Variational sparse Gaussian-process method for detecting spatially variable genes and cellular interactions in spatial transcriptomics

Zhicong Wang, Jing Li, Liqing Xie, Yiran Wang, Yongtian Wang, Jing Chen, Xuequn Shang, Xingyi Li, Zhaowen Liu, Jialu Hu

Briefings in Bioinformatics, 2026, 27, bbag371
https://doi.org/10.1093/bib/bbag371
Published: 14 July 2026
Problem Solving Protocol

Variational sparse Gaussian-process method for detecting spatially variable genes and cellular interactions in spatial transcriptomics

Zhicong Wang[1,2,3,‡], Jing Li[1,‡], Liqing Xie[3,4,‡], Yiran Wang[1], Yongtian Wang[1], Jing Chen[5], Xuequn Shang[1], Xingyi Li[1], Zhaowen Liu[1], Jialu Hu[1,3,*]

[1]School of Computer Science, Northwestern Polytechnical University, 1 Dongxiang Rd., Chang'an Dist., Xi'an, Shaanxi, 710129, China
[2]Department of Radiology, The First Affiliated Hospital of Xiamen University, School of Medicine, Xiamen University, 55 Zhenhai Rd., Xiamen, 361003, China
[3]German Center for Neurodegenerative Diseases (DZNE), Venusberg-Campus 1, Bonn 53127, Germany
[4]Department of Child Healthcare, Children's Hospital of Fudan University (Xiamen Branch), 92 Yibin Rd., Xiamen, 361006, China
[5]School of Computer Science and Engineering, Xi'an University of Technology, 5 Jinhua South Road, Xi'an 710048, China

*Corresponding author. E-mail: jhu@nwpu.edu.cn
‡Zhicong Wang, Jing Li and Liqing Xie have contributed equally to this work and share first authorship.

Abstract

Advanced spatially resolved transcriptomic (SRT) technologies preserve the spatial context of gene expression within tissues, enabling the study of context-dependent transcriptional regulation. Here, we propose a variational inference-assisted sparse Gaussian-process (VISGP) framework for identifying spatially variable genes (SVGs) and inferring spatially dependent cellular interactions from SRT data. VISGP combines sparse Gaussian-process approximations with variational inference via inducing variables to reduce computational and memory costs while enabling gene-specific adaptation of spatial covariance structures. Across simulated data and four real SRT datasets, VISGP detected more SVGs than existing methods and identified 85 spatially constrained ligand–receptor pairs that were missed by alternative approaches. Together, VISGP provides a scalable and statistically grounded strategy for decoding spatial gene regulation and cell–cell communication, yielding biological insights into cellular heterogeneity and cancer pathology.

Keywords **spatial transcriptomics, spatially variable genes, sparse Gaussian process, variational inference, ligand–receptor interactions, cellular interactions**

Introduction

Spatially resolved transcriptomics (SRT) enables direct measurement of gene expression within intact tissue architecture, transforming analyses of cellular heterogeneity. By preserving spatial context, these technologies have yielded fundamental insights into tissue function [1, 2], normal development [3], and disease pathology [4]. Over the past decade, diverse SRT technologies have emerged, differing in resolution, sensitivity, and transcriptome coverage [5–7]. Imaging-based techniques, such as *in situ* sequencing and multiplexed *in situ* hybridization methods (e.g. MERFISH and seqFISH), provide single-molecule resolution but are limited to predefined target panels. In contrast, sequencing-based methods encode spatial information onto RNA molecules and recover transcriptome-wide expression through next-generation sequencing [8]. The growing scale of these datasets requires computational frameworks to resolve spatial heterogeneity and uncover tissue architecture. A key analytical goal in SRT is to identify spatially variable genes (SVGs), which differ from highly variable genes in scRNA-seq. SVGs show region-specific expression patterns that mirror tissue organization and cellular heterogeneity, providing a quantitative link between transcriptional variation and anatomical structure.

Several statistical modeling methods have been proposed to identify significant SVGs based on various model-based inference. The most pioneering methods include trendsceek [9], SpatialDE [10], SPARK [11], and SPARK-X [12]. Trendsceek models SRT data as marked point processes, and tests for significant dependency between the spatial distributions of points and their associated expression levels. SpatialDE models the data based on Gaussian process (GPR) regression and decomposes expression variability into spatial and nonspatial components using two random effect terms. SPARK directly models spatial count data through generalized linear spatial models, and employs a penalized quasi-likelihood algorithm to approximately

Received: January 31, 2026. **Revised:** June 2, 2026. **Accepted:** June 9, 2026

estimate parameters. SPARK-X builds upon a robust covariance test framework and extends it to incorporating a variety of spatial kernels for nonparametric spatial modeling of sparse count data. GPcounts [13] models spatial expression using GPs with a count-based likelihood, such as a negative-binomial observation model, to account for count-valued and overdispersed measurements. A more recent method, nnSVG [14], employs a nearest-neighbor Gaussian process (NNGP) framework to estimate gene-specific spatial covariance parameters. However, many of these model-based approaches rely on predefined kernel choices, restricted covariance structures, or separate preprocessing/modeling steps, which may limit their flexibility in capturing heterogeneous spatial expression patterns across genes. Apart from statistical modeling, several deep learning-based methods have been developed for SRT data, such as SOMDE [15] and STAMarker [16]. These techniques usually construct spatial dependency structures first, treating graph generation, constraint design, or normalization as a separate module. Any inaccuracies introduced during this initial step tend to be inherited by downstream analyses, potentially leading to less optimal outcomes. These limitations constrain the robustness and generalizability of current approaches in detecting SVG expression.

Here, we present variational inference-assisted sparse Gaussian process (VISGP), a new statistical modeling framework for identifying SVGs. Building on recent GP-based SVG methods, VISGP emphasizes explicit score-based testing and enables spatially constrained ligand–receptor inference. VISGP models each gene's expression as a GP integrating sparse approximation with variational inference [17–19] to achieve both flexibility and computational efficiency. The sparse formulation introduces a set of low-dimensional inducing variables that compress spatial features and reduce computational complexity, while variational inference jointly optimizes the kernel hyperparameters with model parameters, enabling gene-specific adaptation of spatial covariance structures. Furthermore, it employs a score-based statistical test under a null-hypothesis framework [20] to rigorously identify SVGs. We benchmarked VISGP against commonly used and recently developed SVG detection methods on one simulated and three real SRT datasets. It accurately and efficiently detected SVGs, demonstrating strong robustness and scalability. Compared with existing methods, it exhibited superior sensitivity in capturing subtle spatial patterns, particularly in datasets with weak spatial signals. Beyond identifying SVGs, it can also be readily extended to infer spatially dependent ligand–receptor interactions, enabling characterization of spatially organized cell–cell communication. These results highlight the potential of VISGP as a powerful and generalizable tool for the analysis of SRT data.

Materials and methods

Gaussian process model

We model spatial transcriptomics data comprising expression measurements for p genes across n spatial spots. For a given gene, let $\mathbf{y} = (y_1, \ldots, y_n)^\mathsf{T}$ denote the gene expression vector measured at spatial locations $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^\mathsf{T}$, where each coordinate is typically 2D, $\mathbf{x}_i = (x_{i1}, x_{i2})$. We model the expression profile of each gene independently using a GP.

A GP defines a distribution over functions $f(\mathbf{x})$:

$$f(\mathbf{x}) \sim \mathcal{GP}\big(m(\mathbf{x}),\, k(\mathbf{x}, \mathbf{x}')\big), \tag{1}$$

where $m(\mathbf{x})$ is the mean function and $k(\mathbf{x}, \mathbf{x}')$ is the covariance function. We set $m(\mathbf{x}) = 0$ throughout, consistent with gene-wise z-score standardization across spots (zero mean, unit variance) prior to model fitting. In this study, we use a radial basis function (RBF) kernel

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\ell^2}\right), \tag{2}$$

parameterized by the length scale ℓ.

For kernel sensitivity analysis, we additionally evaluated two variants with different kernels: one with a Matérn kernel using a smoothness parameter $\nu = 1.5$, and the other with an additive kernel combining RBF and Matérn covariance functions.

Given observations $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ for a gene, we assume a Gaussian likelihood with additive noise:

$$y_i = f(\mathbf{x}_i) + \epsilon_i, \qquad \epsilon_i \sim \mathcal{N}(0, \sigma_e^2). \tag{3}$$

Let $\mathbf{f} = (f_1, \ldots, f_n)^\mathsf{T}$ denote the latent function values with $f_i = f(\mathbf{x}_i)$. Under the GP prior,

$$p(\mathbf{f} \mid \mathbf{X}) = \mathcal{N}(\mathbf{f} \mid \mathbf{0}, \mathbf{K}_{nn}), \tag{4}$$

where $\mathbf{K}_{nn}$ is the covariance matrix with entries $[\mathbf{K}_{nn}]_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. The joint model factorizes as $p(\mathbf{y}, \mathbf{f}) = p(\mathbf{y} \mid \mathbf{f})\, p(\mathbf{f})$, where

$$p(\mathbf{y} \mid \mathbf{f}) = \mathcal{N}(\mathbf{y} \mid \mathbf{f}, \sigma_e^2 \mathbf{I}). \tag{5}$$

Marginalizing $\mathbf{f}$ yields

$$p(\mathbf{y} \mid \ell, \sigma_e^2) = \mathcal{N}\left(\mathbf{y} \mid \mathbf{0},\, \mathbf{K}_{nn} + \sigma_e^2 \mathbf{I}\right). \tag{6}$$

We further introduce a variance scaling parameter τ to capture the magnitude of spatial effects, resulting in

$$p(\mathbf{y} \mid \boldsymbol{\theta}) = \mathcal{N}\left(\mathbf{y} \mid \mathbf{0},\, \tau \mathbf{K}_{nn} + \sigma_e^2 \mathbf{I}\right), \tag{7}$$

where $\boldsymbol{\theta} = (\ell, \sigma_e^2, \tau)$. In the hypothesis testing section, we show that τ can be interpreted as the variance component governing spatial effects.

In principle, $\boldsymbol{\theta}$ can be estimated by maximizing the log marginal likelihood $\log p(\mathbf{y} \mid \boldsymbol{\theta})$ using gradient-based optimization. However, exact inference requires repeated inversion (or Cholesky factorization) of $\tau \mathbf{K}_{nn} + \sigma_e^2 \mathbf{I}$, leading to computational complexity $\mathcal{O}(n^3)$ and memory cost $\mathcal{O}(n^2)$ for n spots. This limits scalability to medium or large datasets. To address this issue, we adopt a sparse variational GP approximation using inducing variables [17, 21] (described below).

Sparse Gaussian process and variational approximation

Exact GP inference in Equation (7) scales cubically with the number of spatial spots n. To improve scalability, we adopt a sparse variational GP approximation based on inducing variables [17, 18, 21]. Let $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_m)^\mathsf{T}$ denote m inducing locations in the same spatial domain as $\mathbf{X}$, with $m \ll n$. Define the inducing variables $\mathbf{u} = (u_1, \ldots, u_m)^\mathsf{T}$

where $u_j = f(\mathbf{z}_j)$. We place a joint GP prior over the latent function values $\mathbf{f}$ and inducing variables $\mathbf{u}$:

$$\begin{bmatrix}\mathbf{f}\\ \mathbf{u}\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}\mathbf{K}_{nn} & \mathbf{K}_{nm}\\ \mathbf{K}_{mn} & \mathbf{K}_{mm}\end{bmatrix}\right), \tag{8}$$

where $\mathbf{K}_{mm}$ is the covariance matrix evaluated at inducing points, and $\mathbf{K}_{nm} = \mathbf{K}_{mn}^{\mathsf{T}}$ is the cross-covariance between training inputs and inducing points.

From the properties of multivariate Gaussians, the conditional prior $p(\mathbf{f} \mid \mathbf{u})$ is Gaussian:

$$p(\mathbf{f} \mid \mathbf{u}) = \mathcal{N}(\mathbf{f} \mid \mathbf{A}\mathbf{u}, \mathbf{B}), \tag{9}$$

with

$$\mathbf{A} = \mathbf{K}_{nm}\mathbf{K}_{mm}^{-1}, \qquad \mathbf{B} = \mathbf{K}_{nn} - \mathbf{K}_{nm}\mathbf{K}_{mm}^{-1}\mathbf{K}_{mn}. \tag{10}$$

Because the exact posterior $p(\mathbf{f}, \mathbf{u} \mid \mathbf{y})$ remains intractable at scale (it involves $\mathbf{K}_{nn}$ through $p(\mathbf{y})$), we introduce a variational distribution

$$q(\mathbf{f}, \mathbf{u}) = p(\mathbf{f} \mid \mathbf{u})\, q(\mathbf{u}), \qquad q(\mathbf{u}) = \mathcal{N}(\mathbf{u} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{11}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are variational parameters and the inducing locations $\mathbf{Z}$ are optimized jointly with model hyperparameters.

Under this construction, the marginal variational distribution of $\mathbf{f}$ is also Gaussian:

$$q(\mathbf{f}) = \int p(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})\, d\mathbf{u} = \mathcal{N}\left(\mathbf{f} \mid \mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathsf{T}} + \mathbf{B}\right). \tag{12}$$

We maximize the evidence lower bound (ELBO),

$$\mathcal{L} = \mathbb{E}_{q(\mathbf{f})}\left[\log p(\mathbf{y} \mid \mathbf{f})\right] - \mathrm{KL}\left(q(\mathbf{u}) \,\|\, p(\mathbf{u})\right), \tag{13}$$

which satisfies $\log p(\mathbf{y}) \geq \mathcal{L}$. Here $p(\mathbf{u}) = \mathcal{N}(\mathbf{0}, \mathbf{K}_{mm})$ is the inducing prior. For the Gaussian likelihood $p(\mathbf{y} \mid \mathbf{f}) = \mathcal{N}(\mathbf{y} \mid \mathbf{f}, \sigma_e^2\mathbf{I})$, the expected log-likelihood term has a closed form:

$$\mathbb{E}_{q(\mathbf{f})}\left[\log p(\mathbf{y} \mid \mathbf{f})\right] = -\frac{n}{2}\log(2\pi\sigma_e^2) - \frac{1}{2\sigma_e^2}\left(\|\mathbf{y} - \mathbf{A}\boldsymbol{\mu}\|^2 + \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathsf{T}} + \mathbf{B})\right). \tag{14}$$

The KL term between Gaussians also admits a closed form:

$$\mathrm{KL}(q(\mathbf{u}) \,\|\, p(\mathbf{u})) = \frac{1}{2}\left(\log\frac{|\mathbf{K}_{mm}|}{|\boldsymbol{\Sigma}|} - m + \mathrm{tr}\left(\mathbf{K}_{mm}^{-1}\boldsymbol{\Sigma}\right) + \boldsymbol{\mu}^{\mathsf{T}}\mathbf{K}_{mm}^{-1}\boldsymbol{\mu}\right). \tag{15}$$

We jointly optimize model hyperparameters (including ℓ, τ, and σ_e^2 in Equation (7)), inducing locations $\mathbf{Z}$, and variational parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ by maximizing $\mathcal{L}$ using gradient-based optimization. This sparse variational formulation reduces the computational complexity from $\mathcal{O}(n^3)$ to $\mathcal{O}(nm^2)$ and the memory footprint to $\mathcal{O}(nm)$ for $m \ll n$. The number of inducing variables m controls the trade-off between approximation fidelity and computational efficiency in the sparse variational GP. Increasing m improves expressiveness but increases cost, with dominant time complexity $\mathcal{O}(nm^2)$ and memory approximately $\mathcal{O}(nm)$. In practice, we recommend $m \ll n$ and using m as a tuning knob: start with $m = 20$ and increase for larger or more heterogeneous tissues until results stabilize under available resources.

We maximize the ELBO using the Adam optimizer. The variational mean is initialized to zero and the variational covariance is initialized to the identity matrix; inducing locations are initialized by sampling from a Gaussian distribution in the spatial coordinate space and are optimized jointly with kernel hyperparameters during training.

Hypothesis testing

We assess SVGs through a hypothesis testing framework, where the null hypothesis posits the absence of spatial structure in gene expression. In this context, the marginal distribution of gene expression values $\mathbf{y}$ across spatial locations is defined as:

$$\mathbf{y} \sim \mathcal{N}\left(\mathbf{0}, \tau\mathbf{K}_{nn} + \sigma_e^2\mathbf{I}\right), \tag{16}$$

where τ denotes the variance component attributable to spatial dependence. Thus, the presence of a spatial expression pattern can be formally evaluated by testing the null hypothesis:

$$H_0 : \tau = 0, \tag{17}$$

which corresponds to the case where no spatial variation is explained by the covariance structure.

To perform this test, we compute a *P*-value using the Davies method [22] based on the score statistic:

$$Q = \mathbf{y}^T\mathbf{K}_{nn}\mathbf{y}. \tag{18}$$

Based on Equations 1618, under $H_0 : \tau = 0$ the statistic Q is a Gaussian quadratic form; by eigendecomposition of $\mathbf{K}_{nn}$, Q/σ_e^2 follows a weighted sum of independent χ_1^2 variables with weights given by the eigenvalues of $\mathbf{K}_{nn}$ (see Supplementary Text). Genes are subsequently identified as SVGs if their q-value (FDR-adjusted *P*-value [23]) is <0.05. We control FDR using the Benjamini–Hochberg(BH) procedure; although spatial data can exhibit dependence, BH is widely used and remains valid under certain positive-dependence conditions.

Datasets and preprocessing

We obtained mouse olfactory bulb (MOB) and human breast cancer data from the SRT Research database (https://www.spatialresearch.org/). These datasets consist of gene expression measurements derived from read counts across discrete spatial locations, referred to as spots. Our analysis included the MOB replicate 9 dataset, comprising measurements of 15 284 genes across 237 spots, and the BC layer 2 dataset, comprising 14 789 genes across 251 spots. Genes with fewer than 10 total counts were filtered out, and only spots with a minimum of 10 total read counts were retained for subsequent analysis. We analyzed 11 428 genes across 237 spots in the MOB dataset and 9907 genes across 251 spots in the human breast cancer dataset. Human squamous cell carcinoma (SCC) and pancreatic ductal adenocarcinoma (PDAC) data were obtained from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo). For human SCC, we selected Patient 5, replicate 2, which initially comprised 17 690 genes measured in 1934 spots. Based on criteria described in the original study, a subset of 521 spots was retained for downstream analysis. In total, 25 754 genes across 428 spots were analyzed in

PDAC. For control of data quality, we filtered out genes and spots with <10 total counts and performed gene-wise z-score standardization across spots using sklearn.preprocessing.scale. This step was used to construct continuous, comparable-scale inputs for the GP likelihood and to improve numerical stability during GP hyperparameter optimization; it is not intended to replace library-size normalization as a general SRT preprocessing procedure.

Compared methods

We compared VISGP and its variants to several state-of-the-art algorithms for identifying genes with spatial expression patterns, including SpatialDE (v.1.1.3), SPARK (v.1.1.1), SPARK-X (v.1.1.1), GPcounts (v.0.1), and nnSVG (v.1.10.3).

SpatialDE employs a multivariate normal model to assess spatial variability in gene expression using likelihood-ratio tests. In this framework, a Gaussian kernel is selected using the Bayesian information criterion, and its hyperparameter is optimized from a predefined set of values. SPARK adopts a spatial generalized linear mixed model that incorporates multiple spatial kernels, enabling direct modeling of count-based spatial transcriptomic data. SPARK-X further extends this framework by using a nonparametric covariance-testing approach that supports multiple spatial kernels and improves scalability for sparse SRT datasets. GPcounts uses a Gaussian-process model with a negative-binomial observation model, thereby directly accounting for count-valued and overdispersed expression measurements. nnSVG uses an NNGP framework to estimate gene-specific spatial covariance parameters for scalable SVG detection.

For SpatialDE, SPARK, SPARK-X, GPcounts, and nnSVG, we applied consistent filtering criteria to genes and spots before method-specific preprocessing. GPcounts was applied to filtered raw count matrices, whereas nnSVG was applied to normalized log-counts. Statistical power was reflected by counting the number of significant genes (FDR < 0.05 using BH). The experimental benchmark datasets included MOB, human breast cancer, and human SCC datasets. Not all methods or parameter settings were applied to every dataset because some analyses were designed as sensitivity assessments and some methods were computationally prohibitive on larger datasets; dataset-specific settings are summarized in Supplementary Tables S1 andS2 and figure legends.

Evaluation metrics

True positive rate and true negative rate

The true positive rate (TPR) is used to measure the method's sensitivity:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{19}$$

A method with a higher sensitivity has a lower type II error rate. The true negative rate (TNR) could measure the specificity of the method:

$$\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{20}$$

A method with a higher specificity has a lower type I error rate.

Moran's I statistic

Moran's I statistic is a widely used index for assessing the degree of spatial autocorrelation. The Moran's I statistic value is distributed in $[-1, 1]$. When the value is closer to 1, it means that the attribute values on the spatial points are more aggregated into a specific pattern; when the value is more comparable to -1, it means that the attribute values on the spatial points are more scattered; and when the value is equal to 0, it means that the attribute values of the spatial points are completely randomly distributed. Moran's I was used only as an external evaluation metric and was computed in R (`spdep v.1.4.1`) from exported VISGP outputs.

Gene set enrichment analysis

We performed GO term enrichment analysis using the R package clusterProfiler (v.4.14.6). In the package, we used the default "BH" method for multiple testing corrections and screened GO terms with FDR <0.05 as significantly enriched. The results showed the top 10 most enriched GO terms for the three categories of GO terms (cellular component, molecular function, and biological process), respectively.

Results

Overview of variational inference-assisted sparse Gaussian-process

As illustrated in Fig. 1a–d, VISGP integrates gene expression profiles and spatial coordinates to identify SVGs from SRT data. For each gene, we model expression as a GP with a squared exponential (RBF) kernel parameterized by length scale ℓ and signal variance τ, and assume Gaussian noise with variance σ_ϵ^2. All model and variational parameters are jointly optimized by maximizing the ELBO under a sparse variational approximation with inducing variables [17–19]. The inducing locations $\mathbf{Z}$ and variational parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ are learned jointly with the model parameters, enabling scalable inference with time complexity $\mathcal{O}(nm^2)$ and memory complexity $\mathcal{O}(nm)$ for n spots and m inducing points. We used a fixed $m = 20$ across datasets for comparability in benchmarking. We then perform a score-based test [20] under the null hypothesis $H_0 : \tau = 0$ to call SVGs, where the test statistic $Q = \mathbf{y}^{\mathrm{T}}\mathbf{K}_{nn}\mathbf{y}$ follows a mixture of χ^2 distributions (Fig. 1c). Genes passing the significance threshold are visualized together with their spatial expression patterns (Fig. 1d). In practice, we use a Gaussian likelihood on gene-wise z-score standardized expression values across spots, which yields centered values on comparable scales and it is well suited to GP-based spatial modeling. An implementation is available in our Python package `scbean` (version $\geq$ 0.6.0; see Data availability).

Benchmarking on simulated data

To benchmark VISGP against established SVG detection methods in the simulation setting, we applied each algorithm to a simulated dataset designed to evaluate how accurately methods distinguish SVGs under controlled spatial patterns. Following Trendsceek [9], we generated spatial pseudo-count data for the simulation. Spatial coordinates for 1000 spots were sampled from a Poisson point process, and gene expression counts were drawn from a negative binomial distribution. Non-SVGs were randomly distributed without spatial architecture, whereas SVGs exhibited one of thethree spatial patterns—hotspot, streak, or gradient (Fig. 1e). Each simulated dataset contained 400 non-SVGs and 100 SVGs spanning these spatial configurations. For a fair comparison, we evaluated all methods using sensitivity (TPR), specificity (TNR), and false discovery rate (FDR).

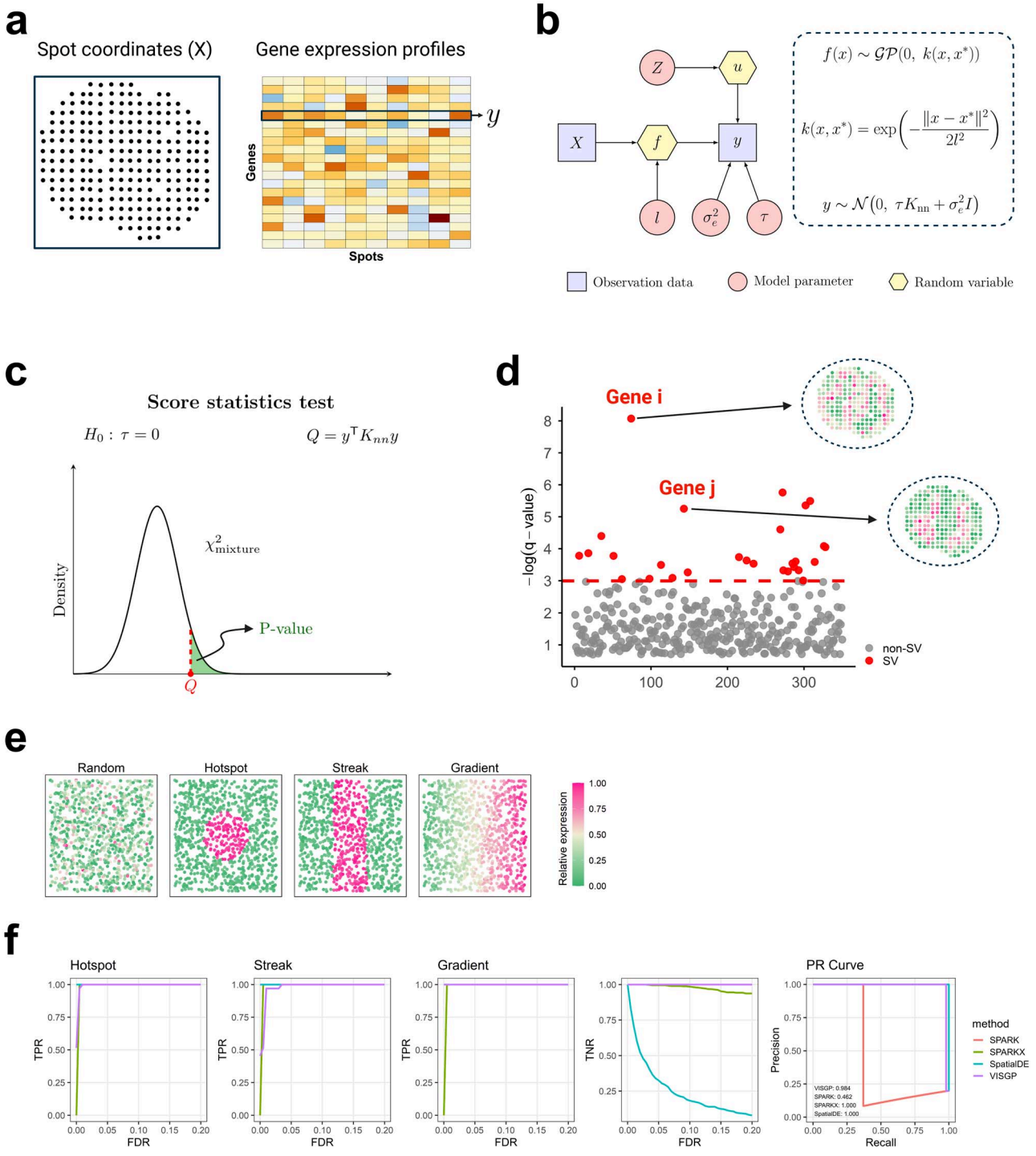

Figure 1 Method overview of VISGP and simulation results. (a) Model inputs. (b) Variational Inference assisted Sparse GP. $\mathbf{K}_{nn}$ is a covariance matrix whose elements are given by the kernel function $k(x, x^*)$ for each pair of spot coordinates. (c) Score statistical significance test. (d) SV genes are screened by q-value (adjusted P-value). The dotted line represents the q-value cutoff of 0.05. (e) Representative genes in the set of simulation show a random pattern and other three spatial expression patterns (hotspot, streak, and gradient). (f) Power plots illustrate the TPR and TNR achieved by different methods across a range of FDR thresholds in the simulations. Precision–recall (PR) curves are also shown, with the corresponding area under the PR curve (AUPRC) reported for each method.

Across patterns, all methods included achieved high sensitivity at conservative thresholds (FDR < 10%). In terms of specificity, SpatialDE showed a marked decline as the FDR threshold increased, likely reflecting limitations of its spatial kernel in modeling certain expression patterns. SPARK-X exhibited only a mild decrease; at FDR = 0.05, the reduction in specificity was minimal (Fig. 1f). Overall, VISGP and SPARK most effectively distinguished SVGs from non-SVGs in this simulation, whereas SpatialDE tended to misclassify a subset of non-SVGs as SVGs. Consistent with the PR analysis, VISGP achieved a near-perfect AUPRC (0.984), indicating high precision across most recall levels. In contrast, SPARK showed markedly lower precision overall (AUPRC = 0.462), while SPARKX and SpatialDE yielded AUPRC values close to 1.0 under this setting. These results motivate applying VISGP to real SRT datasets to identify biologically meaningful spatial structure across diverse spatial patterns.

Variational inference-assisted sparse Gaussian-process yields robust spatially variable genes detection in mouse olfactory bulb data

We next applied VISGP to a real SRT dataset of the MOB [24], comprising 15 284 genes measured across 237 spots (see Datasets and preprocessing for details). Three state-of-the-art methods and two variants of VISGP (employing distinct kernels) were included in the benchmark analysis, all operating under their respective recommended configurations. To obtain a null expectation, we kept the MOB expression matrix unchanged and randomly permuted the spatial coordinates 10 times. Each permuted coordinate file was analyzed using the methods included in the null-calibration analysis. Under this permuted null, VISGP, SPARK, and SPARK-X produced well-calibrated *P*-values, whereas SpatialDE showed pronounced *P*-value deflation (Supplementary Fig. 1a). This permutation-based calibration serves as a sensitivity check for the testing procedure by assessing type-I error control when spatial structure is removed. We then evaluated performance on the real (unpermuted) spatial data. VISGP showed high sensitivity across FDR thresholds ranging from 1% to 10% (Fig. 2a). At FDR = 5%, VISGP (RBF) identified 1607 SVGs, compared with 1640 from VISGP (Matérn), 1625 from VISGP (Multi), 1207 from GPcounts, 462 from nnSVG, 867 from SPARK, 884 from SPARK-X, and 291 from SpatialDE. The Matérn and multi-kernel variants produced detection curves comparable with VISGP (RBF), indicating that VISGP-based SVG detection was robust to kernel choice in this dataset. GPcounts, which uses a negative-binomial count likelihood, also recovered a substantial number of SVGs, supporting the consistency between VISGP and count-based GP modeling, however, this came with substantially longer runtime compared with VISGP (Supplementary Table S2). Among the four methods included in the gene-overlap analysis, VISGP (RBF), SpatialDE, SPARK-X, and SPARK uniquely reported 679, 5, 137, and 402 method-specific SVGs, respectively (Fig. 2b).

We subsequently grouped the 1607 SVGs identified by VISGP into three principal categories (Fig. 2c and d), each representing a distinct spatial expression pattern (Patterns I–III). In the MOB, Pattern I marked genes enriched in the glomerular layer, Pattern II in the mitral cell layer, and Pattern III in the granule cell layer, mirroring the hierarchical neuronal organization of the tissue. These three patterns were further supported by representative marker genes, including *Penk* ($p = 1 \times 10^{-16}$), *Doc2g* ($p = 5.3 \times 10^{-3}$), and *Kctd12* ($p = 1 \times 10^{-16}$) (Fig. 2e). Figure 2e also illustrates the benefit of learning a gene-specific length scale ℓ: *Penk* shows a sharply localized layer-restricted pattern, whereas *Doc2g* and *Kctd12* vary over broader spatial domains. Estimating ℓ per gene helps capture both localized and broad patterns within the same tissue, whereas a fixed/shared scale may oversmooth local signals or underfit broad domains.

To assess cross-method consistency, method-specific SVGs uniquely identified by VISGP, SPARK, and SPARK-X were also clustered into three categories. The spatial patterns defined by VISGP- and SPARK-specific genes remained consistent with the three anatomical layers of the MOB (Supplementary Fig. 2a and b), whereas those from SPARK-X were less well defined (Supplementary Fig. 2c). Corresponding heatmaps of VISGP-specific SVGs revealed clear spatial expression structure (Supplementary Fig. 2d). We further quantified spatial coherence of method-specific SVGs using Moran's *I*, a standard statistic of spatial autocorrelation ranging from −1 (complete dispersion) to 1 (perfect clustering). VISGP-specific SVGs exhibited higher Moran's *I* values than those identified by SPARK and SPARK-X, indicating stronger spatial organization (Supplementary Fig. 2e).

To characterize biological relevance, we performed gene set enrichment analysis (GSEA) on the 1607 SVGs identified by VISGP. In total, 756 Gene Ontology (GO) terms were significantly enriched (BH adjusted $P \leq .05$). As shown in Fig. 2f and Supplementary Fig. 3, the most prominent categories were related to neuronal and synaptic processes, including "neuron-to-neuron synapse" (GO:0098984; adjusted $p = 3.18 \times 10^{-23}$) and "asymmetric synapse" (GO:0032279; adjusted $p = 5.92 \times 10^{-22}$). Additional terms such as "modulation of chemical synaptic transmission" (GO:0050804; adjusted $p = 6.75 \times 10^{-20}$) and "regulation of synapse organization" (GO:0050803; adjusted $p = 3.90 \times 10^{-14}$) further highlighted enrichment of neuronal communication pathways. Together, these results indicate that VISGP effectively captures spatial gene expression patterns associated with neuronal circuitry, consistent with the known layered organization of the olfactory bulb. In addition to the enrichment analysis, we report the empirical runtimes of VISGP and its competitors, all evaluated under the same SLURM environment (Supplementary Table S2).

Variational inference-assisted sparse Gaussian-process reveals biologically meaningful structure in human breast cancer microenvironment

To evaluate VISGP in tumor tissues, we applied it to an SRT dataset of human breast cancer biopsies [24], comprising 14 789 genes measured across 251 spots. The same dataset was analyzed using benchmark methods for comparison. Permutation analysis indicated that VISGP, SPARK, and SPARK-X maintained proper *P*-value calibration under the null, whereas SpatialDE exhibited systematic *P*-value inflation, deviating upward from the theoretical null (Supplementary Fig. 1b). We next evaluated performance on the real spatial data using SpatialDE, SPARK, SPARK-X, GPcounts, nnSVG, and three VISGP kernel settings. VISGP showed high sensitivity across FDR thresholds ranging from 1%to 10% (Fig. 3a). At FDR = 5%, VISGP (RBF) identified 3302 SVGs, compared with 3519 from VISGP (Matérn), 3417 from VISGP (Multi), 3476 from GPcounts, 632 from nnSVG, 2032 from SPARK-X, 469 from SPARK, and 201 from SpatialDE. The Matérn and multi-kernel VISGP variants showed similar or higher detection power relative to the RBF kernel, supporting the robustness of VISGP across kernel specifications. GPcounts produced a comparable detection profile, indicating that VISGP results were broadly consistent with a negative-binomial count-likelihood GP model. Among the four methods included

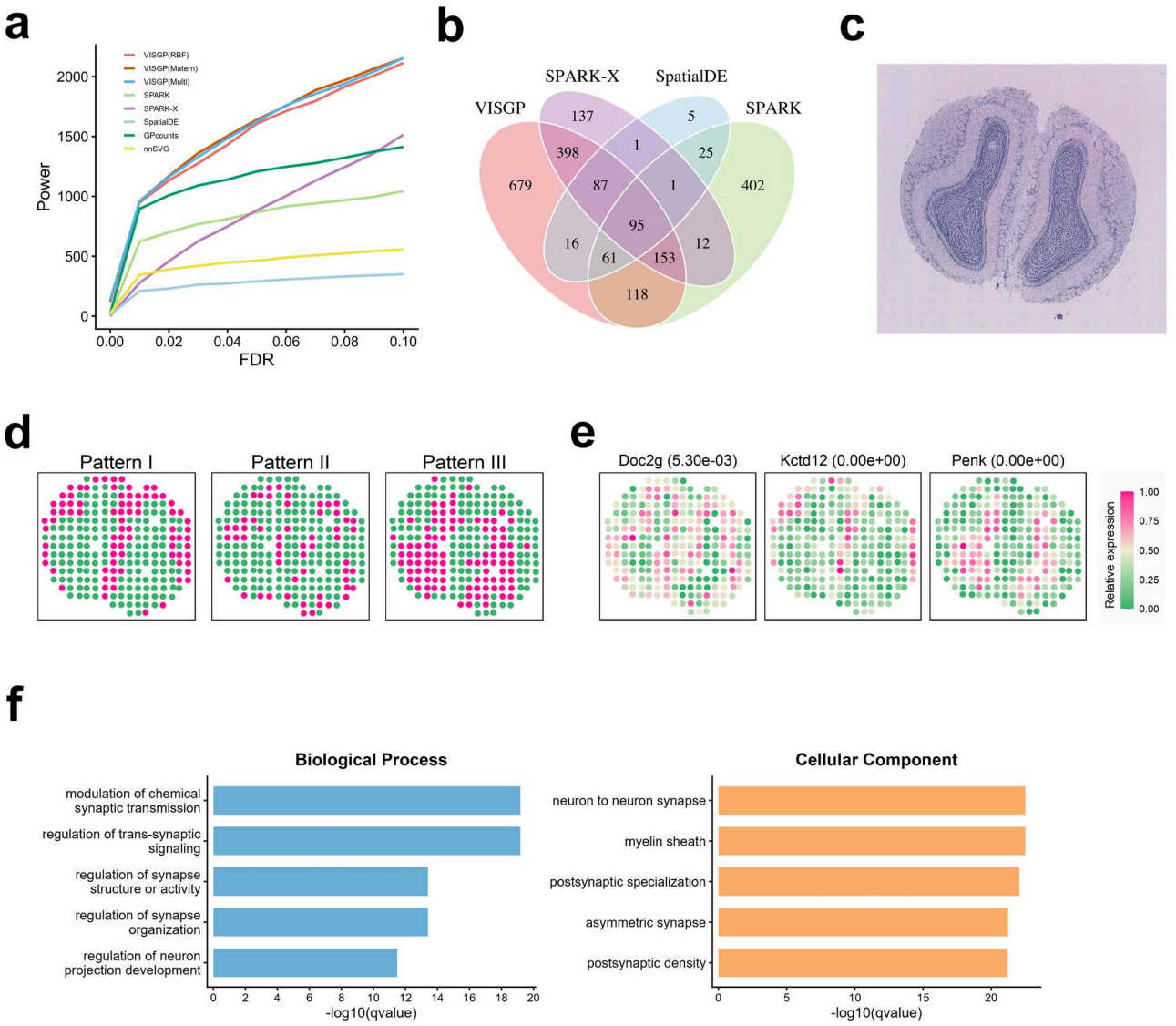

Figure 2 Analyzing the MOB dataset. (a) Power plot shows the number of significant SV genes detected by each method from the MOB dataset. The benchmark includes VISGP with RBF, Matérn, and multi-kernel (RBF + Matérn) settings, together with SpatialDE, SPARK, SPARK-X, GPcounts, and nnSVG. (b) Venn diagram shows the overlap between SV genes identified by VISGP (RBF), SpatialDE, SPARK-X, and SPARK based on an FDR cutoff of 0.05. (c) Image of hematoxylin and eosin-stained MOB from [24]. (d) According to the SV genes identified by VISGP, three different spatial expression patterns are summarized. (e) Spatial expression pattern for some marker genes for different layers in MOB along with their FDR (adjusted P-value) from VISGP (in parentheses). (f) Bar plot for GO enrichment analysis on 1607 SV genes obtained by VISGP based on an FDR cutoff of 0.05.

in the gene-overlap analysis, VISGP (RBF), SPARK-X, SPARK, and SpatialDE uniquely detected 1552, 306, 28, and 39 method-specific SVGs, respectively (Fig. 3b). We then clustered the 3302 SVGs identified by VISGP into two dominant spatial expression patterns (Fig. 3d). Pattern I corresponded to the cancer tissue region visually depicted as the dark area in Fig. 3c, whereas Pattern II mapped to adjacent nontumorous tissue. This spatial contrast highlights VISGP's ability to delineate biologically distinct regions within tumor sections. To assess biological validity, we benchmarked VISGP and competing methods using a curated list of 18 previously reported tumor-associated genes [24–26]. As shown in Fig. 3e and Supplementary Fig. 4, using the methods included in the marker-gene comparison, VISGP identified all 18 genes with well-defined spatial expression patterns, whereas SpatialDE, SPARK, and SPARK-X detected 6, 10, and 12 of them, respectively. To further evaluate the relevance of VISGP-specific discoveries, we analyzed the 1552 SVGs uniquely identified by VISGP. Clustering revealed two distinct spatial expression patterns, both exhibiting clear spatial organization. In contrast, SVGs uniquely identified by SPARK and SPARK-X showed weaker and less well-defined spatial patterns (Supplementary Fig. 5). We next performed GO enrichment analysis on both clusters of VISGP-specific SVGs. At an FDR threshold of 0.05, 370 GO terms were significantly enriched (Fig. 3f). Among them, "cytoplasmic translation" (GO:0002181; $p = 3.13 \times 10^{-21}$) and "cytosolic ribosome" (GO:0022626; $p = 8.26 \times 10^{-23}$) were the most enriched, consistent with elevated protein synthesis activity in tumor cells.

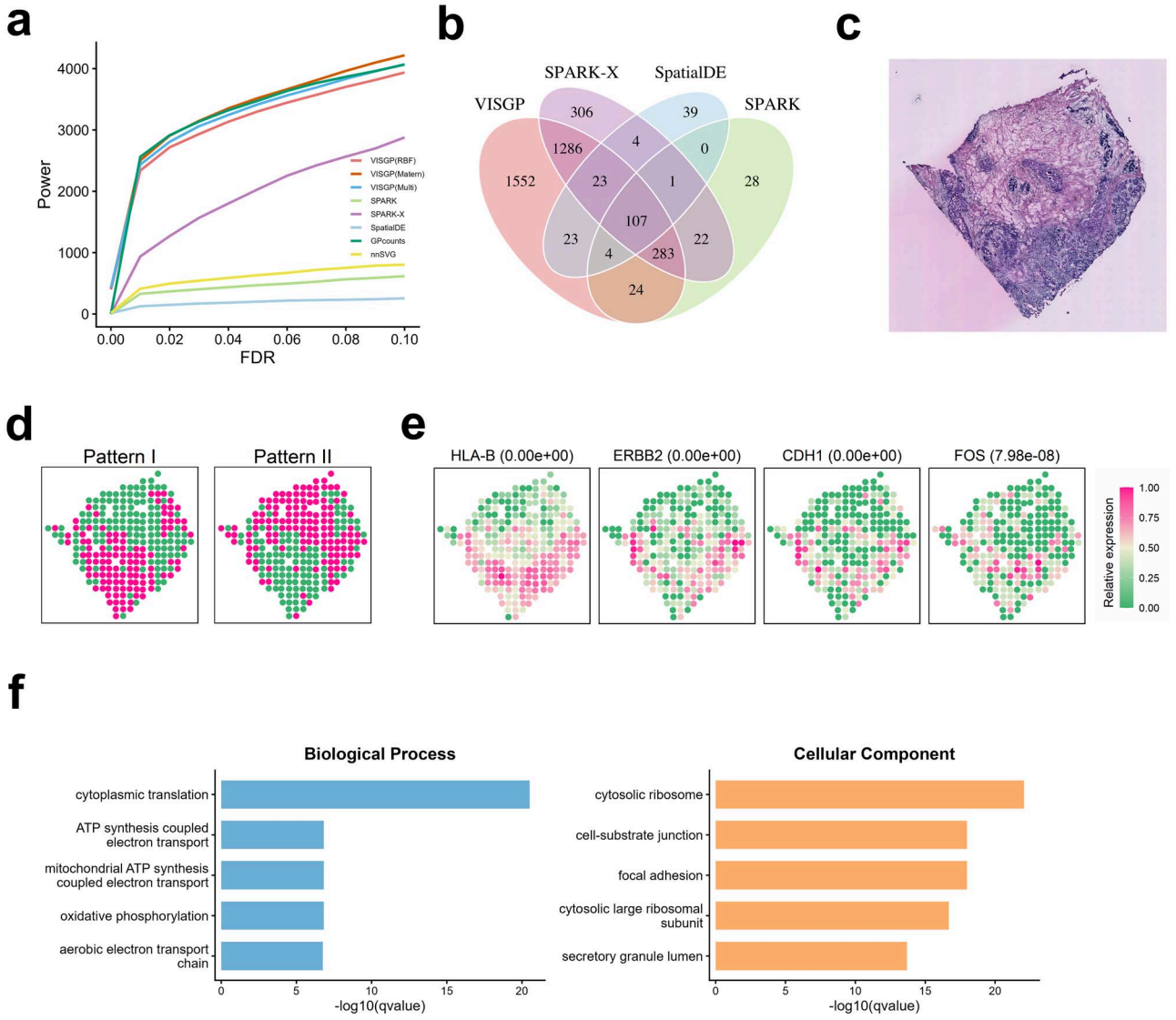

Figure 3 Analyzing the human breast cancer dataset. (a) Power plot shows the number of SV genes detected across FDR thresholds in the breast cancer dataset. The benchmark includes VISGP with RBF, Matérn, and multi-kernel (RBF + Matérn) settings, together with SpatialDE, SPARK, SPARK-X, GPcounts, and nnSVG. (b) Venn diagram shows the overlap between SV genes identified by VISGP (RBF), SpatialDE, SPARK-X, and SPARK based on an FDR cutoff of 0.05. (c) Image of hematoxylin and eosin-stained human breast cancer from [24]. (d) According to the SV genes identified by VISGP, two different spatial expression patterns are summarized. (e) Spatial expression pattern for four known tumor genes. The FDRs for the four genes from VISGP are shown inside parentheses. (f) Bar plot for GO enrichment analysis on 3302 SV genes obtained by VISGP based on an FDR cutoff of 0.05.

We further examined enrichment by spatial pattern (Supplementary Fig. 6). Pattern I genes were associated with "extracellular matrix organization" (GO:0030198; $p = 7.08 \times 10^{-17}$), "focal adhesion" (GO:0005925; $p = 1.72 \times 10^{-19}$), and "blood vessel morphogenesis" (GO:0048514; $p = 4.99 \times 10^{-10}$), reflecting extracellular matrix (ECM) remodeling, angiogenesis, and cell–matrix interactions within the tumor microenvironment. In contrast, Pattern II genes were enriched in "oxidative phosphorylation" (GO:0006119; $p = 1.14 \times 10^{-12}$) and "ATP synthesis coupled electron transport" (GO:0042773; $p = 4.85 \times 10^{-12}$), suggesting preserved mitochondrial metabolism in adjacent nontumorous regions. Together, these analyses demonstrate that VISGP couples sensitive detection of SVGs with clear delineation of biologically meaningful tumor microenvironment structure.

Variational inference-assisted sparse Gaussian-process detects unique spatial patterns in human cutaneous squamous cell carcinoma

In this third application, we applied VISGP to a human cutaneous SCC dataset from a left forearm specimen [27]. In total, 16 383 genes were profiled across 521 spots. Consistent with the breast cancer data, VISGP, SPARK, and SPARK-X showed proper null calibration under the permuted null, whereas SpatialDE did not (Supplementary Fig. 1c). On the real spatial data, VISGP (RBF) identified 3213 SVGs at FDR = 0.05. nnSVG detected 515 SVGs, while SPARK-X, SPARK, and SpatialDE detected 3044, 478, and 216 SVGs, respectively (Fig. 4a and b). We assessed spatial autocorrelation of the SVGs identified by the methods

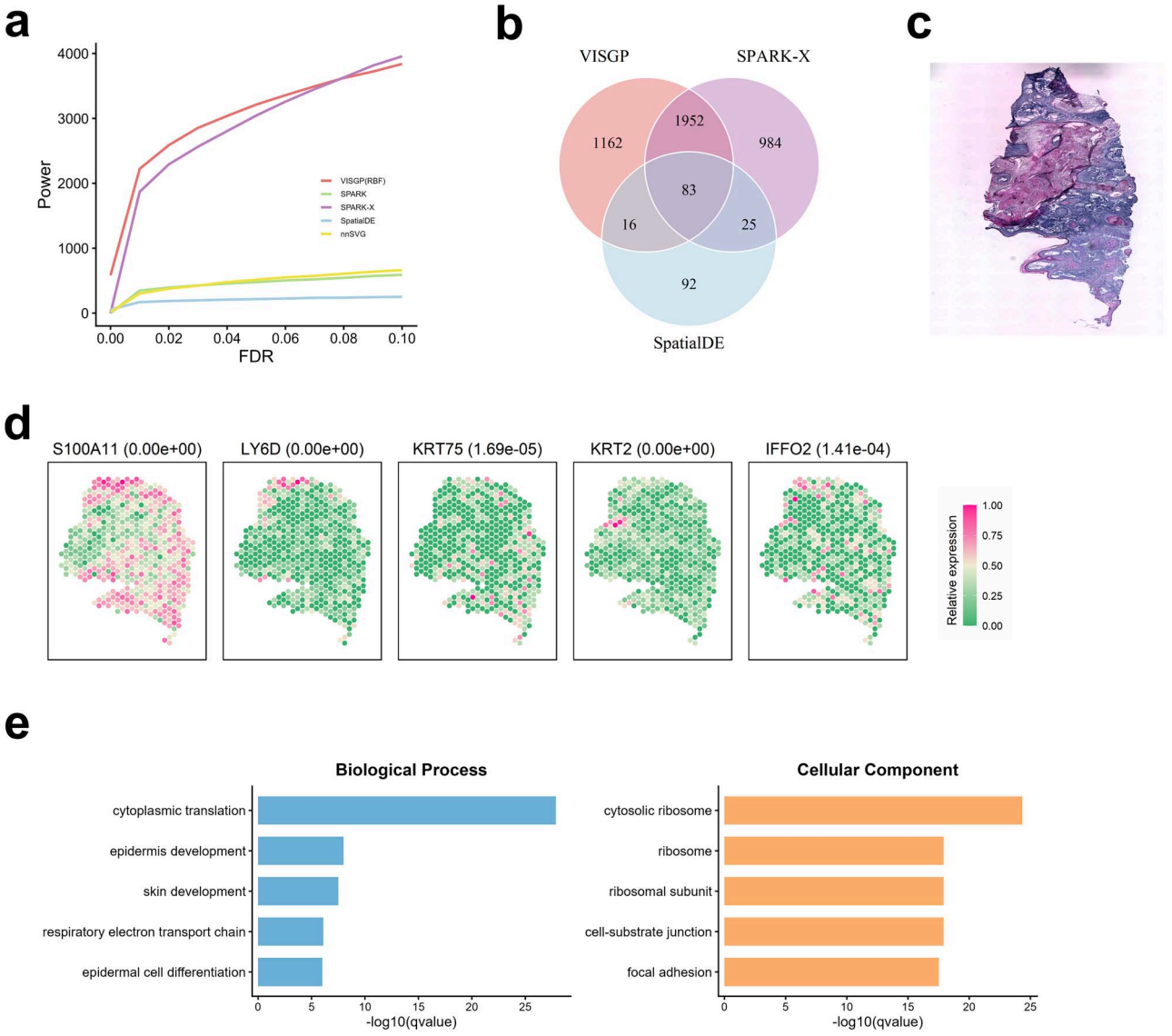

Figure 4 Analyzing the human SCC dataset. (a) Power plot shows the number of SV genes detected by different methods in a series of FDRs in this data. (b) Venn diagram shows the overlap between SV genes identified by VISGP, SpatialDE, and SPARK-X based on an FDR cutoff of 0.05. (c) Image of hematoxylin and eosin-stained human SCC. (d) Spatial expression pattern for some representative genes. The first three genes (S100A11, LY6D, and KRT75) are associated with tumors, and the last two (KRT2, IFFO2) are associated with nontumor adjacent stroma. The FDRs for the five genes from VISGP are shown inside parentheses. (e) Bar plot for GO enrichment analysis on 10 828 SV genes obtained by VISGP on the basis of an FDR cutoff of 0.05.

using Moran's *I*. VISGP yielded consistently higher Moran's *I* values than SPARK-X and SpatialDE (Supplementary Fig. 7a), indicating stronger spatial structure among the detected genes. When restricted to SVGs uniquely identified by each method (Supplementary Fig. 7b), VISGP still produced higher Moran's *I* scores, confirming that its method-specific discoveries showed greater spatial coherence.

Several key genes highlighted in the original study were also recovered by VISGP (Fig. 4d; Supplementary Fig. 8a–c), including tumor-associated, immunosuppression-related, and T cell marker genes. Notably, VISGP uniquely identified additional SVGs that revealed biologically relevant patterns not captured by other methods (Supplementary Fig. 8d). For example, *TNC*, *ITGA5*, and *LAMB2* delineate an ECM–integrin axis consistent with fibroblast-driven matrix remodeling and invasion. Recent studies have summarized *TNC*'s matricellular roles in mechanical and immune niches and discussed emerging TNC-targeting strategies, while *ITGA5* and *LAMB2* have been associated with pro-invasive and prognostic programs across cancers [28–30]. In addition, *CSF1R* and *ENG* were uniquely detected by VISGP as SVGs, revealing distinct myeloid-enriched and angiogenic niches within squamous carcinoma tissues—findings consistent with recent evidence showing that myeloid remodeling can drive resistance to CSF1R inhibition [31] and that *ENG* acts as a key endothelial regulator of VEGF-driven angiogenic sprouting [32].

We next performed GO enrichment analysis on the SVGs identified by VISGP. A total of 335 GO terms were significantly enriched (Fig. 4e). Terms such as "cytoplasmic translation" (GO:0002181; $p = 1.37 \times 10^{-28}$) and "cytosolic ribosome" (GO:0022626; $p = 4.61 \times 10^{-25}$) were highly enriched, indicating enhanced ribosomal

biogenesis and protein synthesis consistent with the proliferative activity of malignant keratinocytes. Processes including "epidermis development" (GO:0008544; $p = 1.06 \times 10^{-8}$), "skin development" (GO:0043588; $p = 3.03 \times 10^{-8}$), and "keratinocyte differentiation" (GO:0030216; $p = 1.43 \times 10^{-6}$) reflected epithelial lineage programs within SCC tissue structure. In addition, enrichment of "cell–substrate junction" (GO:0030055; $p = 1.23 \times 10^{-18}$) and "focal adhesion" (GO:0005925; $p = 3.11 \times 10^{-18}$) indicated active cell–matrix interactions, a hallmark of invasive SCC fronts. Together, these results demonstrate that VISGP captures spatially organized transcriptional programs integrating protein synthesis, epithelial differentiation, and matrix adhesion in squamous carcinoma.

Variational inference-assisted sparse Gaussian-process reveals spatially constrained ligand–receptor interactions in pancreatic ductal adenocarcinoma

To evaluate VISGP for spatially dependent cell–cell communication, we applied it to an SRT dataset of human PDAC [33]. The dataset comprised multiple histologically annotated regions—including cancerous, pancreatic, ductal, and stromal areas—based on H&E staining, and contained expression profiles of 25 754 genes across 428 spatially resolved spots. We modeled the spatial expression of each gene in a curated ligand–receptor database [34] using VISGP and quantified spatial similarity between ligand–receptor pairs using an asymmetric Kullback–Leibler (KL) divergence, which preserves the directional interpretation of ligand–receptor signaling (ligand → receptor). Statistically significant ligand–receptor interactions were determined by permutation testing of KL values (Fig. 5a). To construct the null distribution, we keep each ligand fixed and randomly re-pair it with receptors drawn from the same receptor set in the ligand–receptor database. KL is computed on normalized (sum-to-one) spatial profiles, which reduces sensitivity to overall expression magnitude.

We identified 95 ligand–receptor pairs showing statistically significant spatially coordinated expression ($P < .05$). Spatial heatmaps revealed co-localized or adjacent expression domains for these pairs, including *TIMP1–CD63* (Fig. 5b; additional examples in Supplementary Fig. 9). *TIMP1*, a tissue inhibitor of metalloproteinases, has been linked to poor prognosis in multiple cancers, including PDAC [35]. Its interaction with *CD63* promotes neutrophil extracellular trap formation within the tumor microenvironment, contributing to immune evasion and metastatic potential [36]. Beyond the *TIMP1–CD63* axis, VISGP revealed additional PDAC-relevant ligand–receptor pairs. *LAMC2–ITGB4* marked laminin–integrin adhesion tracks at invasive tumor fronts, where laminin-332 engagement with integrin $\beta 4$ reinforces epithelial–stromal attachment and facilitates epithelial–mesenchymal transition, promoting collective invasion [37, 38]. *SEMA3F–PLXNA3* represented a class-3 semaphorin signaling axis mediating tumor–nerve interactions; binding of *SEMA3F* to *PLXNA3* activates axon-guidance-like pathways that support neural remodeling and perineural invasion within desmoplastic stroma [39].

To assess the added value of spatial modeling, we compared VISGP-inferred ligand–receptor interactions with those derived from a matched single-cell RNA-sequencing (scRNA-seq) dataset from the same patient (PDAC-A) [33] analyzed using CellPhoneDB [40]. This scRNA-seq analysis identified 156 ligand–receptor pairs across diverse cell populations, including acinar cells, Ductal–MHC–Class II cells, Ductal–CRISP3-high centroacinar-like cells, and mDCs. However, none of the four representative ligand–receptor pairs detected by VISGP (Fig. 5b) were recovered by the scRNA-seq analysis. Among the 95 ligand–receptor pairs identified by VISGP, only 10 overlapped with CellPhoneDB results, indicating that VISGP captured 85 spatially constrained and tissue-specific communication events not detectable by conventional scRNA-seq-based inference. This difference is expected because VISGP leverages spatial proximity to prioritize ligand–receptor pairs with co-localized or short-range adjacent expression domains within the tissue. Together, these findings demonstrate that incorporating spatial information substantially enhances discovery of context-dependent signaling interactions within the tumor microenvironment.

Pathway enrichment analysis further showed that VISGP-identified ligand–receptor pairs were significantly associated with key cancer-related signaling pathways, including PI3K–Akt, ECM–receptor interaction, cytokine–cytokine receptor interaction, and focal adhesion (Fig. 5d). Among these, ECM–receptor interaction and focal adhesion exhibited strong spatial dependence, consistent with direct cell–matrix contact and mechanical coupling at the tumor–stroma interface. In contrast, cytokine–cytokine receptor signaling displayed moderate spatial dependence through short-range paracrine gradients within immune and fibroblast compartments, supporting tumor-promoting crosstalk in the local microenvironment [41]. Although PI3K–Akt signaling is primarily intracellular, it acts as a downstream effector of these spatially anchored signaling axes and is frequently hyperactivated in pancreatic cancer, contributing to apoptosis resistance and therapeutic evasion [42].

Discussion and conclusion

Spatial transcriptomics enables quantitative dissection of gene expression within intact tissue structure [1, 2, 5]; however, extracting biologically meaningful spatial organization remains a key analytical challenge [10–12]. In this study, we developed VISGP framework that achieves scalable, gene-specific modeling of spatial variability across tissues and disease contexts. It was developed as a component of the scbean Python toolkit [43], which integrates a range of models for single-cell data analysis [44–46]. Through extensive benchmarking on simulated and real datasets, VISGP demonstrated strong and competitive sensitivity and robustness among established SVG detection methods, while maintaining strict false discovery control. Beyond methodological innovation, VISGP revealed biologically meaningful spatial architectures—ranging from the laminar organization of the MOB [24] to tumor–stroma interaction networks in breast cancer, squamous carcinoma, and PDAC [27, 33]. Together, these findings establish VISGP as a generalizable statistical foundation for decoding spatial gene regulation and intercellular communication in complex tissues.

VISGP introduces two key algorithmic advances that collectively enhance both scalability and flexibility in modeling SRT data. First, by combining sparse GP approximation with variational inference [17, 18], VISGP circumvents the cubic computational bottleneck of conventional GP models, reducing complexity from $\mathcal{O}(n^3)$ to $\mathcal{O}(nm^2)$ while maintaining model accuracy and efficiency. The use of inducing variables allows the model to compress spatial information into a low-dimensional representation. Second, joint optimization of kernel hyperparameters through variational inference provides gene-specific adaptation of spatial covariance structures, improving sensitivity to diverse spatial patterns. Compared with existing methods

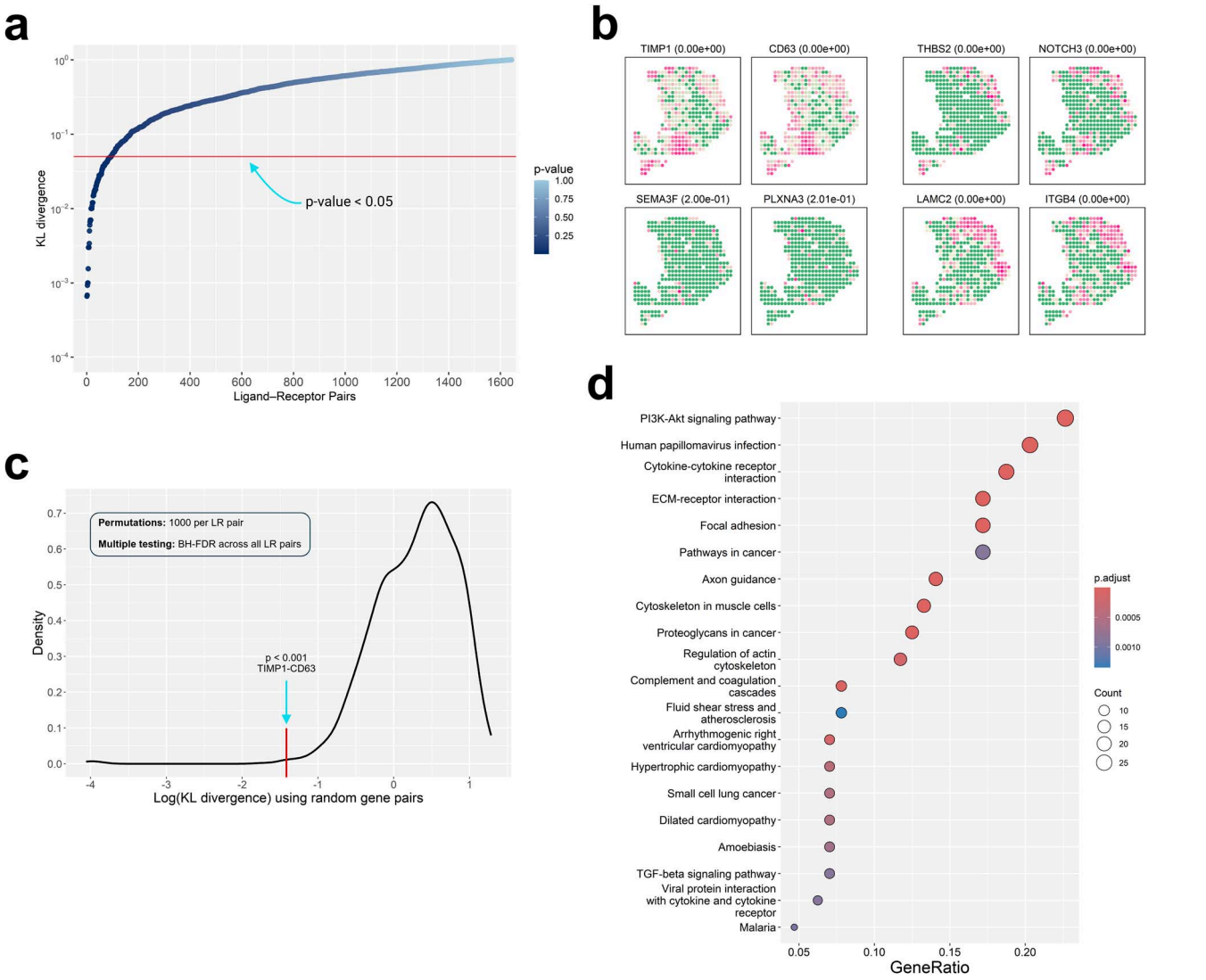

Figure 5 Analysis results of ligand–receptor pairs predicted by VISGP. (a) Scatter plot shows the ligand–receptor pairs and their corresponding KL divergence. The color of the points represent the P-value. (b) Heatmaps of four ligand–receptor pairs that have similar expression distribution in space. (c) The distribution of KL divergence of randomly sampled gene pairs. We performed the statistical test on this background distribution. Permutation testing used 1000 permutations per ligand–receptor pair, and P-values were adjusted across all ligand–receptor pairs using the BH FDR procedure. (d) KEGG enrichment shows that ligand–receptor pairs predicted by VISGP are closely related to human PDAC.

such as SPARK, SpatialDE, and SPARK-X [10–12], which employ fixed or empirically tuned kernels, VISGP dynamically learns the optimal spatial scale (ℓ) and signal variance (τ) for each gene, yielding more accurate representation of heterogeneous tissue structure. Comparative simulation further confirmed that VISGP maintains high sensitivity and strict FDR control across distinct spatial configurations, demonstrating robust statistical performance across parameter regimes.

The performance of VISGP was first validated on the MOB, a well-characterized tissue with distinct laminar organization [24]. By accurately recovering known spatial expression layers—glomerular, mitral, and granule cell zones—VISGP demonstrated strong sensitivity to subtle spatial gradients that mirror true neuroanatomical boundaries. Compared with other approaches, VISGP produced sharper spatial delineation and higher spatial autocorrelation scores, consistent with its ability to model gene-specific covariance structures. Functional enrichment of VISGP-identified SVGs further revealed synaptic and neuronal processes, including neuron-to-neuron and asymmetric synapses, confirming that VISGP captures biologically meaningful spatial variation reflective of tissue structure rather than noise. This analysis establishes VISGP as a reliable approach for dissecting spatially organized gene programs in complex tissues. We next applied VISGP across multiple tumor contexts to evaluate its capacity to resolve spatial transcriptional heterogeneity in malignant tissues. In breast cancer, VISGP delineated two dominant spatial programs distinguishing metabolically active paracancerous regions from ECM–remodeling tumor cores, capturing the metabolic–stromal dichotomy characteristic of invasive carcinomas [24]. In cutaneous SCC, VISGP recovered both known and novel spatial gene modules, including TNC, ITGA5, and LAMB2, which highlight an ECM–integrin axis mediating fibroblast-driven invasion and mechanical coupling [28–30]. These findings parallel emerging evidence that the tumor stroma orchestrates invasion and immune modulation through integrin–laminin signaling [37]. In PDAC, VISGP identified spatially constrained ligand–receptor pairs, including TIMP1–CD63 and LAMC2–ITGB4,

that define localized tumor–stroma communication niches [33, 35, 36]. Together, these analyses reveal conserved spatial axes—ECM–receptor interaction, focal adhesion, and PI3K–Akt signaling—through which VISGP uncovers the molecular structure linking cell–matrix adhesion, metabolic remodeling, and intercellular communication across distinct tumor ecosystems.

Beyond identifying SVGs, VISGP extends naturally to the inference of spatially constrained ligand–receptor interactions. There are currently three main strategies for inferring cell-to-cell communication [47]. The first relies exclusively on scRNA-seq to predict ligand–receptor interactions at the cell-type level, with notable tools including CellChat [48] and CellPhoneDB [40]. However, these methods lack spatial context, which can lead to false positives, as most intercellular communication is restricted to spatially proximal cells. A second class of methods integrates scRNA-seq with SRT to improve spatial resolution. These approaches refine cell-type distributions using spatial data and enhance communication inference accordingly. A third class, represented by tools such as Giotto [49], infers communication directly from ST data. While these spatially informed approaches provide improvements, they often focus on cell-type distributions rather than the spatial expression patterns of individual genes. Notably, among the four major modes of cell-to-cell communication in the human body, three (paracrine, juxtacrine, and autocrine) require spatial proximity between interacting cells [50]. This implies that the expression patterns of functionally interacting ligands and receptors should exhibit spatial overlap. By explicitly modeling the spatial co-expression of ligand and receptor transcripts, VISGP captures signaling relationships that are contingent on tissue structure rather than cell-type proximity. In PDAC, this approach revealed spatially co-localized ligand–receptor pairs, including TIMP1–CD63 and LAMC2–ITGB4, which define distinct communication niches at the tumor–stroma interface [33, 35]. Many of these interactions converge on signaling cascades such as ECM–receptor, cytokine–cytokine receptor, focal adhesion, and PI3K–Akt pathways, which collectively govern oncogenic processes including cell proliferation, adhesion, and stromal communication [27, 41, 42]. These findings highlight how spatial modeling uncovers mechanistically grounded patterns of tumor signaling that would remain undetected in single-cell analyses lacking positional context [40]. Importantly, several of the spatially enriched signaling pathways identified by VISGP, such as PI3K–Akt and focal adhesion, have corresponding therapeutic inhibitors in clinical evaluation, suggesting that spatially localized pathway activation could inform region-specific therapeutic vulnerabilities within the tumor microenvironment. The cross-tissue consistency of ECM–adhesion and PI3K–Akt axes underscores that spatially anchored transcriptional regulation represents a unifying principle across morphologically distinct disease contexts.

Deep learning approaches for spatial transcriptomics can be powerful for capturing complex, nonlinear spatial structure, but their outputs are often model- and training-dependent and may not directly provide gene-level significance measures under an explicit null. In contrast, VISGP offers a statistically grounded and interpretable framework: it models spatial covariance explicitly, learns gene-specific spatial parameters, and supports hypothesis testing with FDR-controlled SVG calling. We view these paradigms as complementary—deep learning excels at feature learning and domain discovery, while GP-based inference provides transparent statistical testing and uncertainty-aware interpretation of spatial effects.

Despite its strong performance, VISGP has several limitations. The inducing-point number m controls the accuracy–efficiency trade-off, with computational cost scaling as $\mathcal{O}(nm^2)$ and memory approximately $\mathcal{O}(nm)$. We selected $m = 20$ as the default value because it offered stable SVG detection while maintaining computational efficiency, although the optimal choice of m may depend on dataset size and spatial complexity. To evaluate this, we assessed the power curves of VISGP across different FDR thresholds using varying inducing-point numbers ($m = 5, 10, 20, 40, 80$) on the MOB and human breast cancer datasets (Supplementary Fig. 1d and e), respectively. VISGP uses a GP likelihood on gene-wise standardized expression values, this standardization is a model-scale requirement for Gaussian GP fitting rather than a replacement for standard SRT library-size normalization. But comparisons with GPcounts, which models filtered raw counts using a negative-binomial likelihood, showed broadly consistent SVG detection patterns in real SRT datasets. Similarly, additional Matérn and RBF+Matérn kernel analyses produced comparable detection profiles, supporting robustness across representative kernel settings, although further kernel customization may be useful for specific tissue architectures. Finally, deep learning-based spatial methods were not included in the main benchmark because many are designed primarily for spatial domain detection or representation learning rather than gene-level SVG testing with calibrated *P*-values and FDR control. Future integration of VISGP with complementary spatial modalities, such as spatial proteomics, imaging mass spectrometry, or multiplexed immunofluorescence, may provide richer insights into coordinated molecular and phenotypic organization.

Key Points

- By leveraging variational inference within a Gaussian-process modeling framework, variational inference-assisted sparse Gaussian-process (VISGP) enables flexible and robust modeling of complex spatial gene expression while maintaining practical computational feasibility.
- VISGP showed strong and competitive performance among established spatially variable genes detection methods while exhibiting well-calibrated statistical behavior.
- VISGP facilitates the discovery of previously unrecognized spatial regulatory programs and cellular communication patterns with direct relevance to cellular heterogeneity and cancer pathology.

Acknowledgements

We thank Prof. Xiang Zhou and his team for insightful discussions on the statistical modeling, which improved this work.

Author contributions

Zhicong Wang (Conceptualization, Investigation, Visualization, Supervision, Writing—original draft), Jing Li (Conceptualization, Writing—review & editing), Liqing Xie (Investigation, Visualization, Writing—original draft), Yiran Wang (Conceptualization, Methodology, Investigation, Writing—original draft), Yongtian Wang (Writing—review & editing), Jing Chen (Writing—review & editing), Xuequn Shang (Writing—review & editing), Xingyi Li (Investigation), Zhaowen Liu

(Writing—review & editing), Jialu Hu (Conceptualization, Methodology, Supervision, Writing—review & editing)

Supplementary material

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

Conflicts of interest

None declared.

Funding

This work was supported by the National Natural Science Foundation of China (grant no. 62572398) and Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (grant no. JYB2025XDXM202).

Data availability

The datasets analyzed in this study are publicly available. The MOB and human breast cancer datasets were obtained from the Spatial Research website (https://www.spatialresearch.org). The human cutaneous SCC and PDAC datasets were obtained from the Gene Expression Omnibus (GEO) under accession numbers `GSE144240` and `GSE111672`, respectively. VISGP is implemented in the Python package scbean (version 0.6.0) available on PyPI, and the source code is available at https://github.com/jhu99/scbean.

References

1. Hickey JW, Neumann EK, Radtke AJ *et al.* Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging. *Nat Methods* 2021;**19**:284–95. https://doi.org/10.1038/s41592-021-01316-y
2. Zhao C, Zhongli X, Wang X *et al.* Innovative super-resolution in spatial transcriptomics: a transformer model exploiting histology images and spatial gene expression. *Brief Bioinform* 2024;**25**:bbae052. https://doi.org/10.1093/bib/bbae052
3. Srivatsan SR, Regier MC, Barkan E *et al.* Embryo-scale, single-cell spatial transcriptomics. *Science* 2021;**373**:111–7. https://doi.org/10.1126/science.abb9536
4. Nolan E, Lindeman GJ, Visvader JE. Deciphering breast cancer: from biology to the clinic. *Cell* 2023;**186**:1708–28. https://doi.org/10.1016/j.cell.2023.01.040
5. 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
6. Teves JM, Won KJ. Mapping cellular coordinates through advances in spatial transcriptomics technology. *Mol Cells* 2020;**43**:591.
7. Zeng Z, Li Y, Li Y *et al.* Statistical and machine learning methods for spatially resolved transcriptomics data analysis. *Genome Biol* 2022;**23**:83. https://doi.org/10.1186/s13059-022-02653-7
8. Larsson L, Frisén J, Lundeberg J. Spatially resolved transcriptomics adds a new dimension to genomics. *Nat Methods* 2021;**18**:15–8. https://doi.org/10.1038/s41592-020-01038-7
9. Edsgärd D, Johnsson P, Sandberg R. Identification of spatial expression trends in single-cell gene expression data. *Nat Methods* 2018;**15**:339–42. https://doi.org/10.1038/nmeth.4634
10. 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
11. Sun S, Zhu J, Zhou X. Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nat Methods* 2020;**17**:193–200. https://doi.org/10.1038/s41592-019-0701-7
12. Zhu J, Sun S, Zhou X. SPARK-X: non-parametric modeling enables scalable and robust detection of spatial expression patterns for large spatial transcriptomic studies. *Genome Biol* 2021;**22**:1–25.
13. BinTayyash N, Sokratia Georgaka ST, John SA *et al.* Non-parametric modelling of temporal and spatial counts data from RNA-seq experiments. *Bioinformatics* 2021;**37**:3788–95. https://doi.org/10.1093/bioinformatics/btab486
14. Weber LM, Saha A, Datta A *et al.* nnSVG for the scalable identification of spatially variable genes using nearest-neighbor gaussian processes. *Nat Commun* 2023;**14**:4059. https://doi.org/10.1038/s41467-023-39748-z
15. Hao M, Hua K, Zhang X. SOMBE: a scalable method for identifying spatially variable genes with self-organizing map. *Bioinformatics* 2021;**37**:4392–8. https://doi.org/10.1093/bioinformatics/btab471
16. Zhang C, Dong K, Aihara K *et al.* STAMarker: determining spatial domain-specific variable genes with saliency maps in deep learning. *Nucleic Acids Res* 2023;**51**:e103–3. https://doi.org/10.1093/nar/gkad801
17. Titsias M. Variational learning of inducing variables in sparse Gaussian processes. In: van Dyk D, Welling M (eds.), *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*. Hilton Clearwater Beach Resort, Clearwater Beach, Florida, USA: PMLR; 2009, 567–74.
18. Hensman J, Fusi N, Lawrence ND. Gaussian processes for big data. In: Nicholson A, Smyth P (eds.), *Proceedings of the 29th Conference on Uncertainty in Artificial Intelligence (UAI 2013)*. AUAI Press; 2013, 282–90.
19. Tran D, Ranganath R, Blei DM. The variational Gaussian process. In: Bengio Y, LeCun Y (eds.), *Proceedings of the International Conference on Learning Representations (ICLR)*. San Juan, Puerto Rico, May 2-4, 2016.
20. Schweiger R, Weissbrod O, Rahmani E *et al.* RL-SKAT: an exact and efficient score test for heritability and set tests. *Genetics* 2017;**207**:1275–83. https://doi.org/10.1534/genetics.117.300395
21. Quinonero-Candela J, Rasmussen CE. A unifying view of sparse approximate gaussian process regression. *The Journal of Machine Learning Research* 2005;**6**:1939–59.
22. Davies RB. Algorithm AS 155: The distribution of a linear combination of χ 2 random variables. *J R Stat Soc: Series C (Applied Statistics)* 1980;**29**:323–33. https://doi.org/10.2307/2346911
23. Thissen D, Steinberg L, Kuang D. Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. *J Educ Behav Stat* 2002;**27**:77–83. https://doi.org/10.3102/10769986027001077
24. 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
25. Berx G, Becker K-F, Höfler H *et al.* Mutations of the human E-cadherin (CDH1) gene. *Hum Mutat* 1998;**12**:226–37. https://doi.org/10.1002/(SICI)1098-1004(1998)12:4⟨226::AID-HUMU2⟩3.0.CO;2-D
26. Verma IM. Proto-oncogene fos: a multifaceted gene. *Trends Genet* 1986;**2**:93–6. https://doi.org/10.1016/0168-9525(86)90191-5
27. Ji AL, Rubin AJ, Thrane K *et al.* Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* 2020;**182**:497–514.e22. https://doi.org/10.1016/j.cell.2020.05.039

28. Xie Y, Liu F, Yunfei W *et al.* Inflammation in cancer: therapeutic opportunities from new insights. *Mol Cancer* 2025;**24**:51. https://doi.org/10.1186/s12943-025-02243-8
29. Xiao Y, Tao P, Zhang K *et al.* Myofibroblast-derived extracellular vesicles facilitate cancer stemness of hepatocellular carcinoma via transferring ITGA5 to tumor cells. *Mol Cancer* 2024;**23**:262. https://doi.org/10.1186/s12943-024-02170-0
30. Hong J, Jin HJ, Choi MR *et al.* Matrisomics: beyond the extracellular matrix for unveiling tumor microenvironment. *Biochim Biophys Acta Rev Cancer* 2024;**1879**:189178.
31. Watson SS, Zomer A, Fournier N *et al.* Fibrotic response to anti-CSF-1R therapy potentiates glioblastoma recurrence. *Cancer Cell* 2024;**42**:1507–1527.e11. https://doi.org/10.1016/j.ccell.2024.08.012
32. Sharma S, Ehrlich M, Zhang M *et al.* NRP1 interacts with endoglin and VEGFR2 to modulate VEGF signaling and endothelial cell sprouting. *Commun Biol* 2024;**7**:112. https://doi.org/10.1038/s42003-024-05798-2
33. Moncada R, Barkley D, Wagner F *et al.* Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol* 2020;**38**:333–42. https://doi.org/10.1038/s41587-019-0392-8
34. Ramilowski JA, Goldberg T, Harshbarger J *et al.* A draft network of ligand–receptor-mediated multicellular signalling in human. *Nat Commun* 2015;**6**:1–12. https://doi.org/10.1038/ncomms8866
35. Schoeps B, Eckfeld C, Prokopchuk O *et al.* TIMP1 triggers neutrophil extracellular trap formation in pancreatic cancer. *Cancer Res* 2021;**81**:3568–79. https://doi.org/10.1158/0008-5472.CAN-20-4125
36. Wang C-A, Hou Y-C, Hong Y-K *et al.* Intercellular TIMP-1-CD63 signaling directs the evolution of immune escape and metastasis in KRAS-mutated pancreatic cancer cells. *Mol Cancer* 2025;**24**:25. https://doi.org/10.1186/s12943-024-02207-4
37. Nonnast E, Mira E, Mañes S. The role of laminins in cancer pathobiology: a comprehensive review. *J Transl Med* 2025;**23**:83. https://doi.org/10.1186/s12967-025-06079-0
38. Fangqiu Y, Zeng G, Yang L *et al.* LAMB3: central role and clinical significance in neoplastic and non-neoplastic diseases. *Biomed Pharmacother* 2024;**178**:117233.
39. Thielman NRJ, Funes V, Davuluri S *et al.* Semaphorin 3D promotes pancreatic ductal adenocarcinoma progression and metastasis through macrophage reprogramming. *Sci Adv* 2024;**10**:eadp0684. https://doi.org/10.1126/sciadv.adp0684
40. Efremova M, Vento-Tormo M, Teichmann SA *et al.* CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc* 2020;**15**:1484–506. https://doi.org/10.1038/s41596-020-0292-x
41. Hartupee C, Nagalo BM, Chabu CY *et al.* Pancreatic cancer tumor microenvironment is a major therapeutic barrier and target. *Front Immunol* 2024;**15**:1287459. https://doi.org/10.3389/fimmu.2024.1287459
42. Asano T, Yamasaki S, Tashiro K. Targeting the PI3K/AKT pathway in pancreatic cancer: rationale and progress. *Cancers* 2022;**14**:177.
43. Zhang H, Wang Y, Lian B *et al.* Scbean: a python library for single-cell multi-omics data analysis. *Bioinformatics* 2024;**40**:btae053. https://doi.org/10.1093/bioinformatics/btae053
44. Wang Y, Lian B, Zhang H *et al.* A multi-view latent variable model reveals cellular heterogeneity in complex tissues for paired multimodal single-cell data. *Bioinformatics* 2023;**39**:btad005. https://doi.org/10.1093/bioinformatics/btad005
45. Hu J, Chen M, Zhou X *et al.* Effective and scalable single-cell data alignment with non-linear canonical correlation analysis. *Nucleic Acids Res* 2022;**50**:e21–1. https://doi.org/10.1093/nar/gkab1147
46. Hu J, Zhong Y, Shang X. A versatile and scalable single-cell data integration algorithm based on domain-adversarial and variational approximation. *Brief Bioinform* 2022;**23**:bbab400. https://doi.org/10.1093/bib/bbab400
47. Almet AA, Cang Z, Jin S *et al.* The landscape of cell–cell communication through single-cell transcriptomics. *Curr Opin Syst Biol* 2021;**26**:12–23. https://doi.org/10.1016/j.coisb.2021.03.007
48. Jin S, Guerrero-Juarez CF, Zhang L *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* 2021;**12**:1–20. https://doi.org/10.1038/s41467-021-21246-9
49. 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
50. Armingol E, Officer A, Harismendy O *et al.* Deciphering cell–cell interactions and communication from gene expression. *Nat Rev Genet* 2021;**22**:71–88. https://doi.org/10.1038/s41576-020-00292-x

Briefings in Bioinformatics, 2026, **27**, bbag371
https://doi.org/10.1093/bib/bbag371
Problem Solving Protocol