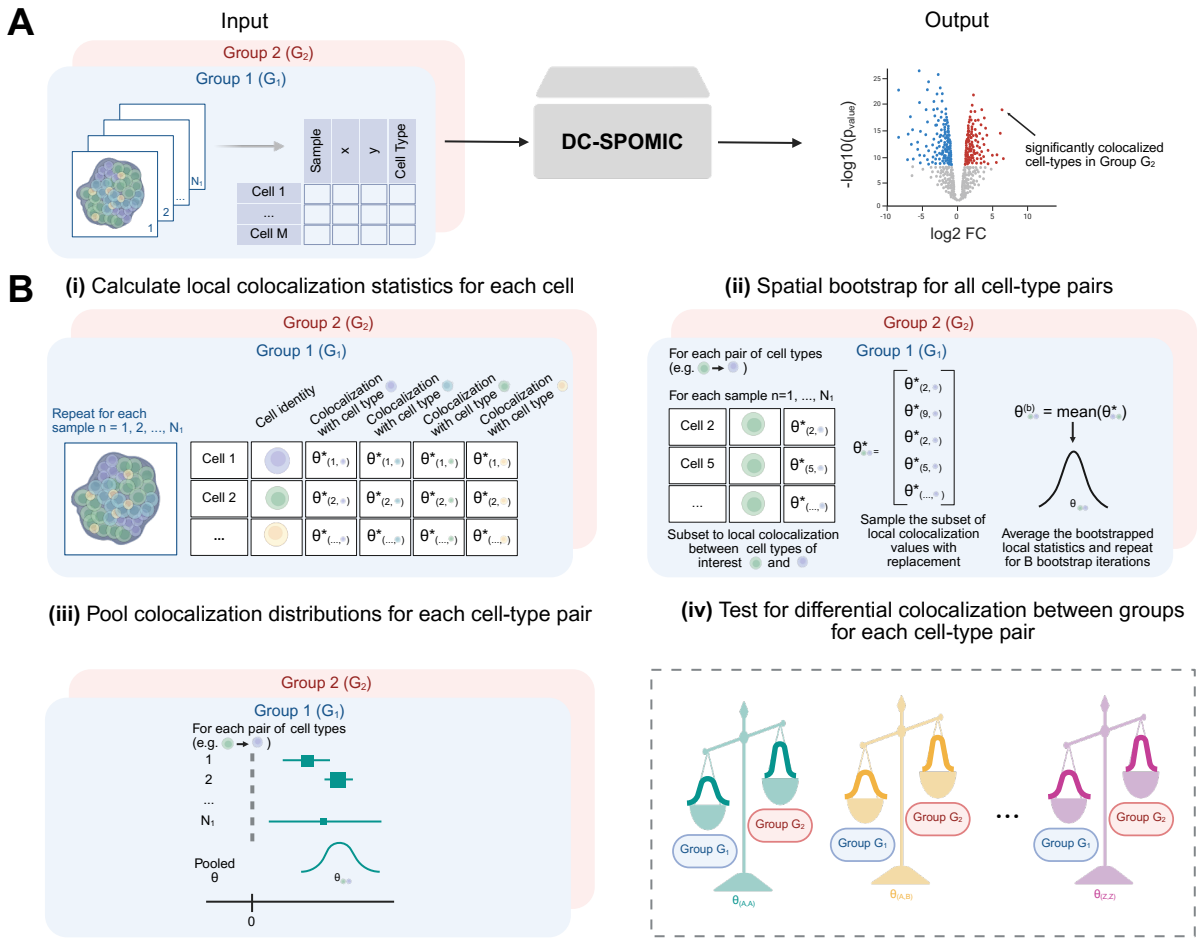

Figure 1. Overview of DC-SPOMIC framework. **(A)** Inputs are groups of data frames representing single-cell spatial omics samples from containing cell locations and cell type annotations. DC-SPOMIC identifies differentially colocalized cell type pairs between groups. **(B)** (i) For each sample, the colocalization between cell types is calculated for every cell type pair. (ii) For each cell type pair, spatial bootstrapping is performed by sampling local spatial statistics with replacement. The mean of these bootstrapped spatial statistics represents the global estimate of colocalization for the cell type pair in the sample. In other words, through many bootstrap iterations a sampling distribution is generated and the variance of this distribution is taken as an estimate for the uncertainty and intra-sample variance of the colocalization statistic. (iii) For each cell type pair, the estimates of colocalization mean and intra-sample variance can be pooled across samples within a group using random-effects meta-analysis. (iv) For each cell type pair, the pooled colocalization value in each group can be tested for differential colocalization across groups.

Benchmarking DC-SPOMIC performance with spatial simulation studies

We evaluated the performance of DC-SPOMIC using two spatial simulation studies in which the ground truth for colocalization between cell type pairs is empirically derived. In both studies, spatial point process (SPP) models were used to generate synthetic spatial omics data. A SPP is well-suited for image generation as it is a stochastic process defined over a two-dimensional spatial domain D and parameterized by an intensity function $\lambda_i(x, y) \in \mathbb{R}^+$ for each cell type i, specifying the expected number of such cells at a given location. By superimposing the intensity functions for multiple cell types, we created multi-type spatial point patterns with specified spatial structure.

Evaluating intra-sample colocalization variance estimation via spatial bootstrapping

Since the first step of DC-SPOMIC is to quantify intra-sample variability in colocalization, our first simulation study assessed whether spatial bootstrapping could recover the true intra-sample variance colocalization from only a single observed point pattern. We defined three distinct SPP configurations, each representing a different spatial organization of $m = 5$ cell types, including both common and rare cell types (Fig. 2A). In the homogeneous pattern, all cell types were distributed randomly over the spatial domain. In the divergent pattern, two cell types localized to opposite regions, while the others were randomly distributed. In the clustered pattern, one cell type was placed randomly, and a second type was positioned as offspring around each of its points, mimicking processes such as cell recruitment, with the remaining types again distributed randomly. See Supplementary Methods for details.

For each sample, represented by a specified SPP, Ω, we generated 1000 independent point pattern realizations $\omega_1, \ldots, \omega_{1000}$, providing empirical estimates over large sample size of the sample-specific colocalization means $\theta_{p,k}$ and inta-sample variances $\sigma^2_{p,k}$ for each cell type pair p in sample k, serving as the ground truth. To mimic realistic spatial omics data, in which only one realization per sample is observed, we then applied spatial bootstrapping with 100 iterations to a single realization, yielding bootstrap estimates $\hat{\sigma}^2_{p,k}$ of intra-sample variance. We compared these estimates to the ground truth by computing the Pearson correlation across all cell type pairs, and repeated this procedure 1000 times (spatially bootstrapping each of the $\omega_1, \ldots .\omega_{1000}$ to assess variability in the correlations (Supp. Fig. B.2). As shown in Fig. 2B, the bootstrap-derived variance estimates showed high concordance with ground truth for the homogeneous and clustered patterns (median correlations $r = 0.85, 0.87$, respectively), and moderate concordance for the divergent pattern (median correlation $r = 0.62$). A follow-up analysis was performed by repeating the steps without considering rare cell types, and found that Pearson correlations remained moderate in the absence of rare cell type colocalization pairs, but with tighter distributions around the median (median correlations $r = 0.77, 0.44, 0.62$ for the homogeneous, divergent, and cluster patterns, respectively) (Fig. 2C).

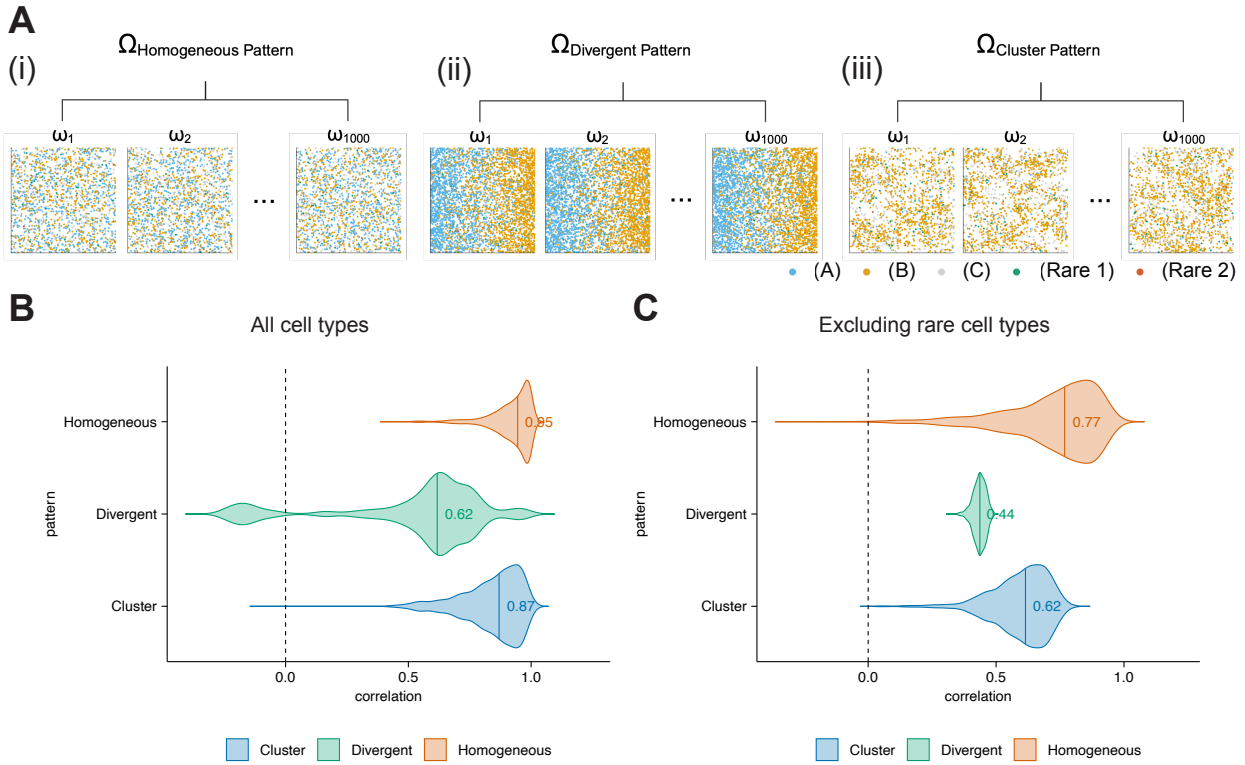

Figure 2. Simulation study for validation of intra-sample variance estimates from spatial bootstrapping. **(A)** Examples of sampling multiple realizations from three different spatial point processes. (i) Homogeneous spatial pattern where cell types are homogeneously (randomly) distributed. (ii) Divergent spatial pattern where cell types A and B localize to opposite sides and cell types C, Rare 1, and Rare 2 are homogeneously distributed. (iii) Cluster spatial pattern where cell type A is sample homogeneously, cell type B are sampled as offspring around each type A point, and cell types C, Rare 1, and Rare 2 are homogeneously distributed. **(B)** Ground-truth colocalization variance for each cell type pair was compared to estimates of the variance produced by DC-SPOMIC with access to only a single point pattern. The Pearson correlation between ground-truth and DC-SPOMIC estimates were taken over 1000 iterations, producing the violin plots. **(C)** The process in (B) was repeated without rare cell types.

Evaluating inter-sample variability estimation in heterogeneous settings

The second simulation study extended the validation to DC-SPOMIC's hierarchical pooling step, which jointly estimates the group-level mean colocalization μ_p and inter-sample variance τ_p^2. We designed a hierarchical simulation framework to emulate inter-sample heterogeneity in a controlled manner (Fig. 3A). Starting from a baseline spatial pattern $\mathbf{\Omega}$, we introduced stochastic spatial characteristics to create K variant spatial point process models $\Omega_1, \ldots, \Omega_K$, each representing a different sample with its own underlying colocalization structure. Multiple realizations $\omega_1, \ldots, \omega_N$ were drawn from each variant model to represent repeated measurements or replicates from the same sample, thereby capturing both intra- and inter-sample variability.

For a given cell type pair p, the mean colocalization within sample k was calculated as

$$\theta_{p,k} = \frac{1}{N} \sum_{n=1}^{N} \tilde{\theta}_{p,k}^{(n)} \tag{8}$$

where N is the number of repeated realizations and $\tilde{\theta}_{p,k,n}$ is the colocalization statistic for cell type pair p for realization n of sample k. The group mean colocalization is then

$$\mu_p = \frac{1}{K} \sum_{k=1}^{K} \theta_{p,k} \tag{9}$$

and the inter-sample variance is

$$\tau_p^2 = \mathrm{Var}\left([\theta_{p,k}]_{k=1,2,\ldots,K} \right) \tag{10}$$

These quantities served as the ground truth for benchmarking, aided by the large sample sizes possible through the simulation framework. DC-SPOMIC was then applied in a realistic setting where only one realization from each sample $\omega_k^{(1)}$ was available, using spatial bootstrapping to estimate intra-sample variance and random-effects meta-analysis to estimate $\hat{\mu}_p$ and $\hat{\tau}_p^2$ (Fig. 3B). For comparison, the sample mean and variance of colocalization from single realizations $\tilde{\theta}_{p,k}^{(1)}$ provide naive estimators, where p is the cell type pair, k is the sample, and the superscript (1) denotes that the colocalization was computed on a single realization.

$$\hat{\mu}_{\mathrm{Naive}_p} = \frac{1}{K} \sum_{k=1}^{K} \tilde{\theta}_{p,k}^{(1)} \tag{11}$$

$$\hat{\tau}_p^2 = \mathrm{Var}([\tilde{\theta}_{p,k}^{(1)}]_{k=1}^{K}) \tag{12}$$

We evaluated performance for sample sizes ranging from $K = 2$ to $K = 100$. Across all sample sizes, DC-SPOMIC consistently achieved lower root mean squared error than naive estimators for both μ_p and τ_p^2 for $p \in \{(A,A),(A,B),(B,A),(B,B)\}$ (Fig. 3E–F), demonstrating its ability to recover more accurate group-level parameters in heterogeneous settings.

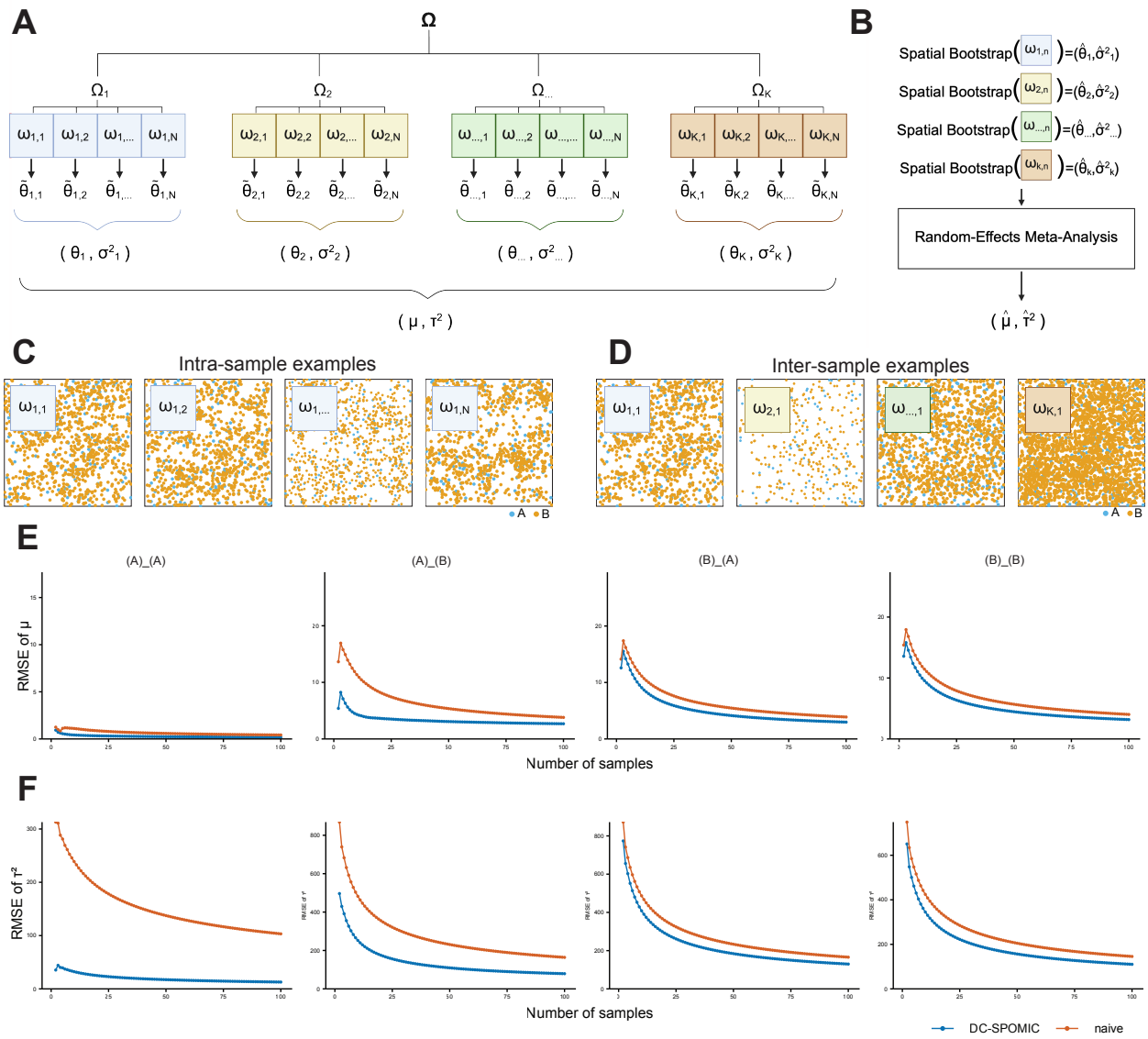

Figure 3. Hierarchical simulation framework for estimates of inter-sample variability. (A) Graphical representation of the simulation and generation of ground truth colocalization mean, intra-, and inter-sample variation. (B) Graphical representation of how DC-SPOMIC estimates colocalization mean, intra- and inter-sample variation in the simulation study using spatial bootstrapping and random-effects meta-analysis. (C) Four realizations from a fixed point process model, indicating the intra-sample variation that can occur from a single spatial generative process. (D) Four realizations from point process models that have random noise added, indicating the inter-sample variation that can occur between samples or patients even when derived from an original underlying process. (E) Ground truth estimates of μ and were compared to DC-SPOMIC estimates of $\hat{\mu}$ and naive sample mean colocalization for sample sizes ranging from 2 to 100 samples. (F) Ground truth estimates of τ^2 were compared to DC-SPOMIC estimates of $\hat{\tau}^2$ and naive sample variance estimation for sample sizes ranging from 2 to 100 samples.

Estimating colocalization differences in colorectal cancer spatial omics data

We applied DC-SPOMIC to a colorectal cancer (CRC) tissue microarray (TMA) imaged with CODEX [20]. The dataset comprises archival formalin-fixed, paraffin-embedded sections from the invasive front of CRC tumors representing two immune phenotypes: Crohn's-like reaction (CLR) and diffuse immune infiltration (DII). In total, 35 cores were analyzed (17 CLR, 18 DII). Cell typing was taken from Schurch et al. and contains 13 different cell types, yielding 169 distinct cell type pairs. Using the cross L-function at a 75 μm radius, DC-SPOMIC identified seven cell-type pairs with significantly higher colocalization in CLR than in DII (Fig. 4A). No pairs showed greater colocalization in DII. To further assess robustness, we repeated the analysis across different radii; the same patterns persisted over a range of spatial scales (Supp. Fig. B.4). Additionally, we compared the results of DC-SPOMIC as opposed to t-tests for a variety of sample sizes, and found that DC-SPOMIC reliably detected strong spatial colocalization behaviors more reliably and with fewer samples (Supplementary Table 1).

We examined the two most significant cell type pairs in greater detail. CLR shows higher colocalization of B cells with other B cells and CD4 T cells with T regulatory cells (Fig. 4B). We use forest plots to summarize, for each pair, the intra-sample uncertainty estimated by spatial bootstrapping and the inter-sample heterogeneity captured by random-effects meta-analysis. Each point corresponds to a single core of the TMA, representing a single sample, with point size proportional to its weight in the pooled estimate; diamonds indicate pooled means with standard errors. The dashed gray line marks the theoretical value of the cross L under complete spatial randomness at 75 μm, providing a visual reference. To place these statistics in a tissue context, we display representative cores from each group whose colocalization values lie closest to their respective pooled group means. By selecting representative cores closes to the group means, we avoid a tendency to bias the data visualization by selecting the most extreme differences. These representative examples of mean colocalization per group illustrate the elevated spatial co-occurrence in CLR relative to DII (Fig. 4C–D).

To probe higher-order spatial organization beyond pairwise effects, we constructed a differential colocalization network from the DC-SPOMIC tests. Since the cross L statistic is asymmetric, the differential test yields two directed Z-scores per unordered pair: $Z_{(i,j)}$ and $Z_{(j,i)}$, where positive values indicate stronger CLR over DII colocalization in the stated direction. For ease of downstream clustering, we converted these into an undirected, weighted graph as follows. First, we form a square matrix M where the rows and columns each represent the distinct cell types in our data. The elements of this matrix are filled by the Z-scores from the DC-SPOMIC differential testing. Second, the matrix is symmetrized $M_{\text{Symmetric}} = \frac{1}{2}(M + M^\top)$. Then, to ensure compatibility with standard community-detection algorithms that assume non-negative weights, we focused on CLR-enriched associations by retaining edges with positive weights, and setting negative weights to zero. Finally, we pruned the graph by keeping edges that met a FDR threshold of $\alpha = 0.1$. Leiden clustering on the resulting symmetric, weighted graph revealed coherent subnetworks of cell types

that tend to colocalize more strongly in CLR than in DII (Supp. Fig. 5A), highlighting potential coordinated spatial arrangements of multiple cell types that are characteristic of the CLR immune phenotype [20].

One of the identified clusters was composed of various immune populations, many of which form the significant pairs identified by DC-SPOMIC. To test the significance of this higher-order subnetwork, for each relevant immune cell, we calculated the distance of the nearest neighbor with all the other cell types in the subnetwork. For example, for every B cell in a sample, we calculated the distance of the nearest T regulatory cell, CD4 T cell, CD8 T cell, and next nearest B cell. The average of these distances was taken across all relevant cells in the sample. We compared the average distances between CLR and DII samples, to test the hypothesis that the average distances in CLR samples is lower than in DII samples. A single-sided t-test demonstrates that the mean of average minimum distance amongst the immune cell types in cluster 1 was statistically significantly lower ($p = 0.03$) in the CLR as opposed to the DII group (Fig. 5B). Two samples from each group closest to the mean are visualized (Fig. 5C(i-ii)).

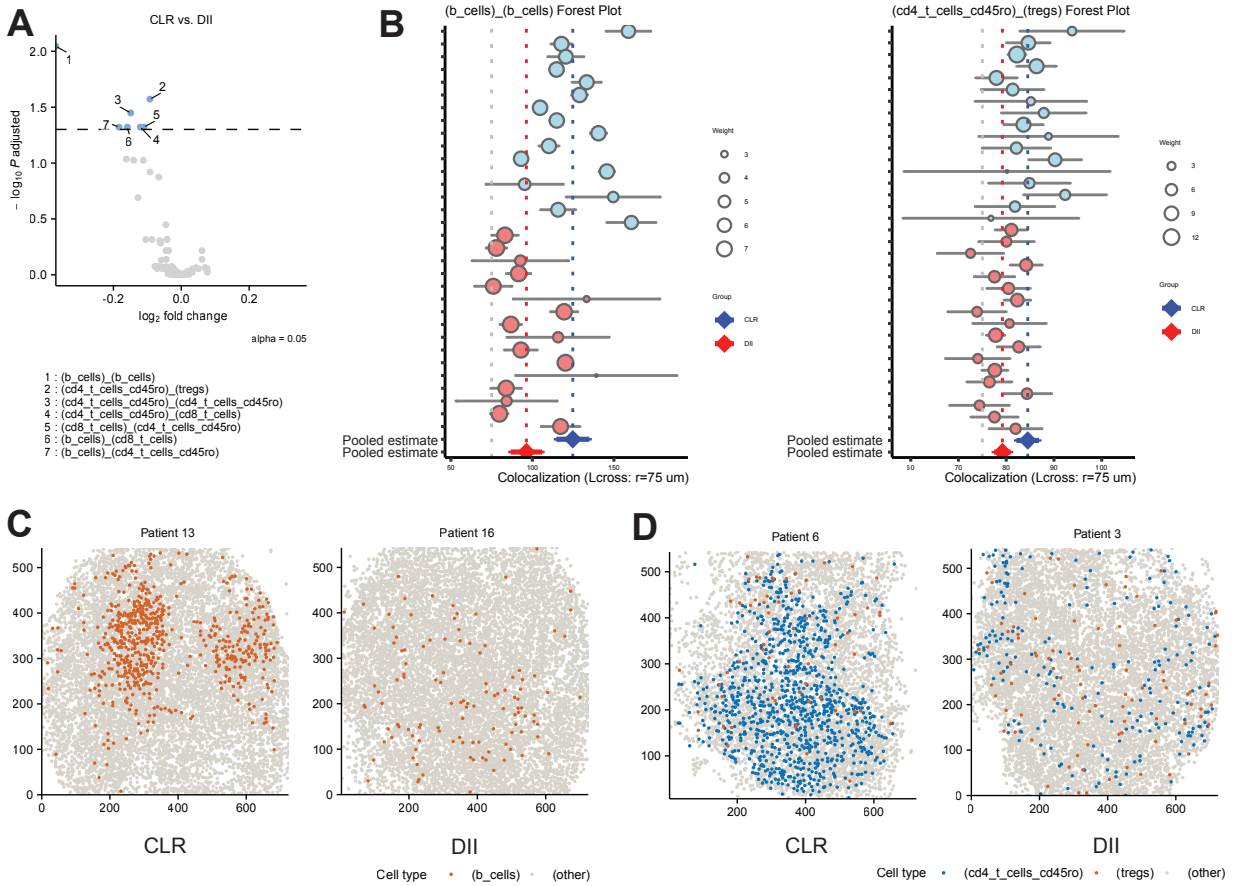

Figure 4. Application of DC-SPOMIC to colorectal cancer dataset from Schurch et al. 2020. (A) Volcano plot of the differential colocalization results from DC-SPOMIC when comparing CLR to DII TMA cores. Seven significant cell pairs (blue) were found to be more strongly colocalized in CLR as opposed to DII TMA cores. (B) Forest plots of the top two significant hits from DC-SPOMIC. The forest plots show the intra-sample variance estimates of colocalization for each cell pair per sample (light blue and pink circles) as well as the pooled colocalization value with standard error (red and blue diamonds). The size of the circles shows the relative impact of the sample on the pooled group estimate. The gray dotted line shows the theoretical value of the L-statistic when considering a 75 μm radius. (C) Visualization of the top significant hit from samples in the CLR and DII groups that are closest to the pooled colocalization means (blue and red dotted lines in (B)). (D) Visualization of the second most significant hit from samples in the CLR and DII groups that are closest to the pooled colocalization means (blue and red dotted lines in (B)).

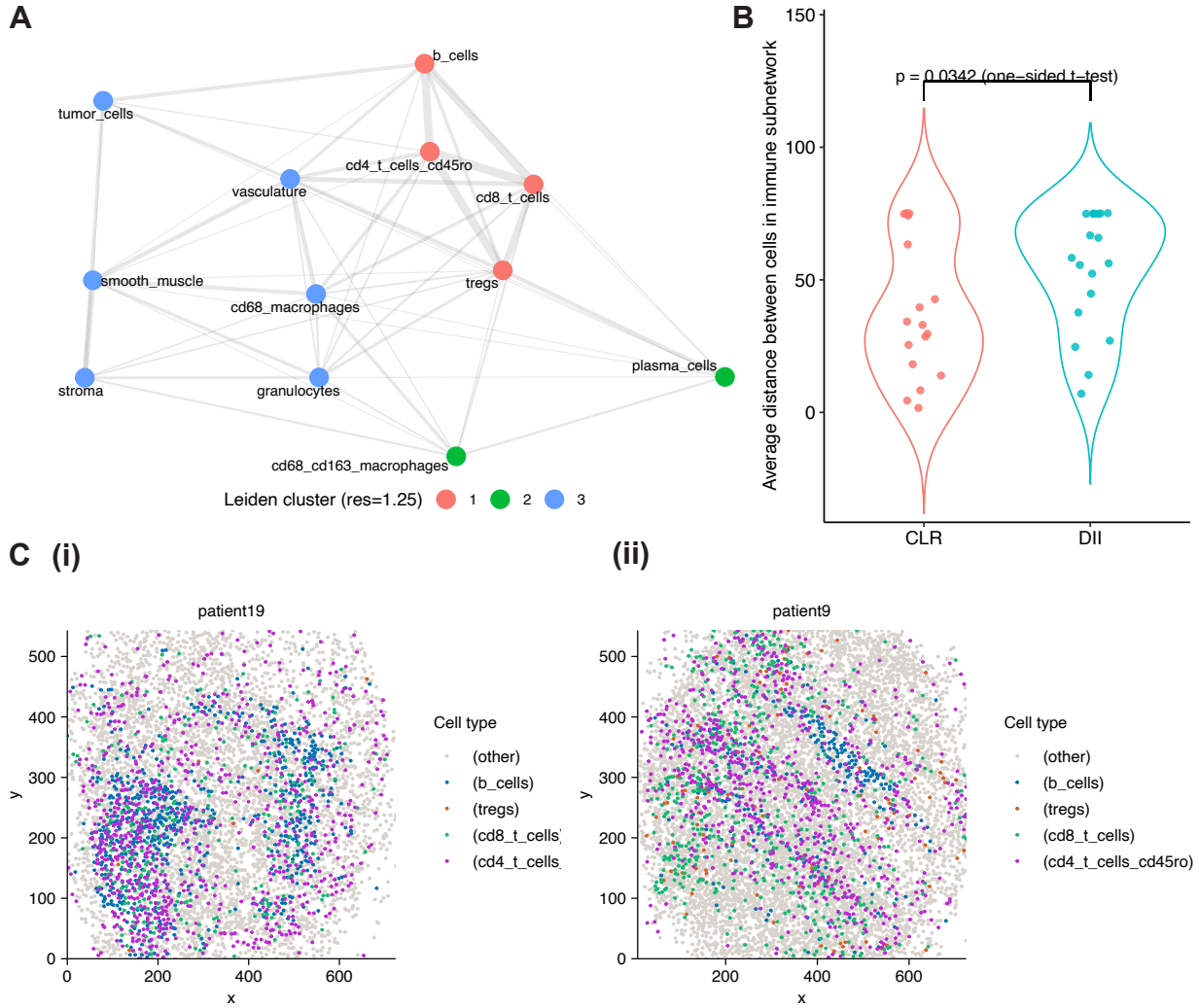

Figure 5. Network analysis of spatial colocalization. **(A)** The Z-scores calculated from differential testing were used to construct a weighted colocalization network between cell types. Each edge between cell-types represents the Z-score from testing CLR vs. DII. The three node colors are determined via Leiden clustering at resolution=1.25. **(B)** To assess the significance of cluster 1, each of the cell types forming cluster 1 in a given sample had the distance to their nearest B cell, CD4 T cell, CD8 T cell, and T regulatory cell calculated. The distances to these nearest cells were averaged to represent the dispersion of cluster 1 cell types around other cluster 1 cell types. These distances were then averaged across the relevant cells in the sample, representing the closeness of cell types belonging to cluster 1 in a given sample. These numbers were compared across CLR and DII via a one-sided t-test. **(C)** The samples closest to the mean average minimum distance in the CLR and DII samples were selected for data visualization.

Discussion

DC-SPOMIC addresses a central statistical gap in spatial omics analysis: the ability to make rigorous group-wise comparisons while accounting for both intra- and inter-sample variability. Existing approaches often treat cell-level measurements in spatial data as independent and overlook the nested structure of variance, which can lead to inflated significance and degraded reproducibility. By combining spatial bootstrapping for intra-sample variance estimation with random-effects meta-analysis for pooling across samples, DC-SPOMIC explicitly models this hierarchy of variability, enabling reliable inference even in studies constrained by small sample sizes and substantial biological heterogeneity. This framework is platform-agnostic and metric-flexible, supporting established measures such as the homogeneous and inhomogeneous cross K and cross L functions, the colocation quotient, and any user-defined colocalization statistic. Such adaptability ensures its utility across diverse spatial technologies and evolving analytical preferences.

Beyond introducing a robust analysis method, this work contributes a reproducible simulation framework for benchmarking spatial statistics under different tissue architectures. Using multi-type spatial point process models, we generated patterns with controllable intra- and inter-sample variability, enabling ground-truth evaluation of both mean and variance colocalization components. Using these simulations enables us to demonstrate that DC-SPOMIC reliably recovers true parameter values in settings where replicate samples are unavailable, as is common in patient-derived datasets. Intra-sample variance estimation was accurate for homogeneous and clustered patterns but diminished for divergent patterns with sharply partitioned spatial domains, suggesting that estimation may be most reliable for contiguous or expert-segmented tissues. Inter-sample variance estimation consistently outperformed naive methods across sample sizes, highlighting the advantage of hierarchical pooling when between-sample variability is nontrivial. More broadly, our simulation design offers the community a foundational benchmark for developing and stress-testing spatial methods under well-controlled conditions.

Applied to multiplexed immunofluorescence TMAs of colorectal cancer, DC-SPOMIC identified seven cell type pairs with significantly greater colocalization in Crohn's-like reaction (CLR) compared to diffuse immune infiltration (DII). These findings were stable across spatial scales and aligned with known enrichment of tertiary lymphoid structures in CLR tumors, many of which were missed by naive testing approaches. By leveraging the uncertainty estimates inherent in the framework, we could visualize both the spread of colocalization values across individual samples and the pooled group estimates, making the magnitude and reliability of differences explicit. Extending beyond pairwise comparisons, we used the CLR–DII differential Z-scores to construct a symmetrized colocalization network, revealing multi-cell-type subnetworks more tightly associated in CLR. This higher-order analysis illustrates how DC-SPOMIC outputs can be integrated into network-based frameworks to explore coordinated multi-cell type spatial organization at the tissue-system level.

In summary, DC-SPOMIC offers a statistically rigorous, flexible, and reproducible approach for

spatial colocalization analysis in single-cell spatial omics. By quantifying uncertainty at multiple hierarchical levels, supporting diverse data modalities and metrics, and enabling both pairwise and higher-order analysis, it addresses multiple limitations of current practice. As spatial omics datasets expand in size, complexity, and clinical relevance, methods like DC-SPOMIC will be essential for ensuring that reported spatial associations are robust, reproducible, and biologically meaningful. Additionally, the flexibility of the framework to work on spatially-indexed data has broad relevance not only to spatial omics, but also to other fields such as the geospatial sciences, where spatial heterogeneity and limited replication also complicate inference. Moreover, the simulation framework introduced here provides a template for benchmarking new spatial analysis methods, fostering a stronger methodological foundation for rigorously evaluating advances in the spatial field.

Data Availability

The colorectal cancer TMA data was previously published by Schurch et al. [20].

Code Availability

The source code for the statistical framework developed in this study is released in the `dcspomic` package https://github.com/plevritis-lab/dcspomic. Additional supporting functions for spatial analysis are available in the `spomic` package https://github.com/plevritis-lab/spomic. All scripts used to conduct the analyses and generate the figures presented in this work are provided in https://github.com/chang-jacob/dcspomic_paper, enabling full reproducibility of the results.

Acknowledgements

We thank Weiruo Zhang and Rohit Khurana for their helpful insights and scientific discussion.

References

1. Dezem FS, Arjumand W, DuBose H, Morosini NS, and Plummer J. Spatially Resolved Single-Cell Omics: Methods, Challenges, and Future Perspectives. Annual Review of Biomedical Data Science 2024 Aug; 7:131–53. DOI: 10.1146/annurev-biodatasci-102523-103640
2. Palla G, Fischer DS, Regev A, and Theis FJ. Spatial components of molecular tissue biology. Nature Biotechnology 2022 Mar; 40:308–18. DOI: 10.1038/s41587-021-01182-1
3. Yuan Z, Zhao F, Lin S, Zhao Y, Yao J, Cui Y, Zhang XY, and Zhao Y. Benchmarking spatial clustering methods with spatially resolved transcriptomics data. Nature Methods 2024 Apr; 21:712–22. DOI: 10.1038/s41592-024-02215-8

4. Zhu J, Wang Y, Chang WY, Malewska A, Napolitano F, Gahan JC, Unni N, Zhao M, Yuan R, Wu F, Yue L, Guo L, Zhao Z, Chen DZ, Hannan R, Zhang S, Xiao G, Mu P, Hanker AB, Strand D, Arteaga CL, Desai N, Wang X, Xie Y, and Wang T. Mapping cellular interactions from spatially resolved transcriptomics data. Nature Methods 2024 Oct; 21:1830–42. DOI: 10.1038/s41592-024-02408-1
5. Yeh CY, Aguirre K, Laveroni O, Kim S, Wang A, Liang B, Zhang X, Han LM, Valbuena R, Bassik MC, Kim YM, Plevritis SK, Snyder MP, Howitt BE, and Jerby L. Mapping spatial organization and genetic cell-state regulators to target immune evasion in ovarian cancer. Nature Immunology 2024 Oct; 25:1943–58. DOI: 10.1038/s41590-024-01943-5
6. Hickey JW, Haist M, Horowitz N, Caraccio C, Tan Y, Rech AJ, Baertsch MA, Rovira-Clavé X, Zhu B, Vazquez G, Barlow G, Agmon E, Goltsev Y, Sunwoo JB, Covert M, and Nolan GP. T cell-mediated curation and restructuring of tumor tissue coordinates an effective immune response. Cell Reports 2023 Dec; 42:113494. DOI: 10.1016/j.celrep.2023.113494
7. Seferbekova Z, Lomakin A, Yates LR, and Gerstung M. Spatial biology of cancer evolution. Nature Reviews Genetics 2023 May; 24:295–313. DOI: 10.1038/s41576-022-00553-x
8. Dos Santos Peixoto R, Miller BF, Brusko MA, Aihara G, Atta L, Anant M, Atkinson MA, Brusko TM, Wasserfall CH, and Fan J. Characterizing cell-type spatial relationships across length scales in spatially resolved omics data. Nature Communications 2025 Jan; 16. DOI: 10.1038/s41467-024-55700-1
9. Canete NP, Iyengar SS, Ormerod JT, Baharlou H, Harman AN, and Patrick E. spicyR: spatial analysis ofin situcytometry data in R. Bioinformatics 2022 May; 38. Ed. by Kelso J:3099–105. DOI: 10.1093/bioinformatics/btac268
10. Emons M, Gunz S, and Robinson M. spatialFDA: A Tool for Spatial Multi-sample Comparisons. 2025. DOI: 10.18129/B9.bioc.spatialFDA
11. Martinelli AL and Rapsomaniki MA. ATHENA: analysis of tumor heterogeneity from spatial omics measurements. Bioinformatics 2022 May; 38. Ed. by Peng H:3151–3. DOI: 10.1093/bioinformatics/btac303
12. Marusyk A, Almendro V, and Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nature Reviews Cancer 2012 May; 12:323–34. DOI: 10.1038/nrc3261
13. Baddeley A, Rubak E, and Turner R. Spatial Point Patterns: Methodology and Applications with R. Boca Raton, FL: Chapman and Hall/CRC Press, 2015
14. Loh JM. A Valid and Fast Spatial Bootstrap for Correlation Functions. The Astrophysical Journal 2008 Jul; 681:726–34. DOI: 10.1086/588631
15. Loh JM. Bootstrapping an inhomogeneous point process. Journal of Statistical Planning and Inference 2010 Mar; 140:734–49. DOI: 10.1016/j.jspi.2009.09.003
16. Röver C, Knapp G, and Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. BMC Medical Research Methodology 2015 Dec; 15:99. DOI: 10.1186/s12874-015-0091-1
17. IntHout J, Ioannidis JP, and Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard

DerSimonian-Laird method. BMC Medical Research Methodology 2014 Dec; 14:25. DOI: 10.1186/1471-2288-14-25

18. Sidik K and Jonkman JN. Simple Heterogeneity Variance Estimation for Meta-Analysis. Journal of the Royal Statistical Society Series C: Applied Statistics 2005 Apr; 54:367–84. DOI: 10.1111/j.1467-9876.2005.00489.x
19. Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software 2010; 36:1–48. DOI: 10.18637/jss.v036.i03
20. Schürch CM, Bhate SS, Barlow GL, Phillips DJ, Noti L, Zlobec I, Chu P, Black S, Demeter J, McIlwain DR, Kinoshita S, Samusik N, Goltsev Y, and Nolan GP. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. Cell 2020 Sep; 182:1341–1359.e19. DOI: 10.1016/j.cell.2020.07.005

Appendices

A Supplementary Methods

Colocalization statistics

K-function and L-function

For an homogeneous Poisson process, the value of the K-function under complete spatial randomness is

$$K(r) = \pi r^2 \tag{13}$$

The L-function performs a variance-stabilizing transformation of the K-function

$$L(r) = \sqrt{\frac{K(r)}{\pi}} \tag{14}$$

From these equations, we can observe that for the L-function, the value under CSR is

$$L(r) = \sqrt{\frac{\pi r^2}{\pi}} \tag{15}$$

$$= r \tag{16}$$

Thus, when interpreting the value of the L-function, $L(r) > r$ indicates higher colocalization of cell types than expected under random organization, while a value $L(r) < r$ indicates that cell types are less colocalized than expected by chance [1, 2].

Local cross-K and cross-L functions

The local versions of the cross-K and cross-L functions calculate the colocalization of an individual cell in the data with other cell types. For example, if there are 100 cells in a dataset composed of 5 different cell types, then the local colocalization values would form a matrix $\Theta^*_{100\times 5}$. A useful property of the local K and local L functions is that the average of local statistics is equivalent to the global K and L function values [1].

$$K^*_{(z,j)}(z,r) = \frac{|W|}{N_j - 1} \sum_{N_j} \left[1\{d(x,y) \leq r\} \cdot e(z,j)\right] \tag{17}$$

The local cross K-function, denoted $K^*_{(z,j)}$ is calculated for a cell z. For z, the local colcoalization with cell type j is computed by calculating the distances of all type j cells, N_j and counting the number of instances when cell type j within a radius r of z. There is an edge correction factor to account for the shape of the spatial domain. The count is then normalized by the area of the spatial region $|W|$, the total population size of cell type j.

The local L-cross function, which is used by default by DC-SPOMIC is a transformed version of the local cross-K function.

$$L^*_{(i,j)}(x,r) = \sqrt{\frac{K^*_{(i,j)}(x,r)}{\pi}} \tag{18}$$

Edge correction

The K andd L-functions count cells of a certain type within various radii. These counts are affected near the edges of the spatial region, where neighborhoods of increasing radii may extend beyond the bounds of the observed spatial point pattern, potentially leading to biased underestimation of interaction or clustering.

To account for this, various edge correction methods are available in the `spatstat` R, each based on different assumptions about the unobserved region beyond the window. By default, DC-SPOMIC uses translation edge correction, which assumes that the spatial point process is stationary. This method corrects for edge effects by adjusting each point pair's contribution based on whether the displacement vector between them could have been fully observed elsewhere in the window. For example, if two points are 10 units apart horizontally, this displacement could only occur when the first point is at least 10 units from the right edge. Translation correction accounts for this reduced visibility by upweighting such pairs, ensuring that edge effects do not bias the estimation of spatial statistics.

Loh spatial bootstrap

To quantify uncertainty in spatial summary statistics without relying on parametric model assumptions, we used the Loh spatial bootstrap method, as implemented in the `spatstat` package. The Loh bootstrap is a nonparametric bootstrap method for spatial point processes. It resamples the local spatial statistics of each cell to compute confidence intervals for spatial summary functions such as the K and L-functions. These confidence intervals are converted to estimates of intra-sample variance, which are used downstream to fit random-effects meta-analyses.

Random-effects meta-analysis

Sidik-Jonkman estimator for τ^2

Many random-effects estimators exist to provide an estimator $\hat{\tau}^2$. Restricted maximum likelihood (REML) and the DerSimonian-Laird (DL) methods are very commonly used, but have been

demonstrated to commonly result in inappropriate type I error rates (IntHout et al., 2014). The Sidik-Jonkman (SJ) estimator has been demonstrated to produce more adequate error rates over a variety of ample sizes and degrees of between-sample heterogeneity, thus DC-SPOMIC uses SJ estiamtion for $\hat{\tau}^2$ by default. Depending on user-knowledge of the samples, DC-SPOMIC also allows any $\hat{\tau}^2$ estimator incorporated into the `metafor` package.

Intra-sample variance simulation

Three point processes were specified to represent different spatial patterns. These point processes were specified over a 2000×2000 window. The parameters specified below were manually tuned to create spatial patterns and cell type densities similar to those observed in the colorectal cancer dataset.

Homogeneous point process

To create the homogeneous point process, a constant intensity function was specified per cell type and then the intensities were superimposed to form a marked point process.

- $\lambda_{(\mathrm{A})}(x, y) = 2 \times 10^{-4}$
- $\lambda_{(\mathrm{B})}(x, y) = 1.5 \times 10^{-4}$
- $\lambda_{(\mathrm{C})}(x, y) = 1 \times 10^{-4}$
- $\lambda_{(\mathrm{Rare\ 1})}(x, y) = 1 \times 10^{-5}$
- $\lambda_{(\mathrm{Rare\ 2})}(x, y) = 5 \times 10^{-6}$

Divergent point process

To create the divergent point process, two inhomogeneous intensity functions were specified for cell types A and B, while constant intensity functions were specified for the remaining cell types. The intensities were superimposed to form a makerd point process.

- $\eta \sim \mathrm{Uniform}(1 \times 10^{-10}, 5 \times 10^{-3})$
- $\lambda_{(\mathrm{A})}(x, y) = \eta e^{-\eta x}$
- $\lambda_{(\mathrm{B})}(x, y) = \eta e^{-\eta(2000-x)}$
- $\lambda_{(\mathrm{C})}(x, y) = 1 \times 10^{-4}$
- $\lambda_{(\mathrm{Rare\ 1})}(x, y) = 1 \times 10^{-5}$
- $\lambda_{(\mathrm{Rare\ 2})}(x, y) = 5 \times 10^{-6}$

Cluster point process

To create the cluster process, we first sample parent points A from a constant intensity function $\lambda_{(A)}(x, y)$.

- $\lambda_{(\mathrm{A})}(x, y) = 1 \times 10^{-5}$

We then define A_i as the coordinate-location of each type A point from $i = 1, 2, \ldots, N_A$, where N_A is the total count of type A points. For each A_i, we generate offspring points of type B which are located nearby. To do this, we randomly sample the number of offspring $N_{A_i,B}$ associated with each parent A_i from a Poisson distribution with pseudocount to ensure at least one offspring is generated.

- $N_{B_i} \sim \mathrm{Poisson}(7) + 1$

Thus, each parent A_i has $N_{A_i,B}$ offspring points. The locations of each of these offspring are sampled from a normal distribution which is centered at the location of the parent point A_i and has some dispersion σ^2.

- $B_j \sim \mathcal{N}(A_i, \sigma^2)$, where $j = 1, 2, \ldots, N_{A_i,B}$ and $\sigma^2 = 175$.

Finally, we sample the remaining cell types exactly as before in the homogeneous and divergent point processes.

- $\lambda_{(\mathrm{C})}(x, y) = 1 \times 10^{-4}$
- $\lambda_{(\mathrm{Rare\ 1})}(x, y) = 1 \times 10^{-5}$
- $\lambda_{(\mathrm{Rare\ 2})}(x, y) = 5 \times 10^{-6}$

Inter-sample variance simulation

The clustered point processes used in this simulation were specified over a 2000×2000 window. A baseline process is specified with the same parameters as in Section A. This baseline process $\mathbf{\Omega}$ is considered to be the underlying generative spatial process for the population of point patterns that we will simulate.

To emulate variability that can exist between patients or samples with common underlying biology, the baseline process was distorted by noise introduced by random variables that cause deviations from the underlying generative model to produce new processes for each sample Ω_k, for samples $k = 1, 2, \ldots, K$.

We first sample the parent points A. However, unlike our earlier simulation, this parameter will have some slight distortion for each Ω_k in $k = 1, 2, \ldots, K$.

- $\lambda_{(A)}(x, y) = 1 \times 10^{-5} + \mathrm{Uniform}(0, 0.0001)$

Then offspring points are generated as before, but now, the number of offspring points N_{B_i} are sampled with some variation for each Ω_k.

- $N_{B_i} \sim \text{Poisson}(\gamma) + 1$, where $\gamma \sim \text{NegativeBinomial}(10, 2) + 3$.

The dispersion of offspring points around the parent points is also variable for each Ω_k, where now

- $B_j \sim \mathcal{N}(A_i, \sigma^2 + \epsilon)$, where $\epsilon \sim \mathcal{N}(0, 100)$

The layers of random sampling create multiple point processes which are all derived from an original spatial process $\mathbf{\Omega}$. Thus, each Ω_k parameterizes SPPs with cluster processes with variation in cell type numbers, densities, and dispersions, reflecting heterogeneity that can exist between samples.

B Supplementary Tables and Figures

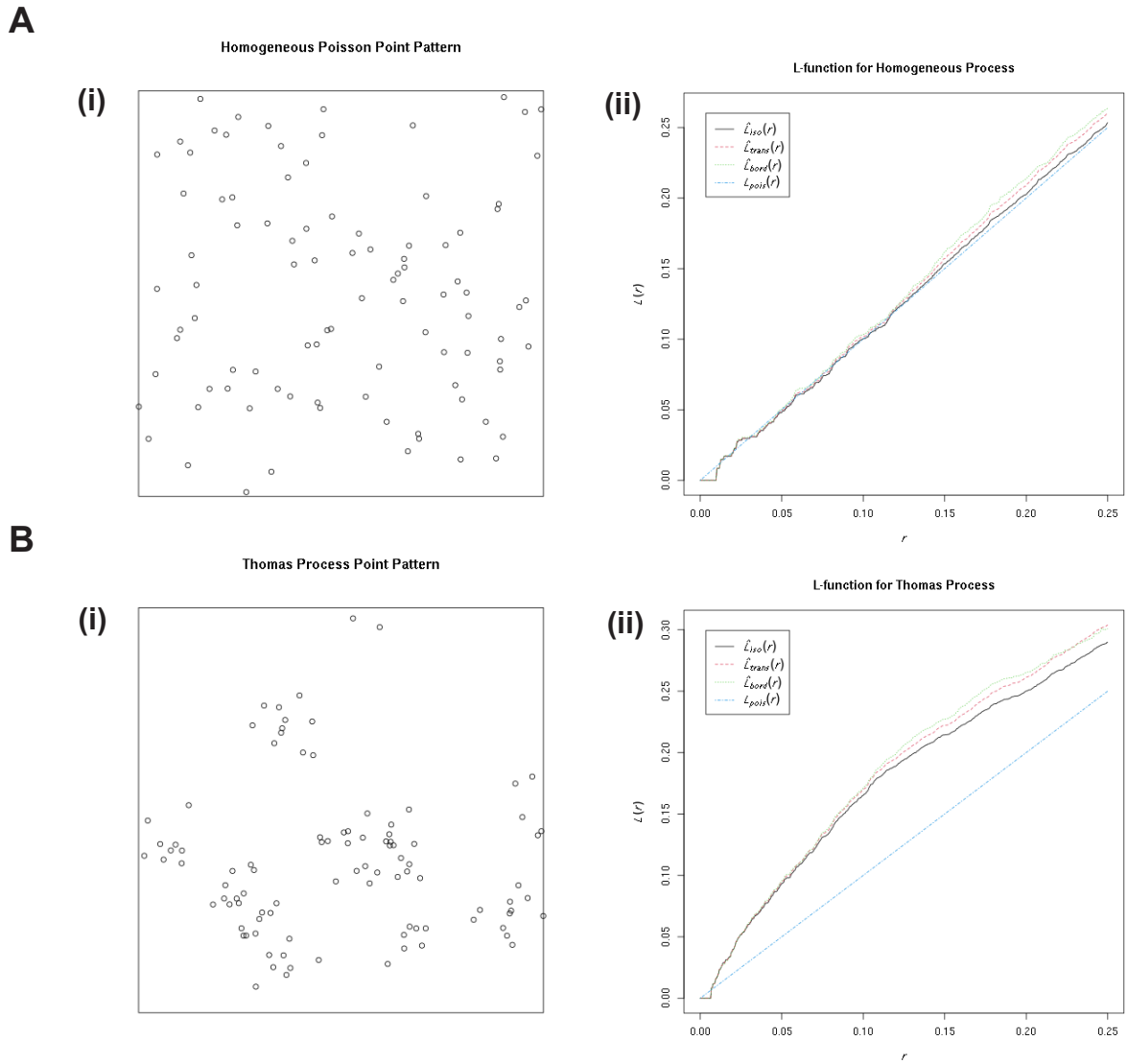

Figure B.1. Example of L-function on two point patterns over the unit square. **(A)**(i) Example of a point pattern from a homogeneous poisson point process. (ii) The L-function with various edge correction methods plotted over a length of radii (r). L_{pois} (blue) refers to the expected number of points under complete spatial randomness. L-function values demonstrate that points are randomly organized. **(B)**(i) Example of a clustered point pattern from a Thomas process. (ii) The L-function with various edge correction methods plotted over a length of radii (r). L_{pois} (blue) refers to the expected number of points under complete spatial randomness. L-function values indicate higher colocalization of points than expected under complete spatial randomness.

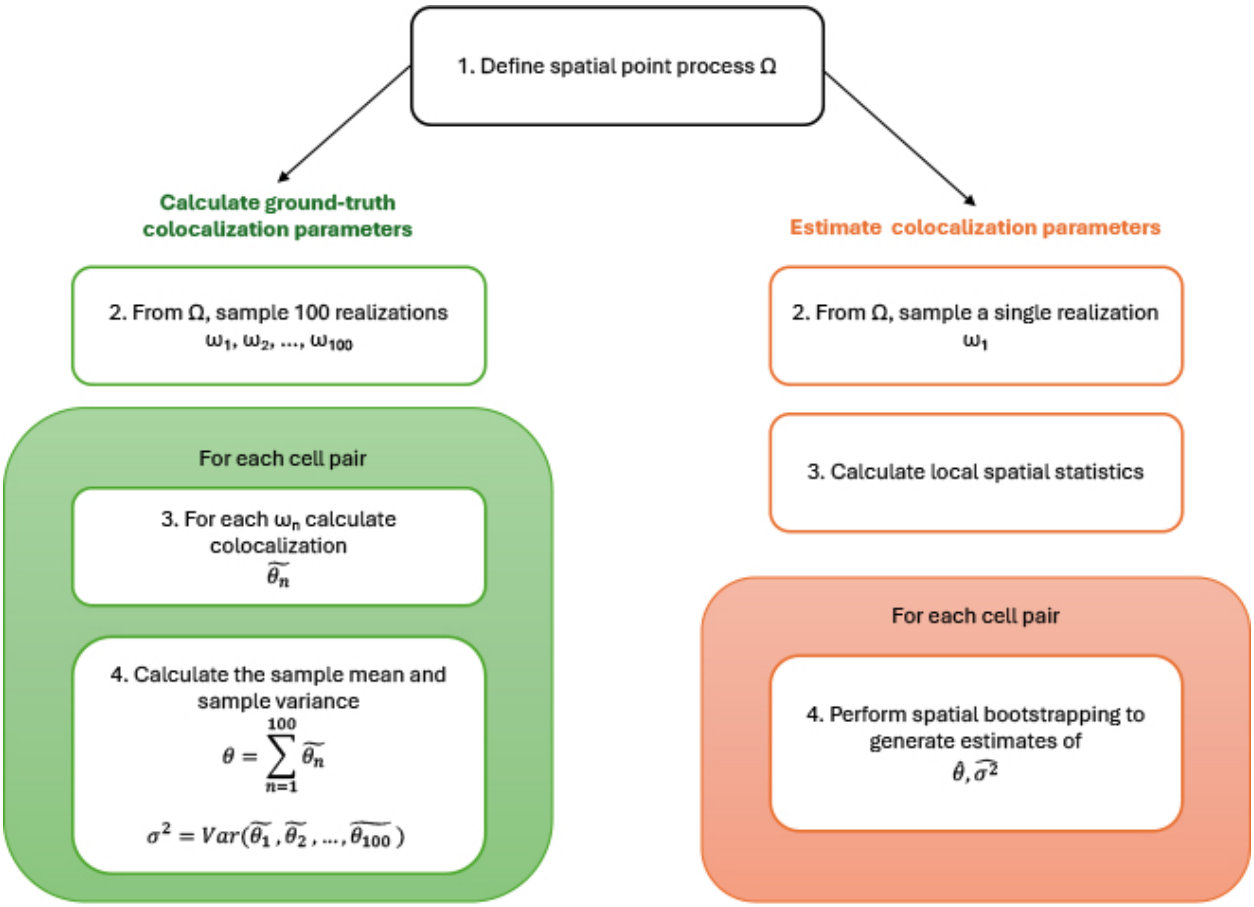

Figure B.2. Schematic for the calculation of ground truth intra-sample variance. First a spatial point process Ω is defined. This stochastic process produces realizations in the form of point patterns. For ground-truth parameter calculation, follow the green side of the diagram. We sample 1000 of these point patterns, and calculate colocalization statistics for each point pattern. The distribution of these values give us empirical estiamtes of the variance which at high sample size we assume to converge to ground-truth parameter values. To estimate colocalization parameters in the absence of unlimited sample size, we sample a single point pattern from Ω, and then calculate the local spatial statistics of each cell in the pattern with the other cell types. These local spatial statistics are spatially bootstrapped hundred of times, and the resulting mean and variance of the bootstrapped distribution produce estiamtes of intrasample variance.

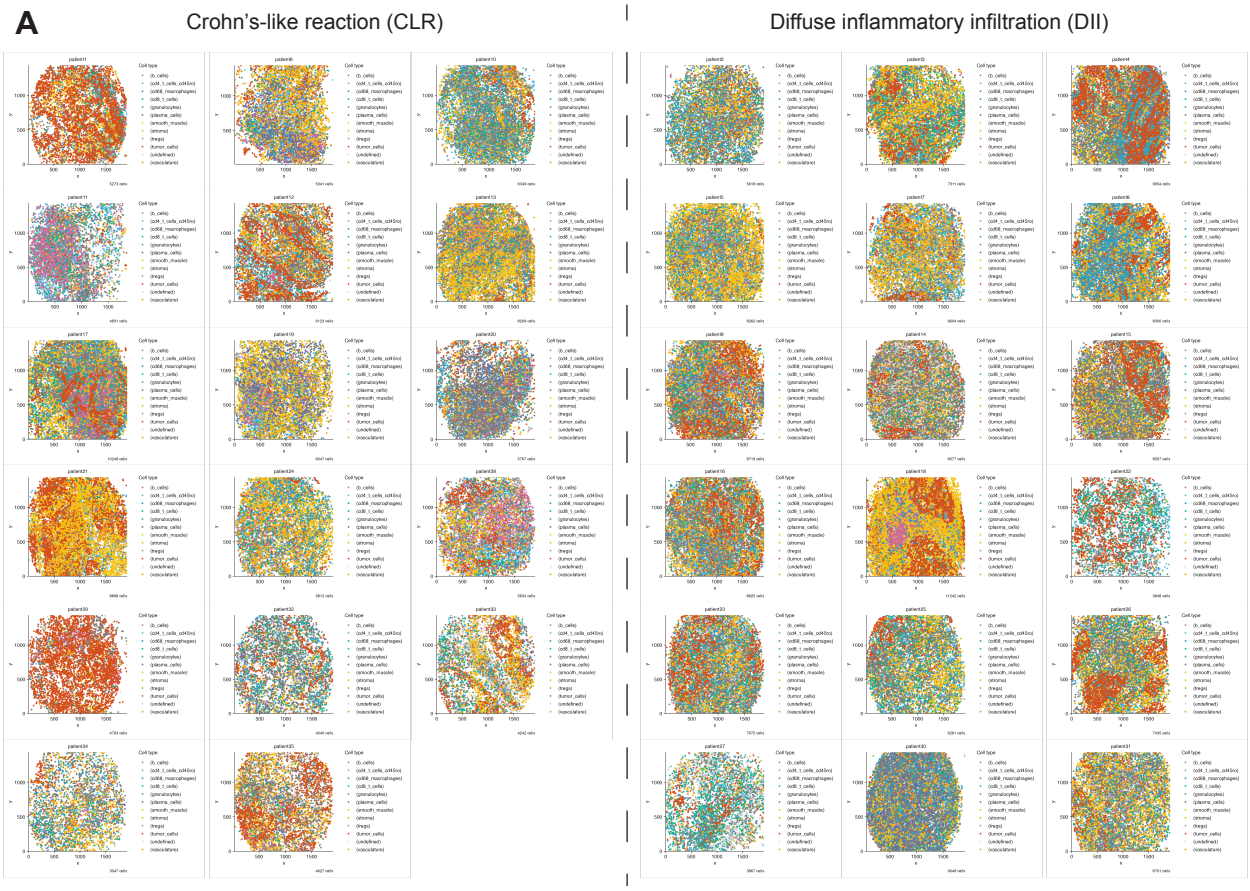

Figure B.3. Visualization of all 35 TMA cores separated into Crohn's-like reaction ($n = 17$) and diffuse inflammatory infiltration ($n = 18$).

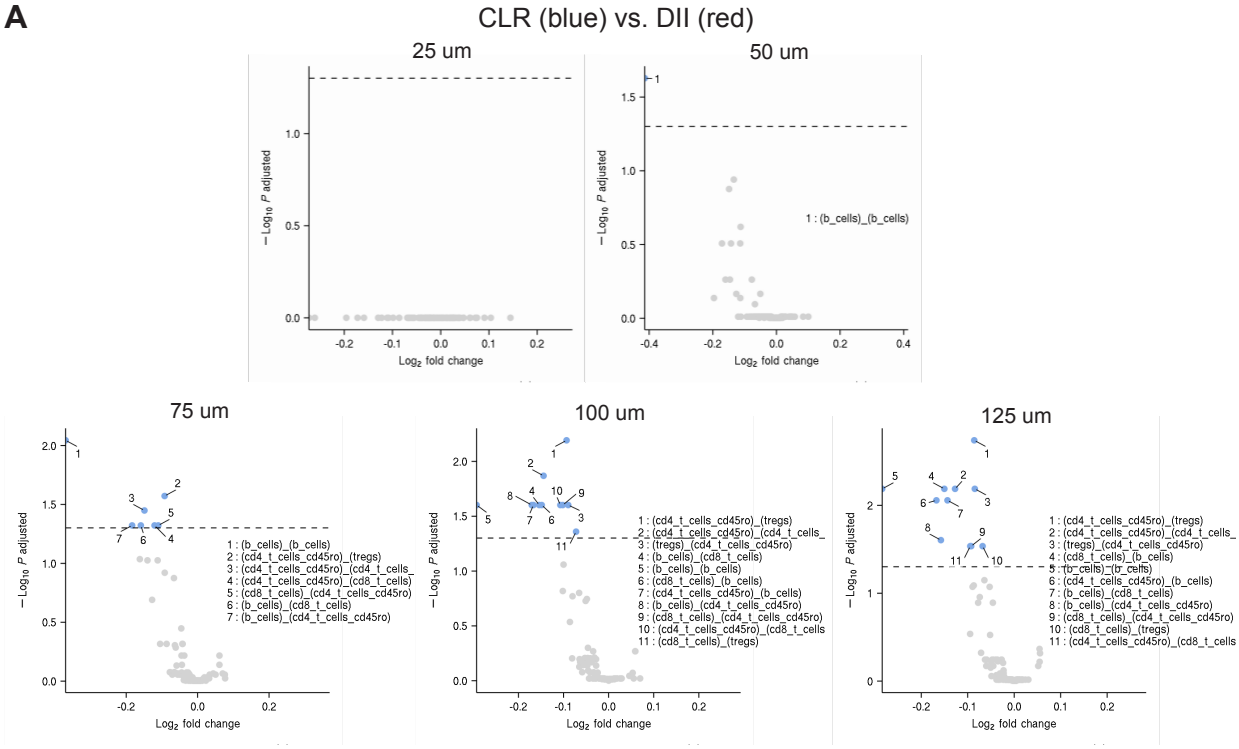

Figure B.4. (A) Volcano plots comparing CLR and DII tumor samples using DC-SPOMIC at multiple length-scales. Colocalization is calculated using the cross-L function with translational edge correction. Black horizontal dotted line denotes the threshold for statistical significance ($alpha = 0.05$ with FDR-adjusted p-values. Dots in blue are more colocalized in CLR tumors than in DII. No pairs of cell types are found to be more colocalized in DII tumors.

CLR sample size	DII sample size	**DC-SPOMIC significant pairs**	**t-test significant pairs**
6	6	NA	NA
7	7	NA	NA
8	8	NA	NA
9	9	NA	NA
10	10	NA	NA
11	11	NA	NA
12	12	NA	NA
13	13	NA	NA
14	14	(b_cells)_(b_cells)	NA
15	15	(b_cells)_(b_cells)	NA
16	16	(b_cells)_(b_cells)	
17	17	(b_cells)_(b_cells), (cd4_t_cells_cd45ro)_(tregs), (cd4_t_cells_cd45ro)_(cd4_t_cells_cd45ro), (cd4_t_cells_cd45ro)_(cd8_t_cells), (cd8_t_cells)_(cd4_t_cells_cd45ro), (b_cells)_(cd4_t_cells_cd45ro), (cd8_t_cells)_(cd8_t_cells), (b_cells)_(cd8_t_cells), (cd4_t_cells_cd45ro)_(b_cells)	(tregs)_(cd4_t_cells_cd45ro), (cd4_t_cells_cd45ro)_(tregs), (b_cells)_(b_cells)
17	18	(b_cells)_(b_cells), (cd4_t_cells_cd45ro)_(tregs), (cd4_t_cells_cd45ro)_(cd4_t_cells_cd45ro), (cd4_t_cells_cd45ro)_(cd8_t_cells), (b_cells)_(cd8_t_cells), (b_cells)_(cd4_t_cells_cd45ro), (cd8_t_cells)_(cd4_t_cells$_{cd45ro)}$	(tregs)_(cd4_t_cells_cd45ro), (cd4_t_cells_cd45ro)_(tregs)

Table 1. Comparison of DC-SPOMIC results with t-tests across varying sample sizes. Starting at 4 samples (2 CLR samples and 2 DII samples), DC-SPOMIC is run at a distance of 75μm alongside t-tests of the colocalization estimates between groups. The colocalization pairs forund to be differentially colocalized after FDR-correction are reported in the table.

References

1. Baddeley A, Rubak E, and Turner R. Spatial Point Patterns: Methodology and Applications with R. Boca Raton, FL: Chapman and Hall/CRC Press, 2015
2. Moraga P. Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny. Biostatistics Series. Chapman & Hall/CRC, 2019