Inferring spatial single-cell-level interactions through interpreting cell state and niche correlations learned by self-supervised graph transformer

Xiao Xiao[1], Le Zhang[2,3], Hongyu Zhao[1] & Zuoheng Wang[1,4]

[1]Department of Biostatistics, Yale University School of Public Health, New Haven, CT, USA.

[2]Department of Neurology, Yale University School of Medicine, New Haven, CT, USA.

[3]Department of Neuroscience, Yale University School of Medicine, New Haven, CT, USA.

[4]Department of Biomedical Informatics & Data Science, Yale University School of Medicine, New Haven, CT, USA.

Correspondence

Zuoheng Wang, email: zuoheng.wang@yale.edu

Abstract

Cell-cell interactions (CCI), driven by distance-dependent signaling, are important for tissue development and organ function. While imaging-based spatial transcriptomics offers unprecedented opportunities to unravel CCI at single-cell resolution, current analyses face challenges such as limited ligand-receptor pairs measured, insufficient spatial encoding, and low interpretability. We present GITIII, a lightweight, interpretable, self-supervised graph transformer-based model that conceptualizes cells as words and their surrounding cellular neighborhood as context that shapes the meaning or state of the central cell. GITIII infers CCI by examining the correlation between a cell's state and its niche, enabling us to understand how sender cells influence the gene expression of receiver cells, visualize spatial CCI patterns, perform CCI-informed cell clustering, and construct CCI networks. Applied to four spatial transcriptomics datasets across multiple species, organs, and platforms, GITIII effectively identified and statistically interpreted CCI patterns in the brain and tumor microenvironments.

Introduction

In multicellular organisms, cell-cell interactions (CCI) play a pivotal role in maintaining tissue development and organ function. These interactions may occur through direct physical contact mediated by cell adhesion molecules between adjacent cells or through longer-range paracrine and autocrine signaling, where secreted factors activate intracellular signaling cascades that modulate gene expression and cellular behavior. CCI are influenced by multiple factors, including the spatial proximity of interacting cells, the local tissue microenvironment of the receiver cell, the strength of ligand-receptor signaling, and the influence of gene regulatory networks on downstream responses. Together, these factors orchestrate cell function and tissue homeostasis.

Single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics (ST) have offered unprecedented opportunities to unravel intercellular communication. Many existing methods focus on constructing CCI networks between cell types by identifying ligand-receptor pairs through which communication occurs[1]. These approaches span statistical tools for detecting significant interactions[2-6], and techniques that incorporate spatial information through graph neural networks (GNN), optimal transport, or distance-based weighting and thresholding[7-18]. Collectively, they have greatly advanced our understanding of ligand-receptor-mediated molecular interactions[6,19-23]. However, most of these methods treat cells of the same type as homogeneous, often overlooking within-cell-type heterogeneity. In reality, cells of the same cell-type pair may interact only when spatially proximal, but not when distant, due to microenvironmental differences or spatial constraints. Moreover, current approaches tend to focus only on the most statistically extreme communication signals, potentially missing subtle interactions that still have downstream effects on gene expression and cell behavior.

ST, which captures both gene expression and the spatial location of cells within tissue, has become a powerful tool for studying CCI at single-cell resolution, particularly in the context of the tissue microenvironment. However, identifying CCI from imaging-based ST[24-26] remains challenging, as these platforms typically measure only a limited set of genes and lack adequate coverage of ligand-receptor pairs. Thus, many CCI analysis tools that rely on ligand-receptor information to detect significant interactions are not applicable to such data. Additionally, spatial information has not been fully exploit in current CCI analysis methods. Some approaches either restrict CCI inference to a small predefined spatial neighborhood[7,14,27,28], or fail to account for the effect of distance on interaction strength[8], or use a fixed distance-scaling model across all signaling pathways[10,15,29], despite the fact that different ligand-receptor signaling may exhibit distinct distance-dependent patterns in biological systems. Furthermore, GNN-based methods for CCI analysis[30-32] often suffer from inadequate graph encoding and limited interpretability, particularly when using multi-layer architecture. These limitations hinder biological interpretation and reduce their utility for uncovering mechanistic insights into CCI.

An important aspect of CCI inference is to estimate the effects of neighboring sender cells on the gene expression of receiver cells through CCI that, to our knowledge, has not been explicitly studied in previous studies. We introduce GITIII (Graph Inductive bias Transformer for Intercellular Interaction Investigation), an interpretable, self-supervised graph transformer-based language model that treats cells as words and their surrounding cellular neighborhood as a sentence to explore CCI within a cellular network using their spatial and transcriptomic information from imaging-based ST. GITIII infers spatial single-cell-level interactions by examining the correlation between a cell's state and the spatial organization, ligand expression, and the types and states of its neighboring cells. Its interpretable architecture allows us to estimate the impact of sender cells

on the gene expression of receiver cells, visualize spatial patterns of CCI, perform CCI-informed cell clustering, construct CCI networks, and compare CCI strength across conditions. Applied to four diverse ST datasets, GITIII successfully identified and quantitatively interpreted CCI patterns that drive within-cell-type heterogeneity, thus advancing our understanding of brain structure and the tumor microenvironment (TME).

Results

Overview of GITIII

GITIII is a lightweight and interpretable deep learning model for inferring spatial single-cell-level interaction patterns. It takes gene expression, spatial coordinates, and cell type annotations as inputs. The model first decomposes the gene expression levels of each cell into two parts: "cell-type expression," representing the average gene expression across cells of the same type, and "cell-state expression," representing the deviation of its gene expression from the average expression of its cell type (Fig. 1a). For each cell, GITIII constructs a cellular neighborhood of size k based on spatial proximity and learns the interactions between the central cell and its neighboring cells. The cellular neighborhood is represented as a spatial subgraph, where nodes are cells and edges are spatial proximity between cells, or alternatively, as a sentence where cells are words and neighboring cells provide context to shape the meaning or state of the central cell. GITIII employs a single-layer graph transformer to learn interaction patterns and computes an influence tensor that quantifies the impact of neighboring cells on the cell-state expression of the central cell (Fig. 1b, 1c).

To investigate CCI patterns across spatial regions, the influence tensor is used to visualize single-cell-level spatial patterns of CCI by plotting the spatial distance and strength of influence

between interacting cells. We then perform cell clustering within each cell type using CCI-informed microenvironmental features (Fig. 1d). At the cell-type level, significant interactions between cell types are identified using Pearson correlation analysis. For each gene, CCI networks are constructed by identifying sender cell types that significantly influence the expression level of this gene in receiver cell types. At the sample level, we compare CCI network strength between conditions using a microenvironment-adjusted network strength metric (Fig. 1e).

GITIII reveals spatial patterns of CCI in mouse brains

We illustrated the main features of GITIII using a mouse primary motor cortex (PMC) ST dataset generated by MERFISH[33]. Slice 201 from mouse 1 was used for demonstration (Fig. 2a). As expected from physical principles, the locally estimated scatterplot smoothing (LOESS) curve showed that the estimated CCI strength decreases with increasing distance between interacting cells (Fig. 2b).

Single-cell-level interactions within the same cell-type pair are expected to be more similar than interactions across different pairs, because the ligand-receptor expression patterns in sender-receiver pairs, as well as the downstream gene regulatory responses in receiver cells, are generally more similar within a given cell-type combination. Fig. 2c shows that the 10 most frequent cell-type interactions from the selected CCI pairs across all cellular neighborhoods formed distinct clusters in the Uniform Manifold Approximation and Projection (UMAP), where each point represents one CCI pair. This confirms that CCI pairs from the same cell-type combination have similar patterns and that GITIII effectively distinguishes interactions among different cell types.

We then used CCI-informed microenvironmental features to cluster astrocytes and L2/3 IT neurons. Astrocytes are known to display layer-specific morphological and molecular differences[34] and to be functionally diverse across brain regions[35]. In the MERFISH study[33],

astrocytes were clustered into three subgroups: one distributed across all layers and two enriched mostly in the white matter (Extended Data Fig. 1b). In comparison, GITIII identified five subgroups with clear layer-specific distributions (Fig. 2d), where subgroup 1 located in layer 1, subgroup 2 located in layers 2 and 3, and subgroup 5 located in the white matter (Fig. 2e, Extended Data Fig. 1a). This pattern is consistent with the reported heterogeneity of astrocytes across cortical layers in both mouse and human brains[36-39]. Subgroups 1 and 2 showed increased expression of *Cux2* and *Lama3*, while subgroup 5 showed increased expression of *Igfbp5* and *Gfap* (adjusted $p < 1 \times 10^{-5}$, Fig. 2f, Extended Data Fig. 1d), suggesting that GITIII captured microenvironment-driven transcriptional variation. We also observed *Luzp2* expression in several subgroups, consistent with a thalamus-enriched (*LUZP2*$^{+}$*DCLK1*$^{+}$) astrocyte subtype found in the human brain[40]. GITIII further revealed that subgroups 1 and 2 were strongly influenced by L2/3 IT neurons, and subgroup 5 was primarily influenced by oligodendrocytes and L6b neurons (Fig. 2g), confirming that intercellular communication is most likely to occur between cells in close physical proximity (Extended Data Fig. 1a, 1c).

For L2/3 IT neurons, the MERFISH study[33] identified five subgroups without clear spatial boundaries (Extended Data Fig. 1f). In contrast, GITIII detected three subgroups with clear separations (Fig. 2d). Subgroup 1 located in the upper region of layers 2 and 3, closer to layer 1, while subgroups 2 and 3 located in the deeper region of layers 2 and 3, closer to layer 4 (Fig. 2e, Extended Data Fig. 1g). Subgroup 1 showed upregulated expression of *Camk2d*, *Trp53i11*, *Trpc6*, and *Otof*, subgroup 2 showed increased expression of *Rorb* and *Unc5d*, while subgroup 3 showed increased expression of *Sulf2* and *Adra1b* (adjusted $p < 1 \times 10^{-24}$, Fig. 2f, Extended Data Fig. 1h). GITIII inferred that subgroup 1 was primarily influenced by astrocytes, Lamp5, Pvalb, and

Vip neurons, while subgroups 2 and 3 were strongly influenced by L4/5 IT neurons as they were physically closer to layer 4, and subgroup 3 was also influenced by Pvalb neurons (Fig. 2g).

To compare our subgroup results with spatial clustering, we applied BANKSY[41], which clusters all cells based on their transcriptomes and local microenvironment. Like other domain detection methods, BANKSY clusters all cells on a tissue slide, thus each cluster may include multiple cell types. In contrast, GITIII clusters within a cell type to identify subgroups that differ in their surrounding cellular neighborhood. Supplementary Fig. 2 shows the clustering results from BANKSY at different resolutions, restricted to clusters containing at least 20 astrocytes or L2/3 IT neurons. Astrocytes and L2/3 IT neurons not included in these clusters were grouped together for visualization. BANKSY detected several tissue domains containing astrocytes and L2/3 IT neurons distributed across cortical layers, especially at higher resolutions. However, GITIII showed clearer layer-specific separation among astrocyte subgroups (Fig. 2e), while BANKSY grouped a small number of astrocytes and L2/3 IT neurons with other cell types. Together, the consistent spatial separation observed in both methods supports that the subgroups identified by GITIII are driven by CCI-informed microenvironmental features.

We further visualized the inferred CCI and constructed CCI networks that influence the expression of three biologically relevant genes, *Rorb*, *Gfap*, and *Cux2*, chosen based on their high predictive accuracy. Pearson correlation coefficients between the observed and predicted cell-state expression of these genes across all cells were greater than 0.58 (Fig. 2h). We then visualized the spatial distribution of the top 5% of CCI pairs that strongly influence the expression levels in receiver cells (Fig. 2i). For example, interactions between L4/5 IT and L2/3 IT cells were associated with increased expression of *Rorb* (Fig. 2j), where we observed more CCI pairs between L4/5 IT and L2/3 IT subgroup 2 (Fig. 2i, Extended Data Fig. 1g, 1i), consistent with the

upregulated expression of *Rorb* in L2/3 IT subgroup 2 (Fig. 2f). Interactions between oligodendrocytes and astrocytes were associated with increased expression of *Gfap* (Fig. 2i, 2j, Extended Data Fig. 1j), consistent with the inferred interactions between oligodendrocytes and astrocyte subgroup 5 in the white matter (Extended Data Fig. 1c) and the upregulated expression of *Gfap* in astrocyte subgroup 5 (Fig. 2f), and supporting the role of astrocyte-oligodendrocyte interaction in myelin regeneration[42]. *Cux2*, a gene encoding a transcription factor involved in neuronal proliferation and differentiation[43], showed increased expression in astrocytes influenced by L2/3 IT neurons (Fig. 2i, 2j). This result agreed with the upregulated expression of *Cux2* in astrocyte subgroup 2, located in layers 2 and 3 (Extended Data Fig. 1a, 1k), which may related to the influence of neurons on astrocyte morphology and proliferation[44].

GITIII identifies CCI patterns within and across patients with Alzheimer's disease

We analyzed the SEA-AD middle temporal gyrus ST dataset generated by MERFISH[45]. Sample 'H20.33.001.CX28.MTG.02.007.1.02.03' was used to demonstrate within-sample performance (Fig. 3a). As expected, interaction strength generally decreased with increasing distance between interacting cells (Fig. 3b). However, adjacent cells can have weaker interactions than more distant cells. The strength of interactions influencing the expression of *RORB* varied substantially within 18 μm and became consistently weak beyond this distance (Extended Data Fig. 2a). This pattern suggests that spatial proximity alone does not determine interaction strength. Even when two cells are physically close, low ligand-receptor expression can lead to weak interactions. The UMAP of the 10 most frequent cell-type interactions showed well-separated CCI pairs, indicating that GITIII effectively identified distinct patterns across interactions among different cell types (Fig. 3c).

To investigate how a cell's local microenvironment influences its state, we focused on microglia, for their involvement in neuroinflammation in AD[46,47], and L2/3 IT neurons, a vulnerable neuronal population in AD[48]. Microglial state is correlated with laminar location in the mouse cortex[49]. The UMAP visualization of microglia using CCI-informed microenvironmental features revealed three clusters, each localized within different cortical layers (Fig. 3d, 3e, Extended Data Fig. 2b), consistent with the reported laminar organization of microglia in human ST studies[50,51]. We observed upregulated expression of *CUX2* in subgroup 1 and *NTNG2* in subgroup 3 (adjusted $p < 1 \times 10^{-24}$, Fig. 3f). GITIII further inferred that subgroup 1 was strongly influenced by L2/3 IT neurons, subgroup 2 by L4 IT and L5 IT neurons, and subgroup 3 by L6 IT and L6 IT Car3 neurons (Fig. 3g). These interactions support that intercellular communication primarily occurs between cells that are physically close (Extended Data Fig. 2b). These identified genes, whose expression was influenced by CCI, may impact microglial functions including proliferation, synapse formation[52], accumulation along axons[53,54], and myelin growth and integrity[55], confirming the role of neurons in determining microglial state[56].

GITIII clustered L2/3 IT neurons into four subgroups, with most cells belonging to the first two subgroups (Fig. 3d). Subgroups 1, 3, and 4 located in the upper region of layers 2 and 3, where most of the vulnerable neuronal populations locate[45], whereas subgroup 2 located in the deeper region of layers 2 and 3, close to layer 4 (Fig. 3e). These subgroups had much clearer spatial separation compared to the gene expression-based subgroups (Extended Data Fig. 2c)[45]. Subgroup 1 showed downregulated expression of *RORB* (adjusted $p < 1 \times 10^{-24}$, Fig. 3f), a marker for selectively vulnerable neurons[48], and was mostly influenced by microglia, while subgroup 2 was primarily influenced by L4 IT neurons (Fig. 3g).

We then constructed CCI networks for two AD risk genes *RORB*[48] and *CD74*[57] (Fig. 3i). For *RORB*, its expression in L2/3 IT subgroup 1 was downregulated by interactions with microglia and VIP neurons, and upregulated in subgroup 2 by interactions with L4 IT neurons (Fig. 3h). *CD74*, a microglia maker involved in immune responses[58], was inferred to be significantly upregulated in cell types interacting with microglia, highlighting the influence of inflammatory microglia on neighboring cells[59].

Finally, we quantified the CCI network strength within each slide and observed some separation among the 69 slides from patients with and without dementia (Extended Data Fig. 2d). Dementia patients had stronger interactions between microglia, oligodendrocytes, and excitatory neurons in deeper cortical layers compared to patients without dementia (adjusted $p < 0.1$, Fig. 3j), indicating more pronounced microglia and oligodendrocyte dysfunction in dementia that may contribute to neuronal impairment and cognitive decline.

GITIII interprets tumor microenvironment in NSCLC

To understand the TME and its impact on cell states, we analyzed the human non-small cell lung cancer (NSCLC) ST dataset generated by CosMx[25]. Sample 'Lung 6' was used for demonstration (Fig. 4a). We observed an overall trend of decreasing interaction strength with increasing distance between interacting cells (Fig. 4b). The CCI pairs among the 10 most frequent cell-type interactions showed distinct separation in a UMAP (Fig. 4c).

Using CCI-informed features, we clustered cells that play important roles in the TME, including macrophages, endothelial cells, and tumor cells. GITIII identified macrophage subgroup 3 and endothelial subgroup 4 as strongly influenced by tumor cells (Fig. 4d-4f). Macrophage subgroup 3 showed increased expression of *SPP1* (adjusted $p < 1 \times 10^{-24}$, Supplementary Fig. 3a), suggesting them as tumor-associated macrophages that contribute to tumor progression and

the TME remodeling[25,60,61]. High expression of *EGFR* and several KRT genes was also observed in this subgroup. While these genes are not marker genes of macrophage, one possible explanation is that macrophages phagocytose nearby tumor cells that express *EGFR* and the KRT genes. The presence of these transcripts within macrophages could therefore reflect RNA from tumor cells. Endothelial subgroup 4 showed upregulated expression of *KRT6A*, *KRT17*, *EGFR*, and *KRT5* (adjusted $p < 1 \times 10^{-4}$, Supplementary Fig. 3a), suggesting tumor-induced activation via EGFR signaling to promote proliferation and angiogenesis[62]. Tumor cells were clustered into two subgroups (Fig. 4d), where subgroup 2, located near the tumor boundary (Fig. 4e), interacted with macrophages and structural cells (Fig. 4f) and showed reduced expression of *GPNMB*, *EGFR*, *KRT19*, *CD9*, and *KRT15* (adjusted $p < 1 \times 10^{-24}$, Supplementary Fig. 3a). These genes are associated with cancer proliferation[63], immunosuppression[64], and metastasis[65], highlighting the suppressive effect of immune cells at the tumor boundary.

We then constructed CCI networks for *SPP1* and *EGFR*. Interactions between tumor cells and macrophages were associated with increased expression of *SPP1* (Fig. 4g), consistent with the spatial distribution of macrophage subgroup 3 around tumor cells (Fig. 4e). Interactions between tumor cells and neighboring cell types, such as tumor cells, macrophages, and structural cells, were associated with increased expression of *EGFR* in these cell types, promoting tumor proliferation and metastasis[66].

To further characterize the CCI-related biological functions and processes, we conducted Gene Ontology (GO) enrichment analysis on genes significantly influenced by interactions between tumor cells and macrophages, fibroblasts, and endothelial cells, using ClusterProfiler[67]. Fibroblasts mainly influenced tumor cells by remodeling the extracellular matrix[68] (Fig. 4h), while tumor cells influenced endothelial cells by promoting cytoskeleton organization and fiber growth[69] (Fig. 4j).

These interactions collectively enhanced tumor proliferation. Macrophages modulated tumor cells through pathways related to cell cycle regulation[70], chromosome organization, and the p53 signal pathway[71] (Fig. 4i). In contrast, tumor cells enhanced the role of macrophages in apoptotic processes[72] and influenced their extracellular organelles[73] (Fig. 4k). Together, these interactions revealed complex interactions in the TME and suggest potential therapeutic targets[74].

We also performed enrichment analyses on cell-type-specific genes for tumor cells, macrophages, fibroblasts, and endothelial cells. Compared with pathways enriched in CCI-influenced genes, the pathways enriched in sender cell-type-specific genes were largely distinct (Supplementary Note 6, Supplementary Fig. 3b), suggesting that the CCI-related biological processes we found are unlikely to be artefacts of lateral spillover but instead reflect intercellular communication.

GITIII uncovers CCI-driven within-cell-type heterogeneity in breast cancer

We analyzed a human breast cancer ST dataset generated by Xenium[26] (Fig. 5a). The first replicate slide from sample 1 was used for demonstration. As expected, interaction strength generally decreased with increasing distance between interacting cells (Fig. 5b). The UMAP showed distinct separation among the CCI pairs from the 10 most frequent cell-type interactions (Fig. 5c).

We then examined how immune, tumor, and structural cells in the TME are influenced by their neighborhood cells. CCI-informed microenvironmental features clustered B cells into two subgroups (Fig. 5d) with distinct transcriptional profiles (Fig. 5f). Subgroup 1 showed increased expression of *MS4A1*, *BANK1*, *PTPRC*, and *CD19*, and decreased expression of *TENT5C*, *SLAMF7*, *ITM2C*, and *NZB1* (adjusted $p < 1 \times 10^{-24}$), and was strongly influenced by CD4+ T cells (Fig. 5g). Macrophages were clustered into five subgroups (Fig. 5d). Subgroup 3, mainly

influenced by stromal cells and invasive tumor cells (Fig. 5g), showed increased expression of *FOXA1*, suggesting that tumor-macrophage interactions may modulate immune responses in cancer[75].

For tumor cells, we focused on a group of ductal carcinoma in situ (DCIS_2). These cells were clustered into three subgroups, where subgroup 1 located on the tumor boundary, and subgroups 2 and 3 were inside the tumor. Subgroup 1 showed increased expression of *MYLK*, *ACTA2*, and *KRT14* (adjusted $p < 1 \times 10^{-24}$), suggesting a progressive role for DCIS at the tumor boundary[76].

We also analyzed structural cells, including endothelial and stromal cells. Endothelial subgroup 2, strongly influenced by invasive tumor cells (Fig. 5g), showed upregulated expression of *CCND1*, *KDR*, *KRT7*, and *TCM* (adjusted $p < 1 \times 10^{-24}$), suggesting tumor-driven angiogenic signaling[77]. Stromal subgroup 3, also influenced by invasive tumors, showed increased expression of *KRT7*, *FOXA1*, *FASN*, and *KRT8* (adjusted $p < 1 \times 10^{-24}$), indicating that tumor-resident stromal cells contribute to tumor invasion by promoting the expression of basal epithelial markers[78].

Consistyency of CCI signals across datasets

To demonstrate the robustness of GITIII, we compared the identified interactions between cell types across datasets that differ by species and platform. In the mouse PMC dataset, GITIII identified interactions between L4/5 IT and L2/3 IT neurons that upregulate the *Rorb* expression in L2/3 IT neurons (Fig. 2i, 2j). Similar interactions were observed in the human SEA-AD dataset, where the interactions between L4 IT and L2/3 IT neurons were associated with the upregulation of *RORB* in L2/3 IT neurons (Fig. 3h, 3i). These interaction patterns were found in the L2/3 IT subgroup adjacent to layer 4 in both datasets (Fig. 2e, 3e). In addition, GITIII consistently identified interactions between tumor cells and a subgroup of macrophages, as well as between

tumor cells and a subgroup of endothelial cells, in both the NSCLC and breast cancer datasets (Fig. 4f, 5g), further indicating its cross-dataset generalizability.

Additionally, we analyzed the composition of interacting cell types for each receiver cell type across the four ST datasets. Consistently across datasets, we observed that the gene expression of certain receiver cell types was influenced by multiple sender cell types, with contributions extending beyond interactions with neighboring cells of the same type (Supplementary Fig. 4).

Benchmarking and sensitivity analysis

We benchmarked GITIII against nine existing CCI analysis tools, COMMOT[8], HoloNet[10], NCEM[30], CellChat[22], CellPhoneDB[21], Giotto[9], SOAPy[16], SpaTalk[14], and stLearn[12], as well as two graph deep learning backbone models, using the four ST datasets. Details of the benchmarked methods are provided in Supplementary Note 7. Performance was evaluated at both single-cell and cell-type levels.

At the single-cell level, we compared GITIII with three methods that estimate CCI signals on receiver cells, COMMOT, HoloNet, and NCEM, as well as graph attention networks and graph transformer. Evaluation was based on the intracell-type variance explained by CCI. Fig. 6a shows that GITIII achieved the highest intracell-type variance explained by CCI across all four datasets. NCEM ranked second among existing CCI analysis methods. The improvements of GITIII over NCEM were 20%, 36%, 44%, and 201%, respectively.

At the cell-type level, we conducted benchmarking by comparing CCI networks inferred by GITIII, CellChat, CellPhoneDB, COMMOT, Giotto, SOAPy, SpaTalk, and stLearn. Performance was evaluated using the Spearman correlation coefficient (SCC) between the aggregated CCI network strength and the Wasserstein distance between sender-receiver cell-type pairs. A more negative SCC (i.e., larger absolute value) indicates better performance in capturing spatial CCI

patterns. The two MERFISH datasets[33,45], however, contain only 254 and 140 genes, respectively. Because all methods we compared relied on their built-in ligand-receptor databases, they identified fewer than 10 ligand-receptor pairs in these datasets, and in some cases none at all. Given the limited number of ligand-receptor pairs detected, we excluded these two datasets from the cell-type-level benchmarking. Fig. 6b shows that GITIII achieved the most negative SSC among all methods in both the NSCLC and breast cancer datasets.

To gain insights into the contribution of individual components to the performance of GITIII, we conducted an ablation study using the NSCLC CosMx dataset. The components or information removed or disabled include ligand expression in the sender cells, cell type information, distance scaler, and all modifications to the GRIT[79] encoder on which our model is based. Among the individual ablations (Fig. 6c), the removal of all modifications to the GRIT block had the most impact on model performance, with the intracell-type variance explained by CCI decreased from 0.0136 to 0.00972, highlighting the importance of our modifications to the original GRIT model in encoding CCI. Removing each input information resulted in worse performance, except for cell type information. The intracell-type variance explained by CCI was 0.01361 when using cell type annotations and 0.01366 without them. This negligible difference suggests that the estimation of CCI influence on cell states in GITIII is robust to the absence of cell type annotations. GITIII's architecture with attention calculation and feature embedding can learn cell type information from the input data to encode CCI. However, we acknowledge that the downstream interpretation analysis and visualization rely on cell type labels.

In the analysis of ST data, most GNN-based models employ multiple GNN layers, while GITIII uses a single-layer GRIT block to preserve interpretability and still outperforms other models. Fig. 6d shows that adding extra encoding layers only led to marginal improvement but would reduce

model interpretability. Therefore, we used a single-layer graph transformer encoder to preserve interpretability. Finally, we evaluated the sensitivity of GITIII to the number of cells in the cellular neighborhood. The intracell-type variance explained by CCI increased as the number of neighboring cells increased and plateaued when the size of the cellular neighborhood reached 50 in the NSCLC dataset (Fig. 6e). To reduce computational resources, we set the size of the cellular neighborhood to 50.

To assess the robustness of GITIII to different quality of cell segmentation and marker quantification, we applied it to the breast cancer Xenium data processed by three cell segmentation methods, Voronoi, JSTA[80], and BIDCell[81]. At the single-cell level, the spatial organization, gene expression, and interacting cell types within the local microenvironment of CCI-informed subgroups, for macrophages, fibroblasts, and SCGB2A2+ malignant cells, were broadly consistent across the three datasets (Supplementary Note 8, Supplementary Fig. 5-7). At the cell-type level, the composition of interacting cell types for each receiver cell type was also generally consistent across the three datasets (Supplementary Fig. 8). Together, these results demonstrate that GITIII is robust to differences in cell segmentation.

Discussion

Imaging-based ST provides spatial gene expression patterns at single-cell resolution, enabling a deep understanding of intercellular communication critical for tissue development and organ function. Disruptions in normal CCI are implicated in disease onset and progression, such as neurodegeneration and cancer. We developed GITIII, a state-of-the-art GNN model that identifies single-cell-level interactions in imaging-based ST data. It integrates MLP-based distance encoding with physics-informed attention calculation, outperforming existing methods in capturing the

complex spatial patterns of CCI. GITIII's interpretable architecture allows estimation of the influence of sender cells on the gene expression of receiver cells, facilitating downstream analyses including visualization of spatial CCI patterns, CCI-informed cell clustering, identification of significant cell-type interactions, CCI network analysis, and comparison of CCI strength across conditions. We applied GITIII to four ST datasets across multiple species, organs, and platforms, demonstrating superior performance over existing GNN models in decoding CCI influences. These applications advanced our understanding of brain structure and the TME.

Despite its advantages, GITIII has several limitations. First, it is designed for imaging-based ST platforms that measure only a limited set of genes, often with sparse or missing ligand-receptor pair information. As a result, the downstream signaling pathways representing functional consequences of CCI remain unexplored. To address this, we plan to integrate scRNA-seq and ST data in future work to characterize the functional influence of CCI. Second, the inferred CCI only indicate the association between cell states and their niche, not a causal interpretation. Thus, the biological results generated by our method require further validation through experimental studies. Third, signal spillover between segmented neighboring cells is a known issue in imaging-based ST, often arising from the imperfect cell segmentation. Such artefacts can confound and complicate the identification of cell types and degrade data quality, thereby affecting the inference of cellular interactions.

Methods

Data preprocessing

Consider an imaging-based single-cell ST dataset that includes a cell-by-gene expression matrix, spatial coordinates of cells, and cell type annotations. We first log-transformed the

normalized expression matrix. Then, we computed a distance matrix using the Euclidean distance between cells. To ensure data quality, we excluded cells that had no neighboring cells within a radius of r pixels (default is 80). These cells, along with their corresponding rows in the expression matrix and both rows and columns in the distance matrix, were removed. Lastly, for each cell type, we calculated the cell-type expression by averaging the gene expression levels across all cells of that type. The cell-state expression of each cell was then calculated as the deviation of its gene expression from the average expression of its respectively cell type.

GITIII algorithm

The GITIII pipeline consists of four components: a deep learning model with an embedding module and a modified graph inductive bias transformer to estimate the influences of sender cells on the gene expression of receiver cells; a CCI pattern visualization module; a cell clustering analysis module; and a network analysis module. Fig. 1 shows the schematic diagram of GITIII. For convenience, a list of notations used below is included in Supplementary Note 1.

Deep-learning model architecture

GITIII uses a lightweight, self-supervised, single-layer graph transformer-based language model to predict the cell-state expression of each cell based on the spatial location and gene expression profiles of its neighboring cells, thereby learning how neighboring sender cells of varying types and states interact with a central receiver cell (Fig. 1a). The ST data can be represented as a spatial graph, where nodes are cells and edges are spatial proximity between cells. For each cell, GITIII first constructs a cellular neighborhood by identifying its $(k-1)$ nearest neighboring cells to form a subgraph. This process converts the original spatial graph into N graphs of size k, where N is the total number of cells in the ST data, and each graph represents a cellular neighborhood. Conceptually, these cellular neighborhoods are treated as sentences, with

each cell representing a word. GITIII processes these subgraphs using a deep learning model consisting of two parts: an embedding module (Fig. 1b) and a single-layer modified graph inductive bias transformer encoder (Fig. 1c). The model outputs an influence tensor that quantifies the impact of each neighboring sender cell on the gene expression of the central receiver cell.

Both CCI within the cellular neighborhood and intracellular biological processes, such as cell differentiation, influence cell states. GITIII is not designed to fully explain cell-state variation, but to capture the component attributable to CCI influence from neighboring cells. The model is trained by minimizing the sum of squared differences between the observed and predicted cell-state expression of receiver cells, where the predicted expression is defined as the aggregated CCI influence from neighboring cells. Importantly, the aggregated CCI influence estimated by GITIII is not intended as a full prediction of cell-state expression, but rather as a way to quantify the influence from neighboring cells.

Model input

The input of GITIII includes three types of information from a cellular neighborhood: (1) the cell-by-gene expression matrix, denoted as $\mathrm{Exp} \in \mathbb{R}^{k \times c}$, where c is the number of genes measured in the ST data; (2) the distance matrix, denoted as $D \in \mathbb{R}^{k \times k}$, the Euclidean distance between cells; and (3) the cell type vector, denoted as $T \in \mathbb{R}^{k}$, representing cell type annotations of cells.

Node feature embedding

The node feature embedding is generated in three steps: expression masking, expression embedding, and cell type embedding. Since GITIII infers CCI by predicting the cell-state expression of a central cell based on the information of its neighboring cells, and cells of the same type in close proximity often share CCI patterns, thus having similar cell-state expression (the closer the more similar), we mask the cell-state expression of neighboring cells of the same type

as the central cell to prevent perfect prediction. Specifically, within a cellular neighborhood of k cells, where the first row in the cell-by-gene expression matrix, $\mathrm{Exp} \in \mathbb{R}^{k \times c}$, corresponds to the gene expression profile of the central cell c_1. For each neighboring cell c_i, if it is of the same cell type as the central cell, its gene expression is replaced by its cell-type expression; otherwise, its observed gene expression is retained. Formally, it is defined as

$$M(\boldsymbol{y}_i) = \begin{cases} \overline{\boldsymbol{y}}_i & \text{if } c_i \text{ and } c_1 \text{ are of the same cell type} \\ \boldsymbol{y}_i & \text{otherwise} \end{cases},$$

where $\boldsymbol{y}_i$ is the gene expression of c_i, and $\overline{\boldsymbol{y}}_i$ denotes the cell-type expression of c_i, the average gene expression of all cells of the same type as c_i. Then, we apply a multilayer perceptron (MLP) to encode the expression features, producing the expression embedding. Finally, we incorporate cell type information by adding cell type embeddings to the node features. The node embedding $X \in \mathbb{R}^{k \times D_V}$ is defined as

$$X = \mathrm{MLP}\big(M(\mathrm{Exp})\big) + \mathrm{Embed}(T),$$

where D_V is the dimension of the node embedding and Embed uses an embedding layer that maps discrete cell type indices to continuous vectors of fixed dimension, analogous to token embeddings commonly used in natural language processing.

Distance embeddings

Given that the interaction between two cells is influenced by their distance, we model the effect of distance on CCI strength using five distance-based features, which differ from the commonly used exponential decay diffusion model, to better describe dynamic processes in biological systems. We then apply MLPs to these features. Specifically, we generate two distance embeddings to be used in edge embeddings and one to be used in the graph transformer encoder. Of these three distance embeddings, two are scaled to a range of 0 to 1 using the sigmoid function.

The design and rationale behind these distance embeddings are explained in detail in Supplementary Note 2.

Edge feature embedding

Edge features are constructed to capture the distance between interacting cells and the effects of ligands from the sender cell on the receiver cell. The ligand effect is modeled as the ligand expression in the sender cell, scaled by one of the distance embeddings. Briefly, we first multiply the ligand expression by the scaled distance embedding and apply a linear transformation. Then, we incorporate the spatial organization by concatenating the resulting tensor with an unscaled distance embedding and apply another linear transformation to produce the final edge embedding, denoted as $E \in \mathbb{R}^{k \times k \times D_E}$, where D_E is the dimension of the edge embedding. Further details on the edge embedding are provided in Supplementary Note 3.

Graph attention calculation

We develop a graph attention method based on the Graph Inductive bias Transformer (GRIT)[79] to capture the complexity of CCI. GRIT is a state-of-the-art graph transformer that does not rely on message-passing. Instead, it incorporates inductive biases through learned relative positional encodings based on random walk probabilities. GRIT encodes node and edge features by integrating position information, node features, and edge features using attention mechanisms. This architecture is particularly well suited for modeling CCI, where gene expression of cells (node features), spatial organization (positional information), and ligand expression in sender cells (edge features) interact with each other and collectively influence the gene expression of receiver cells. To adapt GRIT for encoding cellular neighborhood information in ST, we implemented three main modifications in GITIII: (a) we replaced the random walk graph structure encoding designed for sparse graphs with distance embeddings to better model the dense, two-dimensional spatial

organization of cellular neighborhoods; (b) we incorporated a distance scaler into the attention mechanism to explicitly model the decay of CCI effects with increasing intercellular distance; and (c) we removed the degree scaler, which adjusts for variation in node connectivity, as the cellular neighborhoods in our setting have uniform connectivity by construction. These changes tailor the GRIT framework to the unique spatial and biological characteristics of single-cell ST data and improve its capacity to capture biologically meaningful CCI. In addition, we refine the attention calculation by computing scores only from neighboring sender cells to a central receiver cell and performing a single pass of the attention calculation. This ensures that the aggregated values are solely determined by the features of individual cells, avoiding multi-step contributions from indirect sources. Following is a summary of the key ideas in our graph attention calculation, with a detailed description of the algorithm provided in Supplementary Note 4.

We first construct the Key, Query, and Value matrices using three linear transformations of the node features. The attention matrix is then computed by multiplying the Key and Query matrices to capture how signaling pathway genes, cell type, and other regulatory genes influence the CCI effects. Next, we update the edge features by integrating the edge embeddings with the unscaled distance embeddings and the attention matrix, thereby encoding distance, gene expression, and cell type characteristics relevant to CCI. A linear transformation is then applied to the updated edge features to generate the initial raw attention scores. We discard the self-attention score of the receiver cell and apply a softmax function for normalization. Finally, we update the Value matrix by integrating it with the edge features whose destination is the receiver cell using an MLP. This operation models interactions between edge and node features to enhance the graph representation and encoding capacity.

The Value matrix $V = V[1:,:] \in \mathbb{R}^{(k-1)\times D_E}$ (subscripts starting from 0) represents the attention values of all neighboring cells, where the attention scores from the neighboring sender cells to the receiver cell are denoted as $\alpha \in \mathbb{R}^{1\times(k-1)}$. We compute the influence matrix $i \in \mathbb{R}^{(k-1)\times c}$ by multiplying the Value matrix with the attention scores and applying an MLP to transform the last dimension from D_V to the number of measured genes c. Since each row of the influence matrix is only determined by a cell's type, gene expression features, and its distance to the receiver cell, we interpret it as representing the influence of the $(k-1)$ neighboring cells on the expression levels of c genes in the central cell. Summing over the first dimension of i yields the predicted cell-state expression $\hat{y} \in \mathbb{R}^c$, which is then compared to the observed cell-state expression to infer how neighboring sender cells influence the central receiver cell.

Interpretation of the influence tensor

We construct the influence tensor, denoted as $I \in \mathbb{R}^{N\times(k-1)\times c}$, by concatenating the influence matrices learned from each subgraph representing a cellular neighborhood. Since all non-attention-based transformations in GITIII are performed along the last dimension of the input tensor, and the attention mechanism preserves both the order and representation of each token (cell), the identity of each cell remains intact throughout the encoding process. Thus, the entries in the influence tensor can be interpreted as the impact of the $(k-1)$ neighboring cells on the expression levels of c genes in the central cell. Instead of using attention scores alone to quantify the influence of CCI on receiver-cell expression, the entries in the influence tensor, calculated as the product of the embedding vector and the corresponding attention score, represent the contributions of neighboring cells to the cell-state expression of the central receiver cell. The justification for using the influence tensor as a measure of CCI influence is presented in Supplementary Note 5.

Normalization and aggregation of the influence tensor

To compare the influence of neighboring cells on the gene expression of a central cell within a cellular neighborhood, we calculate a proportional influence tensor, denoted as $I_p \in \mathbb{R}^{N\times(k-1)\times c}$, by normalizing the influence tensor I. Specifically, for each gene in a given cell, we divide the influence values from its neighboring cells by the sum of the absolute values of the influences on that gene across all neighboring cells. If the absolute values of the proportional influence are below a threshold (default is 0.02), we discard these subtle influences by setting the corresponding entries of I to 0, and then recalculate I_p as described above. The default threshold of 0.02 was chosen to ensure that removing weak interactions has minimal impact on the intracell-type variance explained by CCI. Across all four datasets analyzed, applying this threshold led to a reduction of less than 5% in the intracell-type variance explained, supporting its use as a conservative and robust cutoff. Additionally, we compute the average of the absolute proportional influence values across all genes to estimate the overall strength of CCI between two interacting cells. The resulting matrix, denoted as $I_s \in \mathbb{R}^{N\times(k-1)}$, serves as a metric for identifying significant CCI pairs in downstream analysis.

To estimate the effect of different sender cell types on the gene expression of receiver cells, we aggregate the influence values by sender cell type within each cellular neighborhood, resulting in an aggregated influence tensor, $\mathbb{I} \in \mathbb{R}^{N\times t\times c}$, where t is the number of cell types in the ST data, and the entries represent the impact of neighboring cell types on the expression levels of c genes in the central cell. If specific cell types are absent among the neighboring cells in a cellular neighborhood, the corresponding entries are set to 0. Similarly, we calculate a proportional aggregated influence tensor, $\mathbb{I}_p \in \mathbb{R}^{N\times t\times c}$, by normalizing the aggregated influence tensor. For each gene in a given cell, we normalize the aggregated influence values by the sum of the absolute values of the aggregated influences on that gene across all neighboring cell types.

Visualization of CCI patterns

The CCI pattern visualization module in GITIII offers three types of visualizations to understand the spatial patterns and heterogeneity of CCI, including (1) visualization of the relationship between CCI strength and the distance between interacting cells, (2) visualization of CCI pairs, and (3) visualization of the spatial distribution of CCI.

To visualize the relationship between CCI strength and the distance between interacting cells, we flatten the influence tensor I and plot the absolute value of each entry against the distance between the corresponding sender and receiver cells. A locally estimated scatterplot smoothing (LOESS) curve is fitted to demonstrate the decay of CCI strength with increasing distance between interacting cells. For the visualization of CCI pairs, we treat each cell on the slide as a receiver cell and select the top five sender cells from its neighborhood with the strongest interactions based on the overall strength values in I_s to form CCI pairs. These selected CCI pairs are visualized on a UMAP using the estimated influence values on the cell-state expression of the corresponding receiver cell as features, where each triangle represents a CCI pair. Finally, to visualize the spatial distribution of CCI influencing the expression level of a specific gene, we focus on the top 5% of CCI pairs with the largest absolute influence values for that gene among the selected CCI pairs across all cellular neighborhoods. The arrow connecting the corresponding sender and receiver cells is drawn on a spatial map, with the color of each arrow representing the influence value of the interaction.

Cell clustering analysis

The cell clustering analysis module in GITIII provides a quantitative interpretation of the local microenvironment and its impact on cell-state expression, including (1) clustering cells into subgroups based on CCI-informed microenvironmental features, (2) spatial visualization of cell

subgroups, (3) differential expression analysis across subgroups, and (4) investigation of microenvironmental influence on cell-state expression.

To perform cell clustering within each cell type, we first construct CCI-informed microenvironmental features by flattening the last two dimensions of the aggregated influence tensor $\mathbb{I}$ into a feature matrix $F \in \mathbb{R}^{N\times(tc)}$. We then apply the Leiden algorithm[82] to cluster cells into subgroups based on these microenvironmental features and visualize cell subgroups using UMAP. The distinction among these subgroups reflects differences in their surrounding cellular neighborhood, rather than differences in gene expression alone. We expect that CCI-informed subgroups mainly differ in their local microenvironment and tend to locate in spatially distinct regions.

For each cell type, we analyze the differences among its subgroups by examining their spatial organization, gene expression, and interacting cell types within the local microenvironment. Specifically, we visualize the spatial locations of subgroups on a spatial map. Next, we perform differential expression analysis to identify the top marker genes for each subgroup using the Wilcoxon rank sum test. To assess the influence of neighboring cell types on the expression level of a specific gene in receiver cells of different subgroups, we visualize the aggregated influence values of the gene of interest in a heatmap. Additionally, we compute the average of the absolute proportional aggregated influence values across all genes and visualize them in a heatmap to better understand the overall influence of neighboring cell types on receiver cells of different subgroups. These analyses generate a comprehensive view of the contributions of interacting cell types within the local microenvironment that influence cell-state expression across subgroups within each cell type.

Construction of CCI network

If there is a significant interaction between two cell types that influences the gene expression of a receiver cell type, we expect that the proportional influence value is correlated with the cell-state expression of that gene among CCI pairs of these two cell types. To identify such significant interactions, for a fixed gene l and a pair of cell types, say sender cell type A and receiver cell type B, we assess the correlation between the proportional aggregated influence values of neighboring cell type A on gene l in receiver cells of type B and the cell-state expression of gene l in these receiver cells, using the Pearson correlation coefficient. Pearson correlation analysis is performed across all genes and all pairs of cell types, with the corresponding z-scores calculated. To account for multiple hypothesis testing, significance is adjusted using the Benjamini-Hochberg false discovery rate (FDR) control. Sender cell type A is considered to significantly influence the expression level of gene l in receiver cell type B if the FDR-adjusted p-value < 0.05. For interactions that do not reach significance, we set their z-scores to 0. For each gene, these significant interactions between cell-type pairs are used to construct a CCI network that illustrates how cell types influence the expression of that gene.

When multiple slides are available, we aggregate the individual CCI networks across slides using Stouffer's method to compute the combined z-scores and corresponding p-values.

To evaluate the composition of interacting cell types for each receiver cell type, we first calculate the interaction strength between each sender-receiver cell-type pair, defined as the average of the absolute proportional aggregated influence values across all genes and all receiver cells of the same type. We then normalize these interaction strength by the sum of the interaction strength across all sender cell types, illustrating the relative contributions of different interacting cell types for each receiver cell type.

Comparison of CCI network strength between conditions

To quantify CCI network strength, we introduce a microenvironment-adjusted network strength metric to facilitate comparisons across slides. This metric is less susceptible to batch effects compared to the traditional network strength, which is commonly defined as the product of the mean ligand and receptor expression levels[1]. For a fixed gene l and a pair of cell types, say sender cell type A and receiver cell type B, we consider the proportional aggregated influence value for each receiver cell of type B from the corresponding entries in $\mathbb{I}_p$. To adjust for the local microenvironment, we subtract $n_A/(k-1)$, where n_A is the number of neighboring cells of type A. This adjustment accounts for the cell-type composition of the cellular neighborhood. We then aggregate these microenvironment-adjusted proportional aggregated influence values by receiver cell type within each slide to obtain the CCI network strength for that sample. To assess the group-level CCI network strength, we compute the average CCI network strength across all samples within the group. To compare CCI network strength between conditions, we use the Wilcoxon rank sum test, with significance levels adjusted using the FDR to account for multiple hypothesis testing. Interacting cell-type pairs are considered to have significantly different CCI network strengths between conditions if FDR < 0.1.

Summary of the advantages of GITIII

The main advantages of GITIII over existing GNN-based CCI prediction methods are summarized as:

(1) GITIII is designed to estimate the effects of neighboring sender cells on the gene expression of receiver cells through cell-cell interactions, which is a key aspect of CCI inference that has not been explicitly studied by other GNN-based methods.

(2) While existing GNN-based methods rely on the availability of ligand-receptor pairs measured in ST data, GITIII can be applied to imaging-based ST datasets, where only a limited number of genes are measured and the covered ligand-receptor pairs are sparse or absent.

(3) Most existing GNN-based CCI analysis methods consider 2D ST data as sparse graphs, potentially losing interaction information due to limited pairwise distances between cells. In contrast, GITIII models ST data as dense graphs and uses an expressive graph transformer to better encode spatial CCI patterns. Our benchmarking results show that GITIII consistently outperformed other GNN-based methods in explaining the intracell-type variance.

(4) GNN-based CCI prediction models use multiple layers, with each layer repeatedly propagating information from neighboring nodes to a central node. This multi-layer propagation makes it challenging to explicitly quantify the influence of individual neighboring cells on a central receiver cell. GITIII employs a single-layer graph transformer to directly calculate the influence from neighboring sender cells to the central receiver cell. To preserve the interpretability of the influence, GITIII does not apply any nonlinear transformation after aggregation of neighborhood information.

(5) GITIII incorporates statistical analysis to construct CCI networks at the cell-type level and identify sender cell types that significantly influence the gene expression in receiver cell types. This CCI network inference has not been studied by other GNN-based CCI prediction methods.

Implementation details

Data preprocessing was performed using Scanpy[83], and the deep learning model was implemented in Pytorch 2.0.1. To balance model performance and computational efficiency, we selected a modest set of hyperparameters without the need for extensive GPU resources. Across

all four ST datasets, the number of cells in a cellular neighborhood was set to 50 and the embedding dimension of the node features was fixed at 256. The dimensionality of the edge features was dertermined by the number of ligand genes measured in each dataset, resulting in the edge dimensions of 48 for the mouse PMC and SEA-AD datasets and 128 for the NSCLC and breast cancer datasets. We used the AdamW optimizer with a learning rate of 0.0001. A batch size of 256 was used, which required approximately 8 GB of GPU memory, for efficient model training on moderately powered hardware.

For all experiments, each dataset was randomly split into 80% training and 20% validation sets using a fixed random seed. Models were trained for 50 epochs, and the one with the smallest mean squared error on the validation set was selected for downstream analyses.

It took approximately one hour for GITIII to analyze a dataset containing 366,272 cells and 140 genes on a system with an NVIDIA A5000 GPU (24 GB VRAM) and four Intel Xeon CPUs with 100 GB of RAM. The runtime scales linearly with both the number of cells and the number of genes in the dataset.

Variance decomposition into inter- and intra-cell-type variance

For each gene j, the total variance of its expression can be decomposed into intercell-type variance and intracell-type variance:

$$Var(Y_j) = \underbrace{E\left[Var\left(Y_j \mid C\right)\right]}_{intracell-type\ variance} + \underbrace{Var\left[E\left(Y_j \mid C\right)\right]}_{intercell-type\ variance},$$

where C is cell type. The first term represents the variation of gene expression within cell types (intracell-type variance), corresponding to differences in cell state, and the second term represents the variation of gene expression across cell types (intercell-type variance). A gene with a higher proportion of intracell-type variance relative to the total variance $Var(Y_j)$ is more strongly associated with cell state, whereas one with a higher proportion of intercell-type variance is more

strongly associated with cell type. It is important to note that a high proportion of cell-state variation does not necessarily imply that a gene's expression is influenced by CCI. Cell-state expression can also be influenced by intracellular biological processes, which are not captured by our model.

Benchmarking metrics

Single-cell-level evaluation

Since CCI influence cell states and contribute to intracell-type variability, we use the intracell-type variance explained by CCI to evaluate method performance at the single-cell resolution. Specifically, we use the following metric:

$$\mathrm{EV} = \frac{1}{N}\sum_{i=1}^{N}(\|\boldsymbol{y}_i - \overline{\boldsymbol{y}}_i\|^2 - \|\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i\|^2),$$

where $\boldsymbol{y}_i$ is the observed gene expression of cell i, $\overline{\boldsymbol{y}}_i$ is the cell-type expression of cell i, the average gene expression of all cells of the same type as cell i, and $\hat{\boldsymbol{y}}_i$ is the model-predicted expression. This metric, also used in NCEM[30], captures the improvement in predictive accuracy by modeling CCI over simply using the cell-type average expression, and quantifies how much variability in cell states is explained by CCI effects.

Cell-type-level evaluation

We evaluate method performance by comparing the inferred CCI nextworks at the cell-type level. The evaluation metrics based on significant ligand-receptor pairs are not applicable to GITIII, because it neither uses the receptor information nor identifies significant ligand-receptor interactions. Instead, we evaluate performance by computing the Spearman correlation coefficient (SCC) between the aggregated CCI network strength and the Wasserstein distance between sender-receiver cell-type pairs. It is commonly assumed that spatially adjacent cell types are more likely

to interact than distant ones[4,18]. A more negative SCC (i.e., larger absolute value) indicates better performance in capturing spatial CCI patterns.

Evaluation of cell segmentation on GITIII's performance

To assess the robustness of GITIII to differences in cell segmentation, we considered three methods: (1) classical Voronoi expansion, (2) JSTA[80], a joint cell segmentation and cell type annotation approach based on an extension of the watershed algorithm, and (3) BIDCell[81], a biologically-informed deep learning approach. After applying cell segmentations to the breast cancer Xenium dataset, cell type annotations were assigned using scClassify[84], with detailed processing steps described previously[81]. Because the number and labels of cell types were consistent across the three segmentation datasets, but differed from those in the original dataset[26], we focused our evaluation on the consistency of CCI results inferred by GITIII across the three datasets. In the first replicate slide from sample 1, Voronoi, JSTA and BIDCell produced 106,227, 107,131, and 103,209 cells, respectively.

Datasets

The mouse primary motor cortex (PMC) ST dataset generated by MERFISH[33] contains 254 genes measured in 284,098 cells across 64 slices from two mice, with annotations of 24 cell types. We analyzed slice 201 from mouse 1 with 6,137 cells. Among the 254 genes, 47 were identified as ligand genes and no ligand-receptor pairs were found in the ligand-receptor databases CellChatDB[22] and NeuronChatDB[6].

The SEA-AD ST dataset generated by MERFISH[45] contains 140 genes measured in 366,272 cells across 69 slides of the human middle temporal gyrus (MTG) from 27 donors. Sample 'H20.33.001.CX28.MTG.02.007.1.02.03' with 15,225 cells from 24 cell types was used for demonstration. We also analyzed all 69 slides and aggregated CCI networks for group comparison.

Among the 140 genes, 28 were identified as ligand genes and 2 ligand-receptor pairs were found in CellChatDB[22] and NeuronChatDB[6].

The human non-small cell lung cancer (NSCLC) ST dataset generated by CosMx[25] includes 980 genes measured in 769,114 cells across eight formalin-fixed paraffin-embedded (FFPE) slides from five NSCLC tumors, with annotations of 18 cell types. We analyzed sample 'Lung 6' with 89,949 cells.

The human breast cancer ST dataset generated by Xenium[26] contains 313 genes measured across three slides. Two replicates of 5 μm sections from sample 1, which include 274,492 cells from 19 annotated cell types, were analyzed. Among the 313 genes, 33 were identified as ligand genes.

Data availability

The mouse PMC MERFISH dataset is available at ftp://download.brainimagelibrary.org:8811/02/26/02265ddb0dae51de/. The SEA-AD MERFISH dataset is available at https://sea-ad-spatial-transcriptomics.s3.amazonaws.com/middle-temporal-gyrus/all_donors-h5ad/SEAAD_MTG_MERFISH.2024-02-13.h5ad. The NSCLC CosMx dataset is available at https://nanostring.com/products/cosmx-spatial-molecular-imager/ffpe-dataset/nsclc-ffpe-dataset/. The breast cancer Xenium dataset is available at https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast. The ligand-receptor databases of CellChat and NeuronChat are publicly available, and an integrated version is available on our GitHub page.

Code availability

GITIII is implemented in open-source Python using PyTorch. The source code and a detailed tutorial are freely available at https://github.com/lugia-xiao/GITIII. The Code to reproduce the results in the manuscript are available at https://github.com/lugia-xiao/GITIII_reproducible.

Acknowledgments

This study was supported by the National Institutes of Health (NIH) grant R01LM014087. We are grateful to Drs. Jean Y.H. Yang and Yingxin Lin for generously sharing their data generated using different cell segmentation methods.

Author contributions

X.X. and Z.W. conceptualized the study. X.X. designed and implemented the model algorithm, analyzed the spatial transcriptomics data, and created the figures. L.Z., H.Z., and Z.W. helped refine the idea and analyses and aided result interpretation. X.X. and Z.W. wrote the manuscript. Z.W. supervised the project. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Figure legends

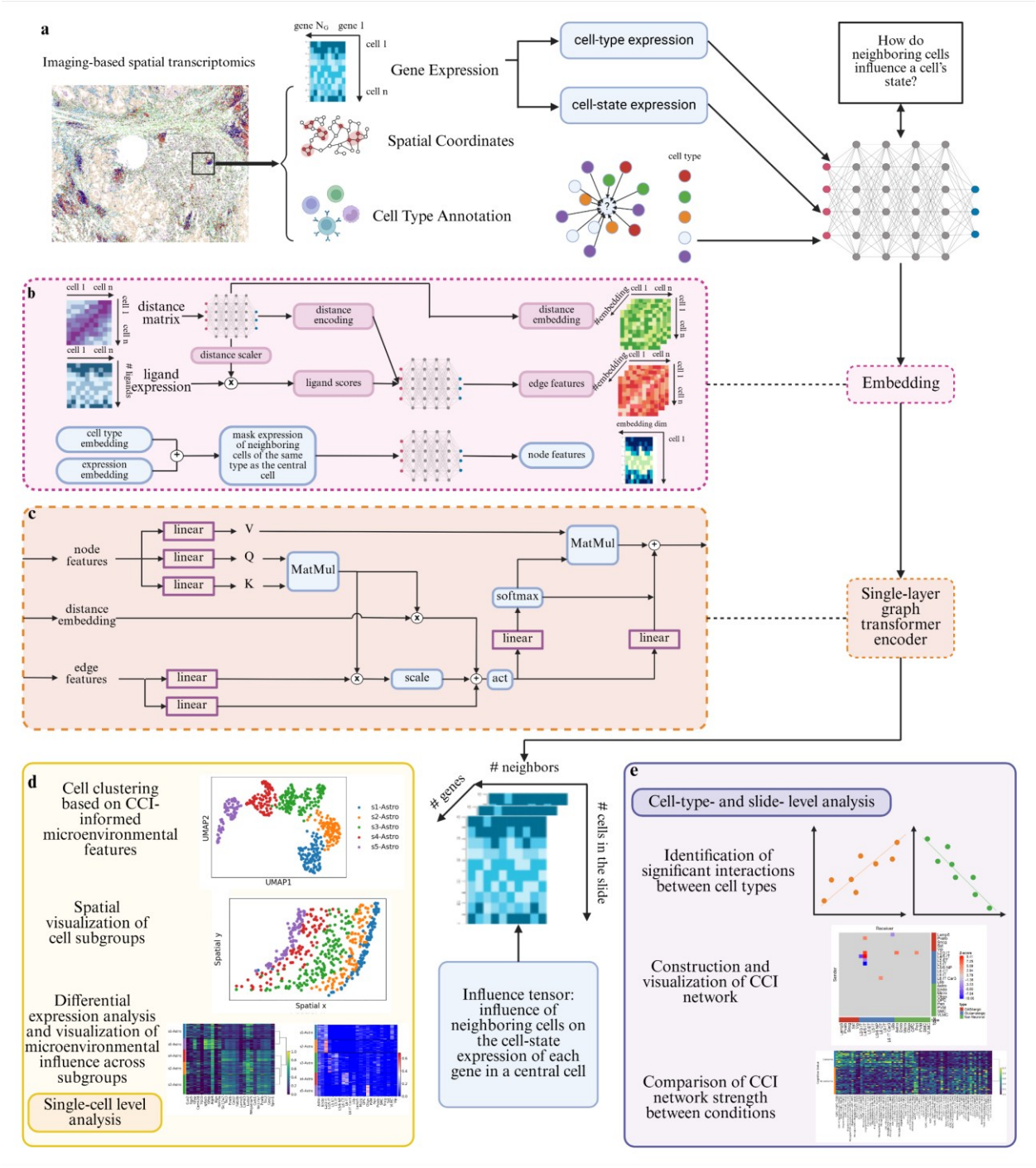


Fig. 1 Schematic diagram of GITIII. a, GITIII Workflow. The model employs self-supervised learning to infer cell-cell interactions (CCI) by predicting the cell-state expression of the central cell from the spatial organization, ligand expression, and the gene expressions of its neighboring cells. We then estimate the influence of neighboring cells on the cell-state expression of the central

cell, and perform downstream analyses. **b,** Detailed view of the input data embeddings. Node features incorporate cell type embeddings and expression encodings, and edge features contain ligand gene expression scaled by distance and the embedding of distance between cells. **c,** Detailed view of the encoder layer. A graph attention mechanism designed to better capture the CCI patterns in the cellular neighborhood. Edge features are updated through linear transformations combined with the attention matrix from the node features and the embedding of distance relationships. Node features are aggregated through attention mechanisms, augmented by edge features weighted by attention scores. **d,** Single-cell level analysis. Clustering cells into subgroups based on CCI-informed microenvironmental features and spatial visualization. Differential expression analysis to interpret differences between subgroups. **e,** Cell-type level analysis. CCI network analysis to identify significant interactions between cell types. Comparison of CCI network strength between conditions.

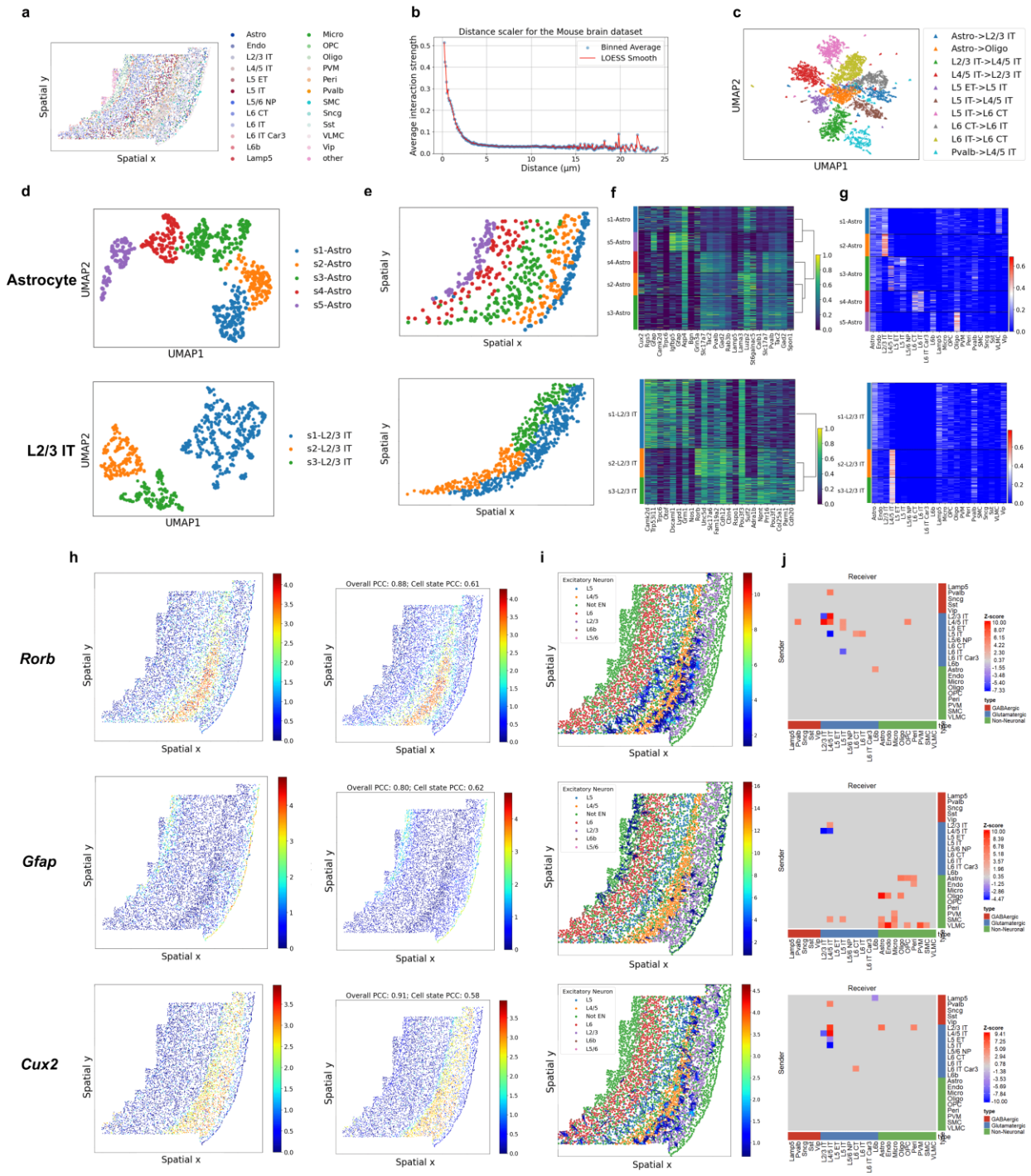


Fig. 2 Spatial patterns of CCI in the mouse PMC dataset. a, Spatial distribution of 24 cell types. **b,** The LOESS curve of the estimated CCI strength decreased with increasing distance between interacting cells. **c,** UMAP visualization of CCI pairs. **d,** UMAP visualization of astrocyte (upper) and L2/3 IT neuron (lower) subgroups. **e,** Spatial distribution of astrocyte (upper) and L2/3 IT neuron (lower) subgroups. **f,** Heatmap of the top differentially expressed genes in astrocyte (upper) and L2/3 IT neuron (lower) subgroups. **g,** Neighboring cell type influence on astrocyte (upper) and L2/3 IT neuron (lower) subgroups. **h,** Visualization of the observed (left) and predicted

(right) expression levels of *Rorb*, *Gfap*, and *Cux2*. **i,** Spatial distribution of CCI**,** showing the top 5% strongest CCI pairs influencing the expression of *Rorb*, *Gfap*, and *Cux2*, with arrows connecting the corresponding sender and receiver cells, and color representing the influence value of the interaction. In the top panel, many CCI pairs between L4/5 IT and L2/3 IT cells, particularly with L2/3 IT subtype 2, as well as CCI pairs within layer 5, influence the expression of *Rorb*. In the middle panel, many CCI pairs, involving non-neuronal cells located in layer 1 and the white matter, influence the expression of *Gfab*. In the bottom panel, CCI pairs, located in layers 2 and 3, influence the expression of *Cux2*. **j,** CCI networks influencing the expression of *Rorb*, *Gfap*, and *Cux2*, with color representing the z-score of the Pearson correlation of interactions between cell types that influence the gene expression of a receiver cell type, positive values indicating upregulation and negative values indicating downregulation.

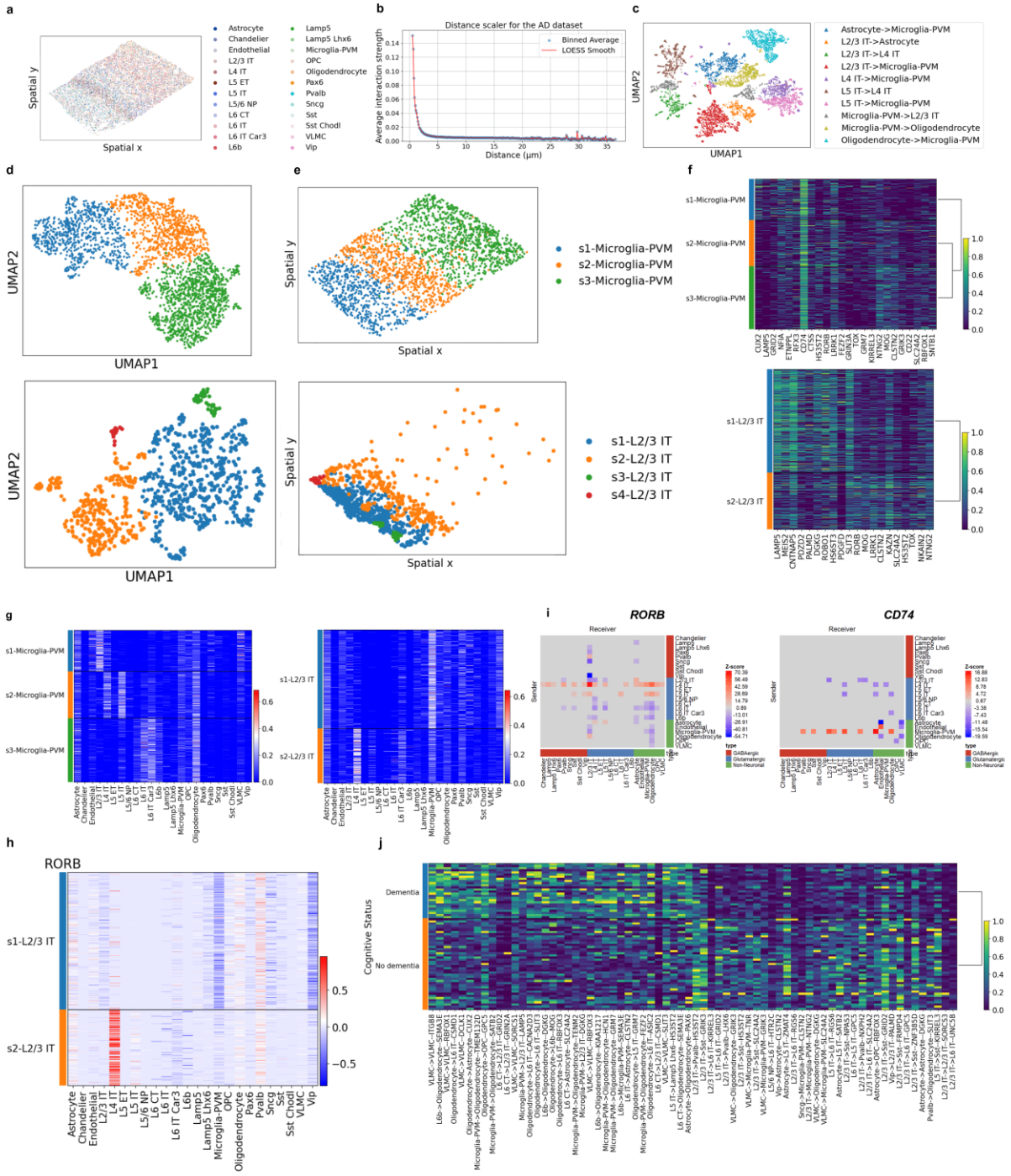
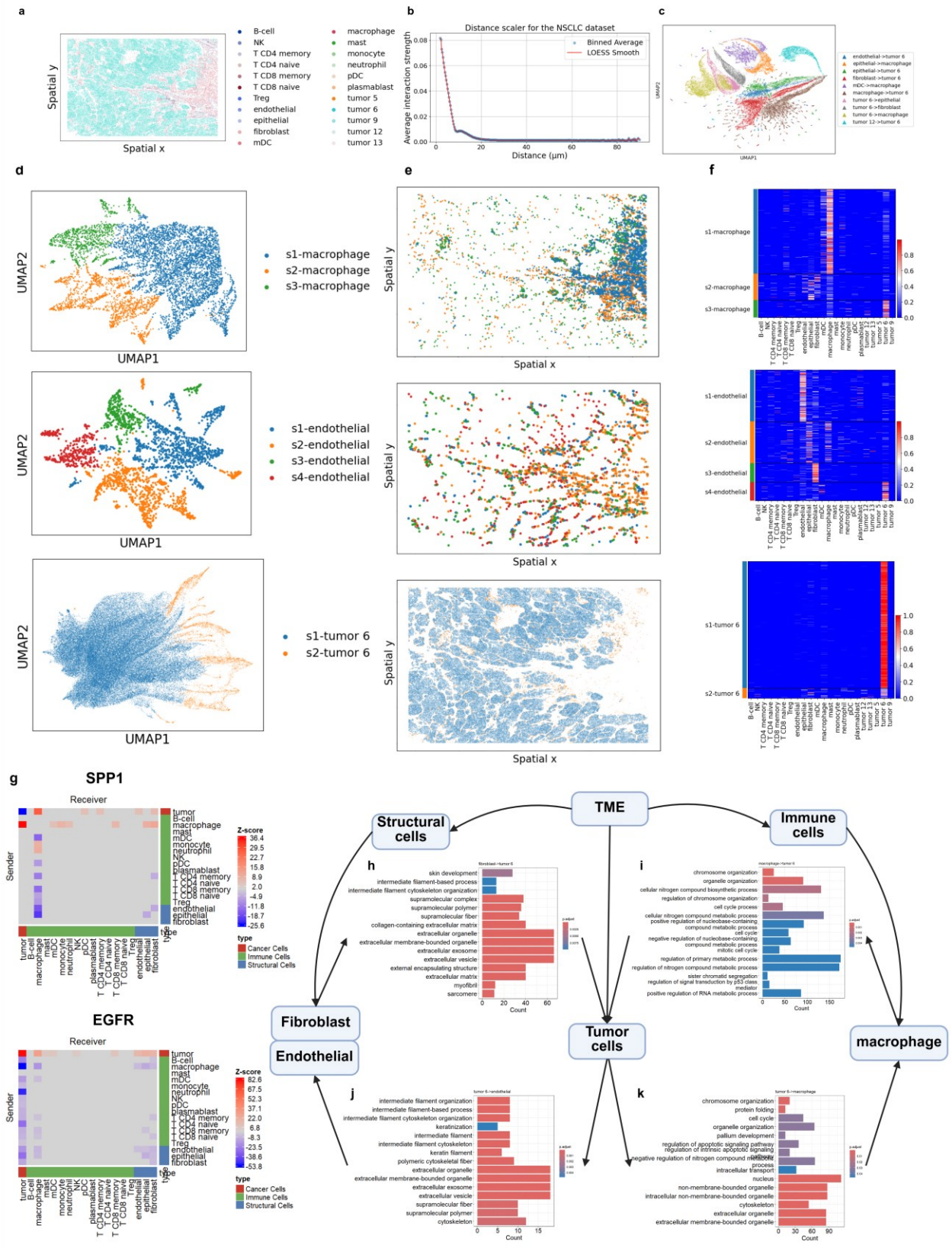


Fig. 3 CCI patterns within and across patients in the SEA-AD dataset. a, Spatial distribution of 24 cell types. **b,** The LOESS curve of the estimated CCI strength decreased with increasing distance between interacting cells. **c,** UMAP visualization of CCI pairs. **d,** UMAP visualization of microglia (upper) and L2/3 IT neuron (lower) subgroups. **e,** Spatial distribution of microglia (upper) and L2/3 IT neuron (lower) subgroups. **f,** Heatmap of the top differentially expressed genes in microglia (upper) and L2/3 IT neuron (lower) subgroups. **g,** Neighboring cell type influence on microglia (left) and L2/3 IT (right) subgroups. **h,** Neighboring cell type influence on the expression of *RORB* in L2/3 IT subgroups, with positive values indicating upregulation and negative values

indicating downregulation. **i,** CCI networks influencing the expression of *RORB* and *CD74*, with color representing the z-score of the Pearson correlation of interactions between cell types that influence the gene expression of a receiver cell type, positive values indicating upregulation and negative values indicating downregulation. **j,** Heatmap of the top differential CCI network strength between patients with or without dementia.


a
Spatial y
Spatial x
B-cell
NK
T CD4 memory
T CD4 naive
T CD8 memory
T CD8 naive
Treg
endothelial
epithelial
fibroblast
mDC
macrophage
mast
monocyte
neutrophil
pDC
plasmablast
tumor 5
tumor 6
tumor 9
tumor 12
tumor 13
b
Distance scaler for the NSCLC dataset
Binned Average
LOESS Smooth
Average interaction strength
Distance (μm)
c
UMAP2
UMAP1
endothelial->tumor 6
epithelial->macrophage
epithelial->tumor 6
fibroblast->tumor 6
mDC->macrophage
macrophage->tumor 6
tumor 6->epithelial
tumor 6->fibroblast
tumor 6->macrophage
tumor 12->tumor 6
d
UMAP2
UMAP1
s1-macrophage
s2-macrophage
s3-macrophage
s1-endothelial
s2-endothelial
s3-endothelial
s4-endothelial
s1-tumor 6
s2-tumor 6
e
Spatial y
Spatial x
f
s1-macrophage
s2-macrophage
s3-macrophage
s1-endothelial
s2-endothelial
s3-endothelial
s4-endothelial
s1-tumor 6
s2-tumor 6
g
SPP1
EGFR
Receiver
Sender
Z-score
type
Cancer Cells
Immune Cells
Structural Cells
TME
Structural cells
Immune cells
Fibroblast
Endothelial
Tumor cells
macrophage
h
i
j
k
Count

1

Fig. 4 Quantitative analysis of tumor microenvironment in the NSCLC dataset. a, Spatial distribution of 18 cell types. **b,** The LOESS curve of the estimated CCI strength decreased with increasing distance between interacting cells. **c,** UMAP visualization of CCI pairs. **d,** UMAP visualization of macrophage (upper), endothelial cell (middle), and tumor cell (lower) subgroups. **e,** Spatial distribution of macrophages (upper), endothelial cells (middle), and tumor cells (lower) subgroups. **f,** Neighboring cell type influence on macrophages (upper), endothelial cells (middle), and tumor cells (lower) subgroups. **g,** CCI networks influencing the expression of *SPP1* and *EGFR*, with color representing the z-score of the Pearson correlation of interactions between cell types that influence the gene expression of a receiver cell type, positive values indicating upregulation and negative values indicating downregulation. **h-k,** GO enrichment analysis of genes whose expression are significantly influenced by CCI **(h)** from fibroblasts to tumor cells, **(i)** from macrophages to tumor cells, **(j)** from tumor cells to endothelial cells, and **(k)** from tumor cells to macrophages.

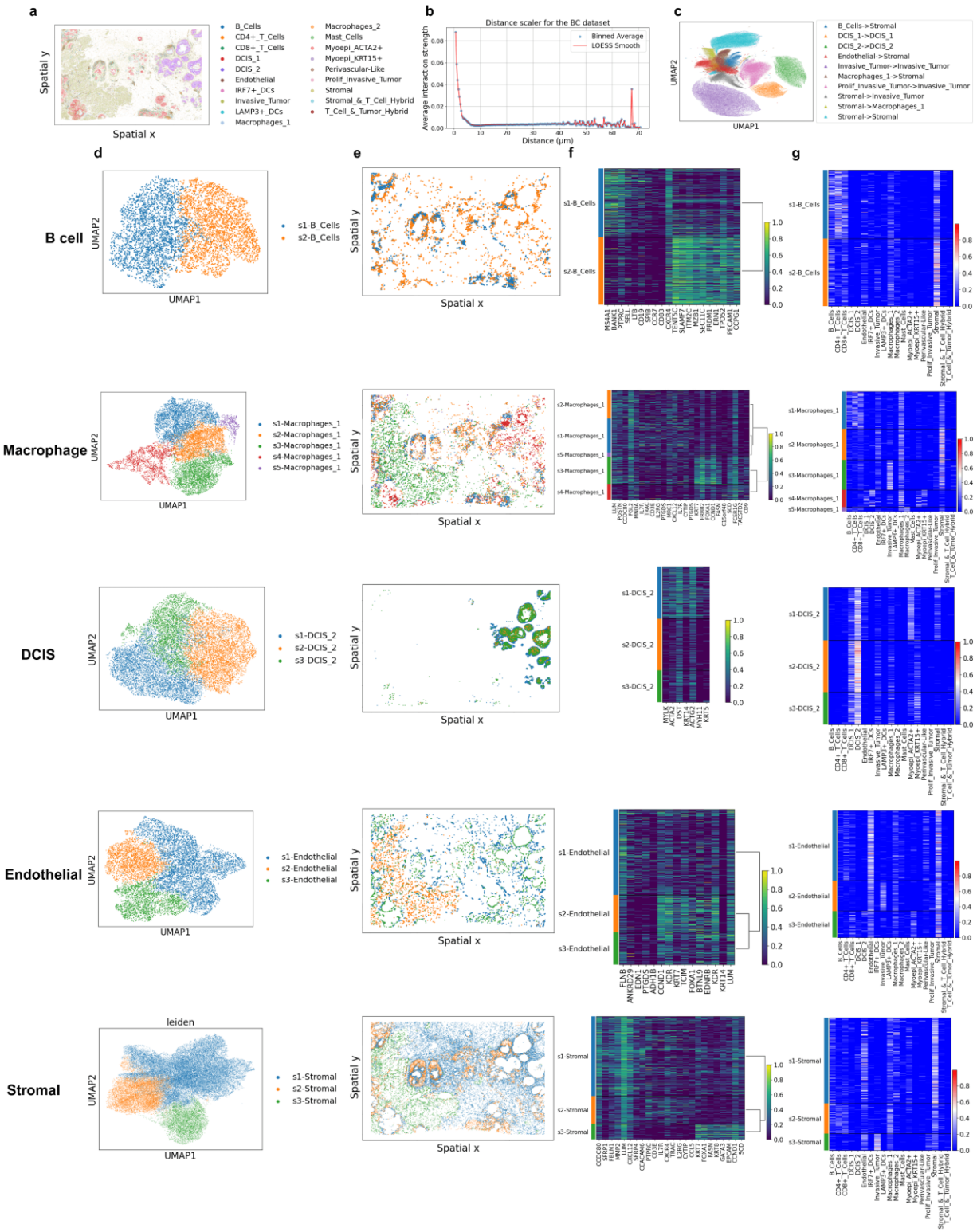


Fig. 5 Analysis of CCI-driven within-cell-type heterogeneity in the breast cancer dataset. a, Spatial distribution of 19 cell types. **b,** The LOESS curve of the estimated CCI strength decreased with increasing distance between interacting cells. **c,** UMAP visualization of CCI pairs. **d,** UMAP visualization of B cell, macrophage, ductal carcinoma in situ (DCIS), endothelial cell, and stromal

cell subgroups. **e,** Spatial distribution of B cell, macrophage, DCIS, endothelial cell, and stromal cell subgroups. **f,** Heatmap of the top differentially expressed genes in B cell, macrophage, DCIS, endothelial cell, and stromal cell subgroups. **g,** Neighboring cell type influence on B cell, macrophage, DCIS, endothelial cell, and stromal cell subgroups.

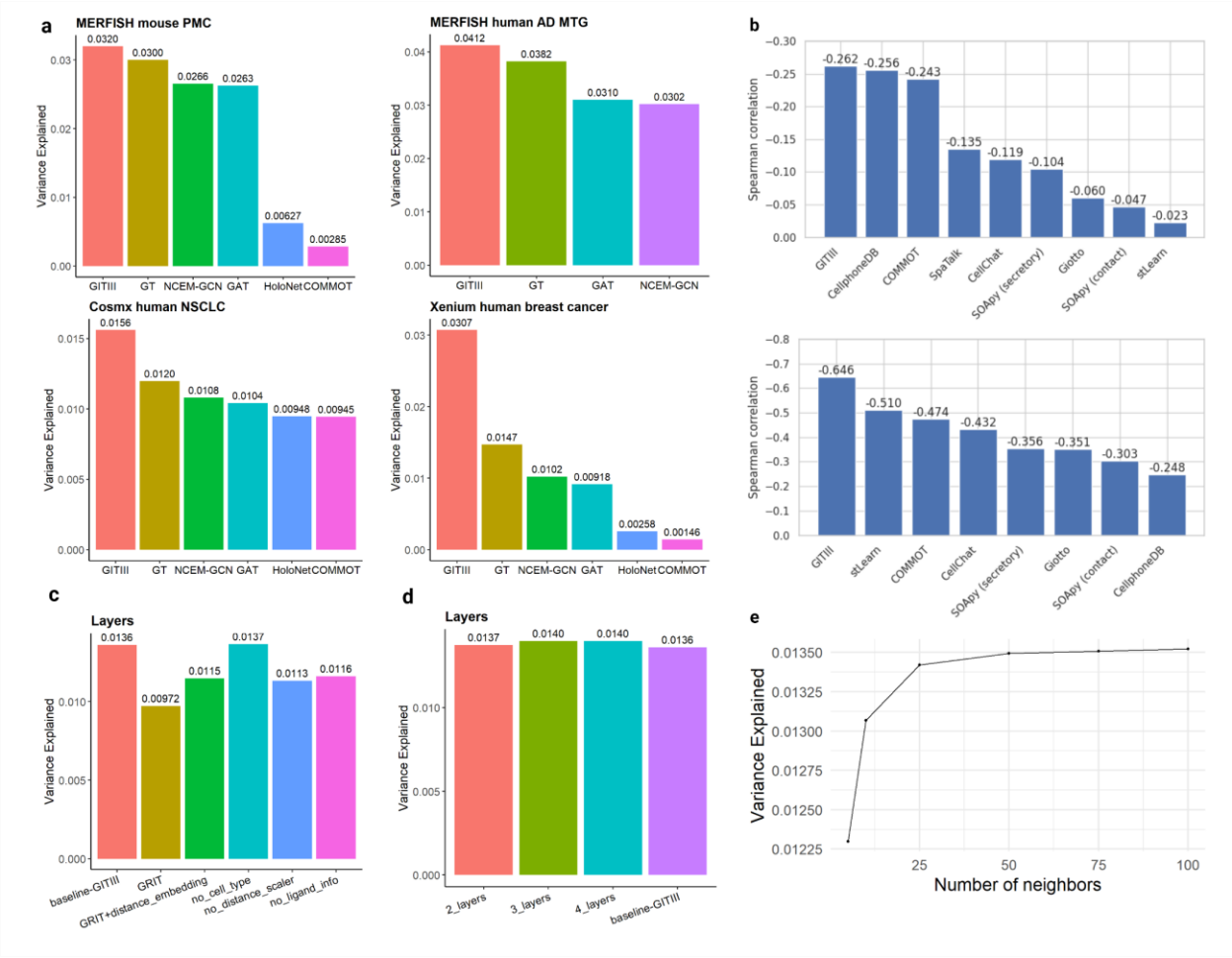
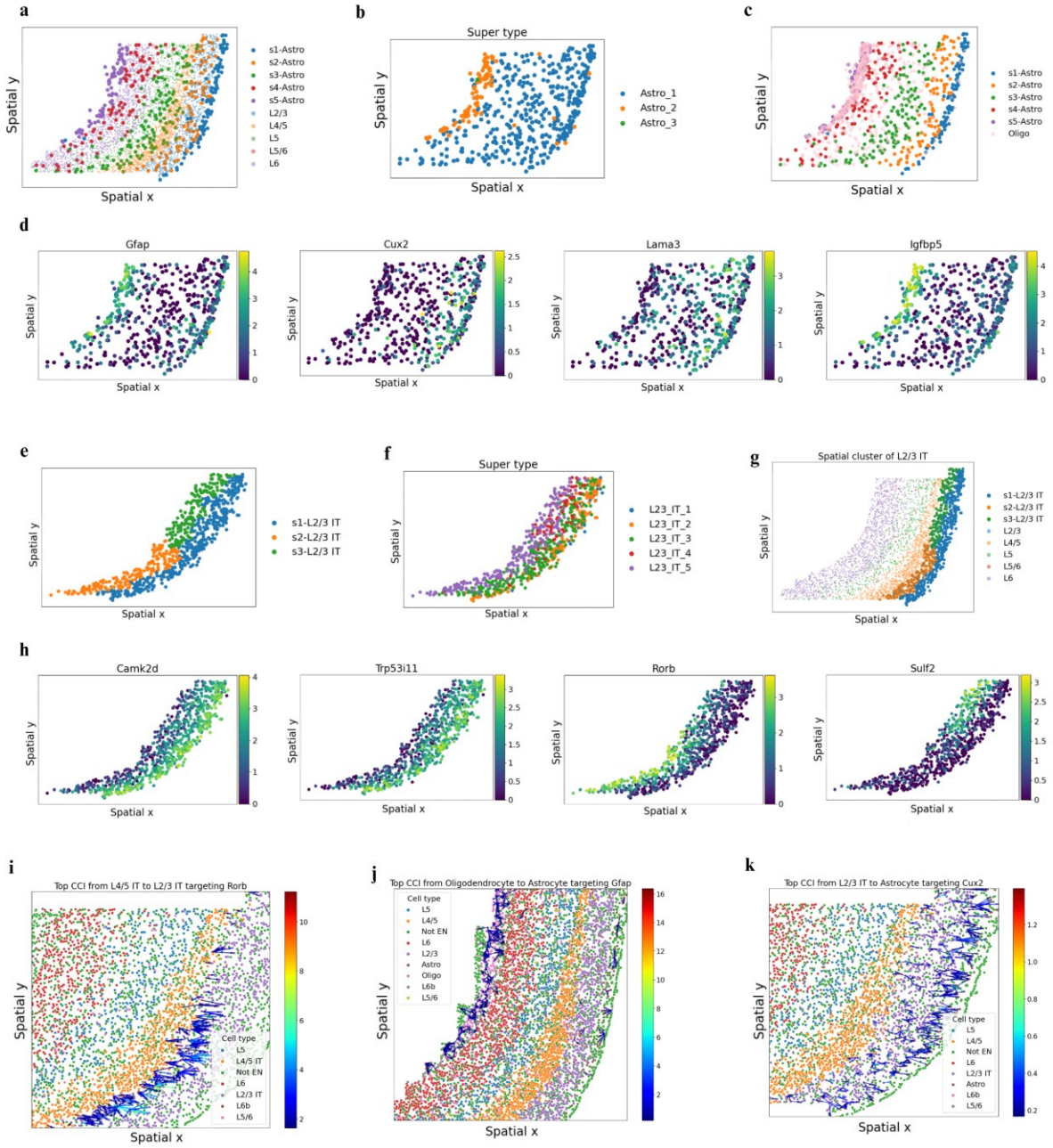
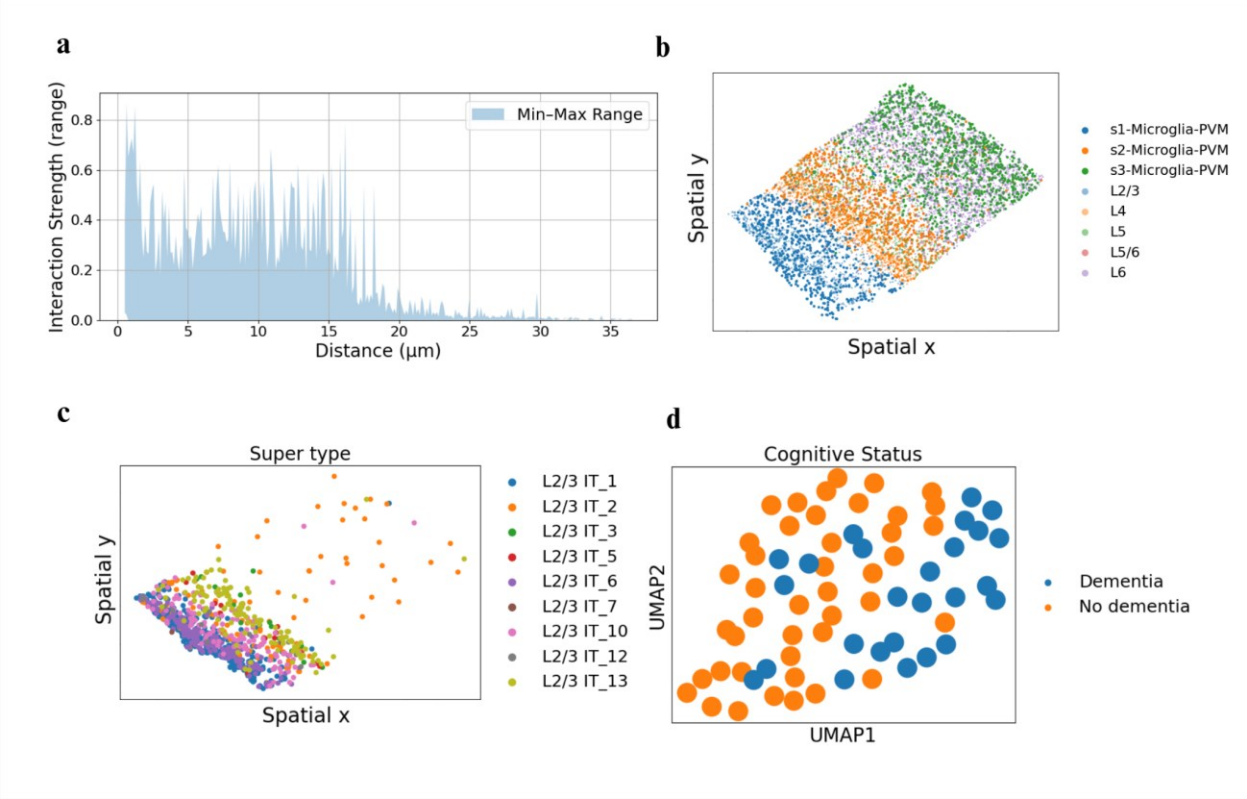


Fig. 6 Benchmarking, ablation and sensitivity analysis. a, Single-cell level comparison of the intracell-type variance explained by CCI inferred from GITIII, COMMOT, HoloNet, NCEM, graph attention networks, and graph transformer in the four spatial transcriptomics datasets. **b,** Cell-type level comparison of the Spearman correlation coefficient between the aggregated CCI network strength and the Wasserstein distance between sender-receiver cell-type pairs inferred by GITIII, CellChat, CellPhoneDB, COMMOT, Giotto, SOAPy, SpaTalk, and stLearn in the NSCLC (upper) and breast cancer (lower) datasets. **c,** Comparison of the intracell-type variance explained by CCI in the ablation study, where the components or information removed or disabled include ligand expression in the sender cells, cell type information, distance scaler, and all modifications to the GRIT block, using the NSCLC dataset. **d,** Intracell-type variance explained by CCI inferred from GITIII using one to four encoder layers in the NSCLC dataset. **e,** Intracell-type variance explained by CCI inferred from GITIII using different numbers of cells in the cellular neighborhood in the NSCLC dataset.

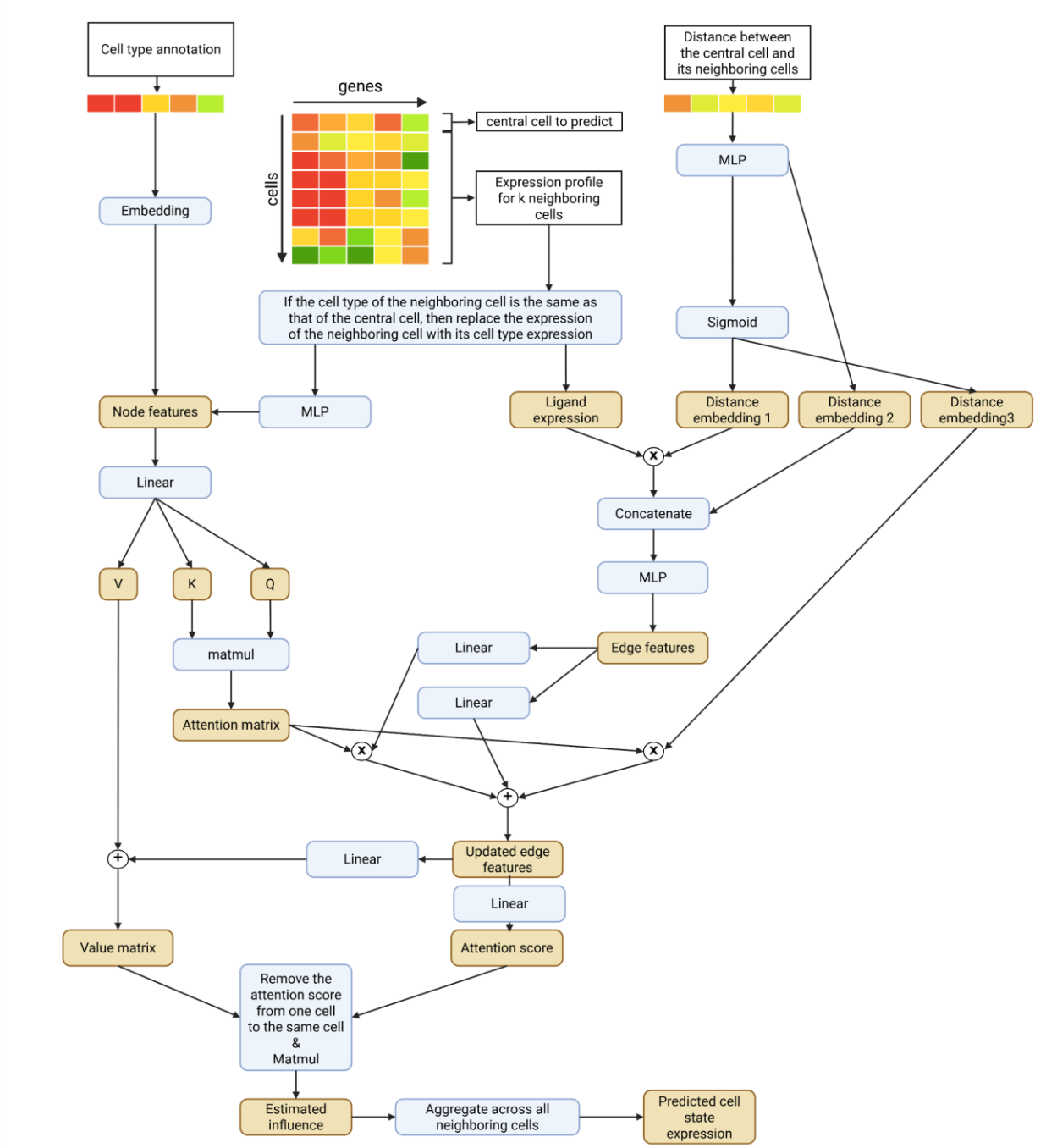


Extended Data Fig. 1 Spatial gene expression in subgroups identified by GITIII and spatial distribution of CCI influencing *Rorb*, *Gfap*, and *Cux2* in the mouse PMC datasets. a, Spatial distribution of astrocyte subgroups and excitatory neurons. Specifically, astrocyte subgroup 1 is located in layer 1, subgroup 2 is located in layers 2 and 3, subgroup 3 is located in layers 4 and 5, subgroup 4 is located in layer 6, and subgroup 5 is located in the white matter. **b,** Super type annotation of astrocytes by Zhang *et al.* **c,** Spatial distribution of astrocyte subgroups and oligodendrocytes. **d,** Spatial expression patterns of *Gfap*, *Cux2*, *Lama3*, and *Igfbp5* in astrocytes.

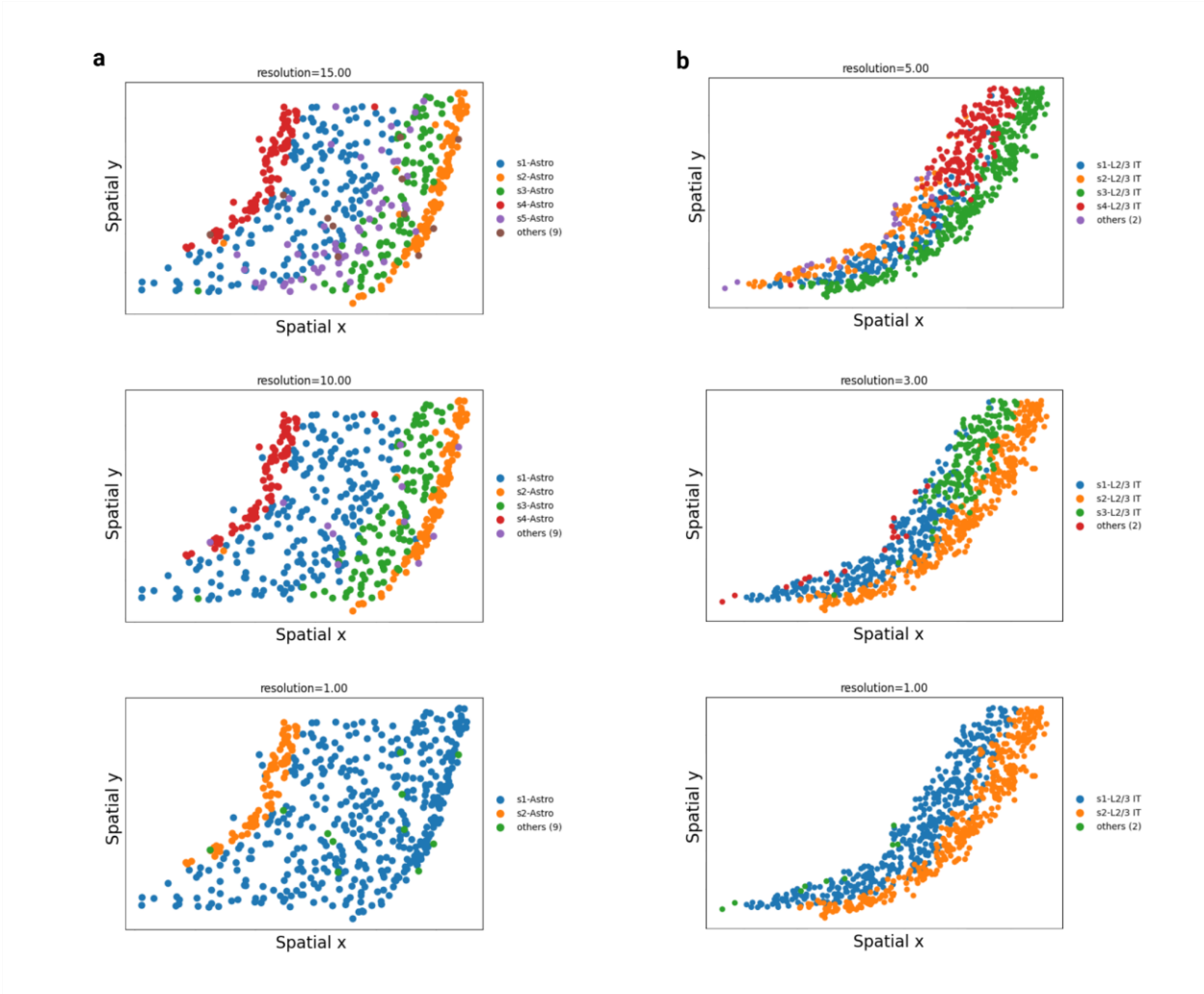
Astrocyte subgroup 5 located in the white matter are physically close to oligodendrocytes and had higher expression of *Gfap* and *Igfbp5* compared to other astrocytes. In contrast, astrocyte subgroup 2 located in layers 2 and 3 are physically close to L2/3 IT neurons and had higher expression of *Cux2* and *Lama3*. **e,** Spatial distribution of L2/3 IT subgroups. Specifically, subgroup 1 is located in the upper region of layers 2 and 3, closer to layer 1, while subgroups 2 and 3 are located in the deeper region of layers 2 and 3, closer to layer 4. **f,** Super type annotation of L2/3 IT neurons by Zhang *et al.* **g,** Spatial distribution of L2/3 IT subgroups and other excitatory neurons. **h,** Spatial expression patterns of Camk2d, *Trp53i11*, *Rorb*, and *Sulf2* in L2/3 IT neurons. L2/3 IT subgroups 2 and 3 located in the deeper region of layers 2 and 3 are physically close to L4/5 IT neurons and had higher expression of *Rorb* and *Sulf2*, respectively, compared to other L2/3 IT cells. **i,** Spatial distribution of interactions from L4/5 IT to L2/3 IT neurons influencing the expression of *Rorb*. **j,** Spatial distribution of interactions from oligodendrocytes to astrocytes influencing the expression of *Gfap*. **k,** Spatial distribution of interactions from L2/3 IT neurons to astrocytes influencing the expression of *Cux2*.



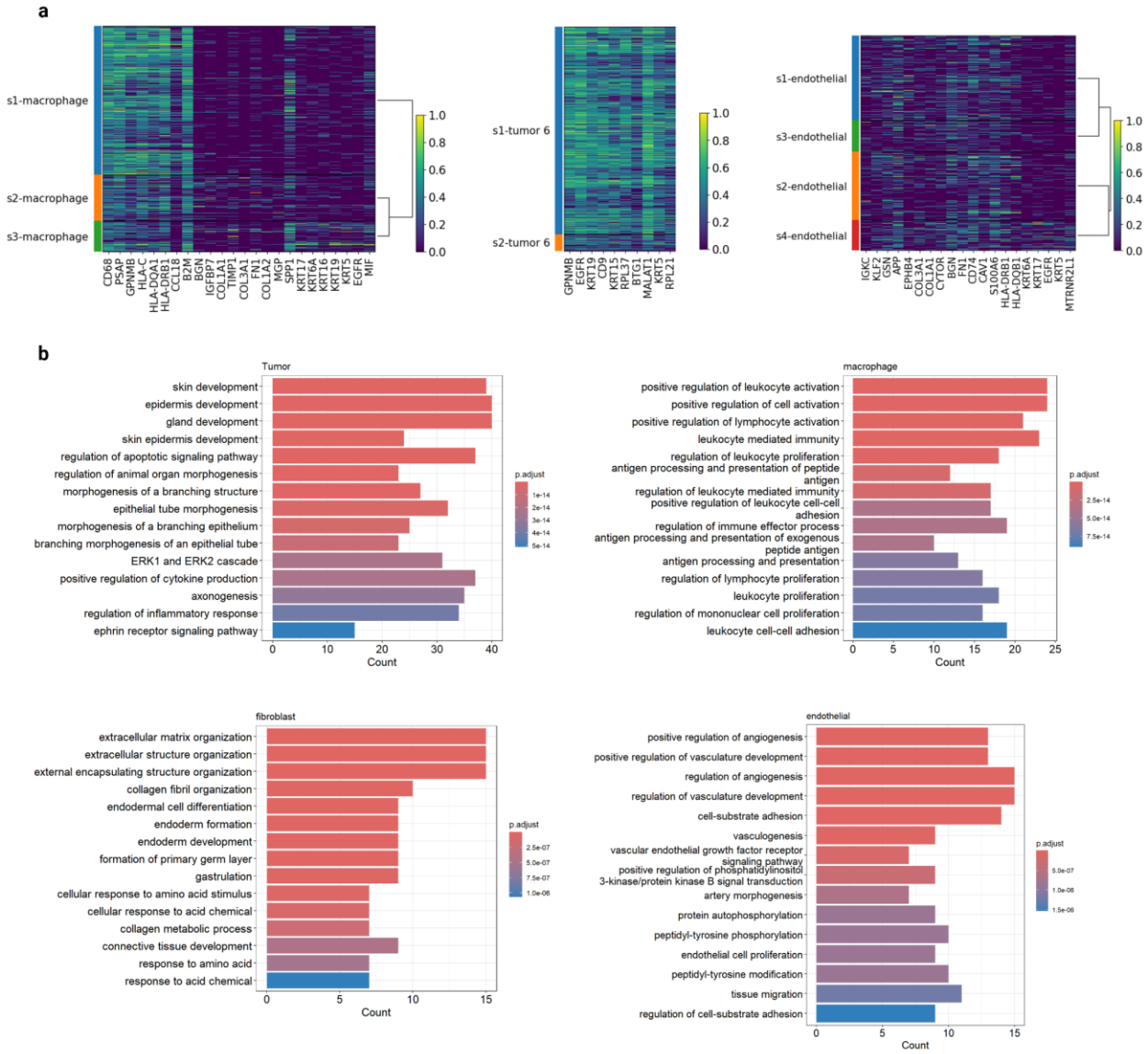
Extended Data Fig. 2 Spatial distribution of Microglia and L2/3 IT neurons in the SEA-AD dataset. a, The strength of interactions influencing the expression of *RORB*. **b,** Saptial distribution of microglia subgroups and excitatory neurons. Specifically, microglia subgroup 1 is located in layers 2 and 3, subgroup 2 in layers 4 and 5, and subgroup 3 in layer 6. **c,** Super type annotation of L2/3 IT neurons in Gabitto *et al.* **d,** UMAP visualization of slides using the inferred network strength as features, with color indicating slide from patients with or without dementia.



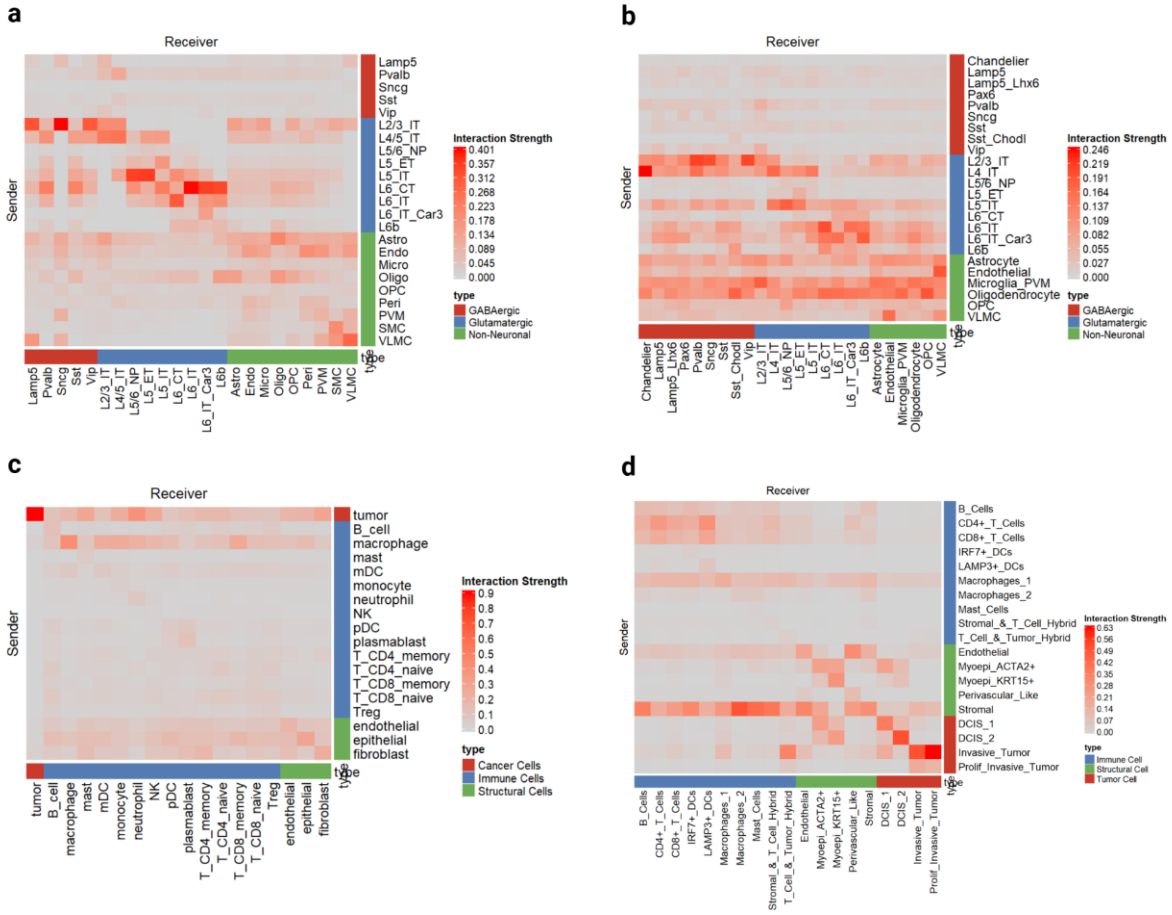
Supplementary Fig. 1 Graphical illustration of the data processing workflow in GITIII.



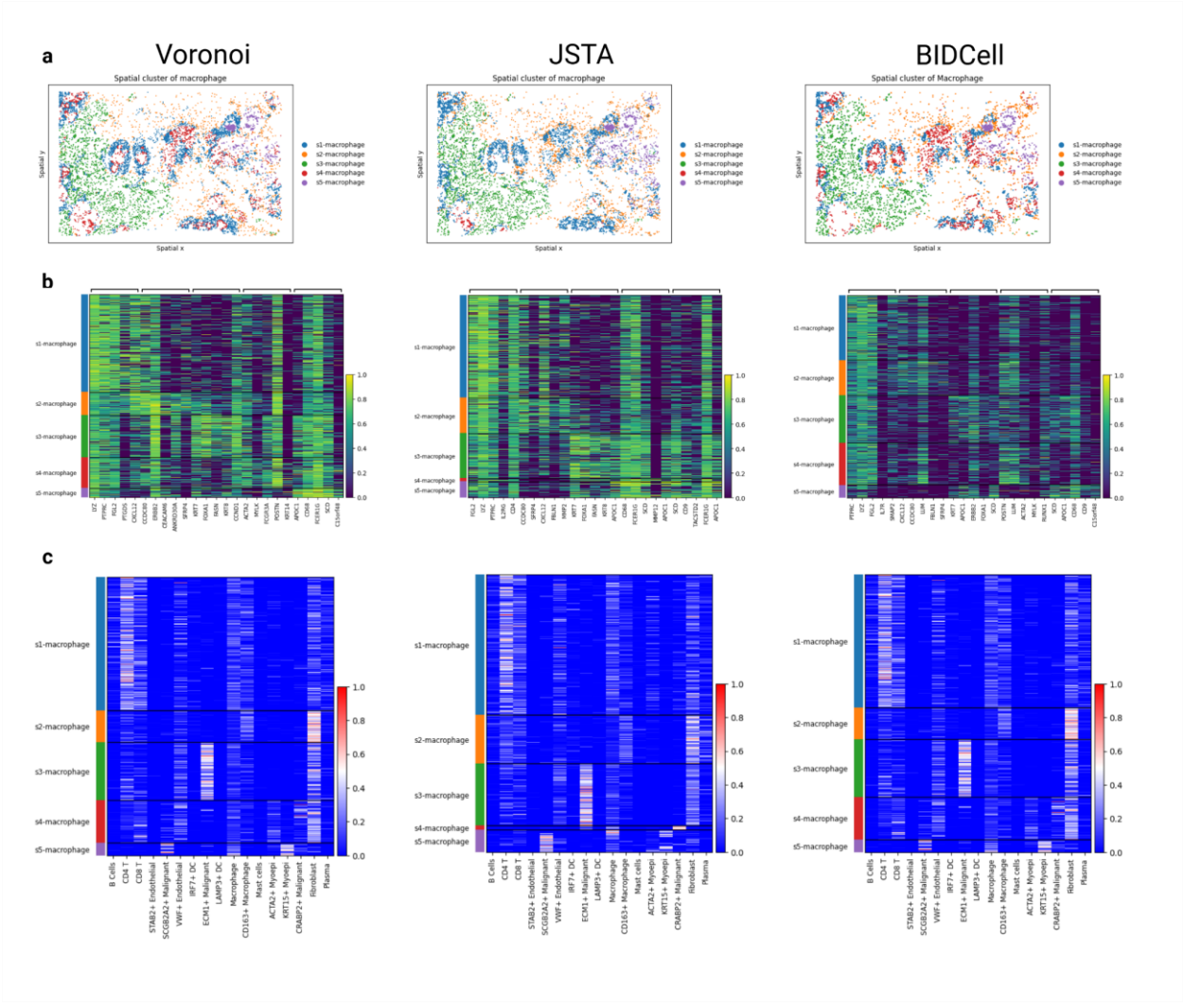
Supplementary Fig. 2 Spatial clustering by BANKSY in the mouse PMC dataset. a, Spatial distribution of astrocyte subgroups at different resolutions. Cells labeled as "others (x)" belong to small subgroups containing fewer than 20 astrocytes, where x denotes the number of such subgroups. **b,** Spatial distribution of L2/3 IT subgroups at different resolutions. Cells labeled as "others (x)" belong to small subgroups containing fewer than 20 L2/3 IT neurons, where x denotes the number of such subgroups.



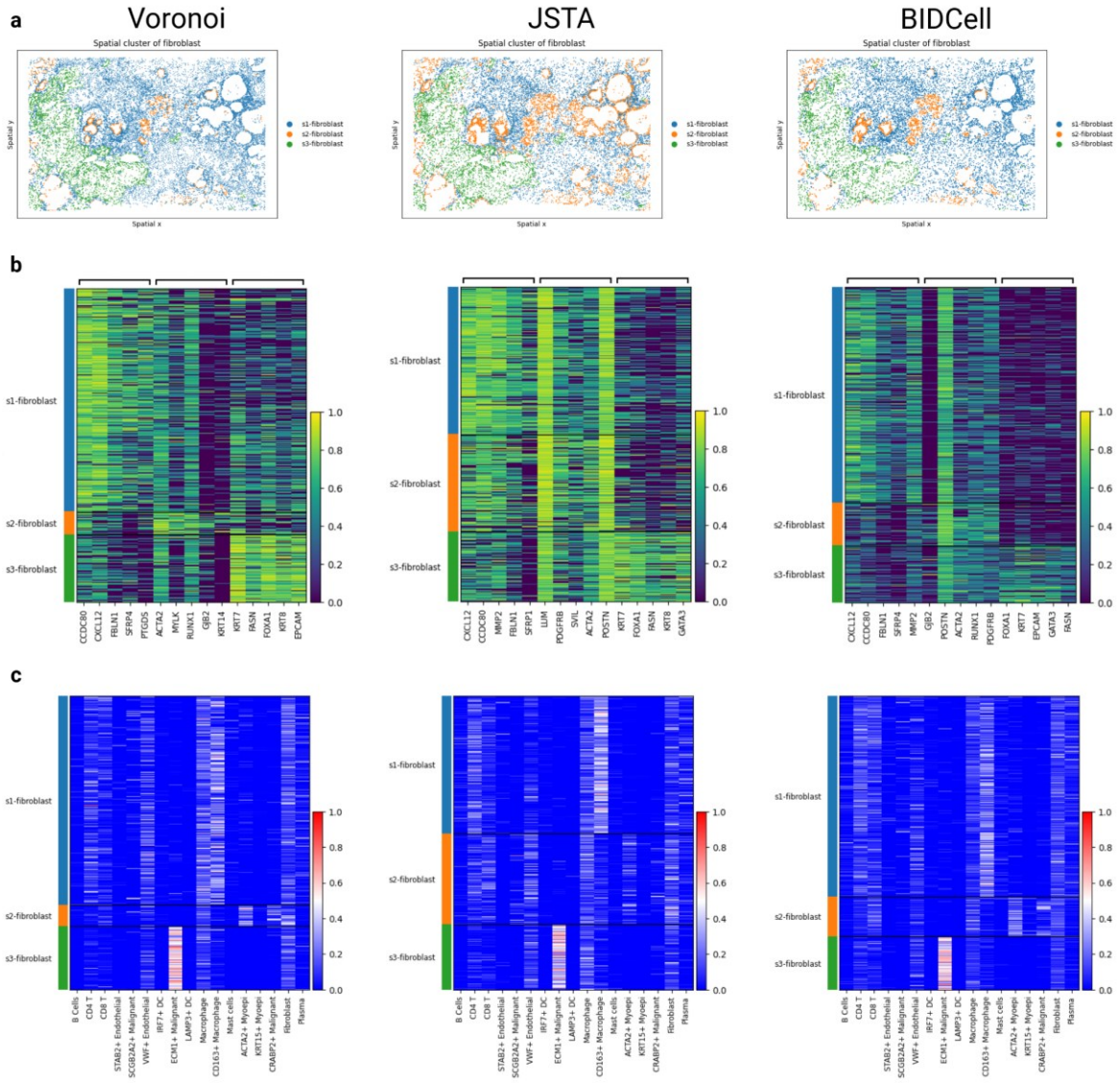
Supplementary Fig. 3 Heatmap of differential expressed genes and pathways of cell-type markers in the NSCLC dataset. a, Heatmap of the top differentially expressed genes in macrophage, endothelial, and tumor cell subgroups. **b,** GO enrichment analysis of the cell-type markers of tumor cell, macrophage, fibroblast, and endothelial cell.



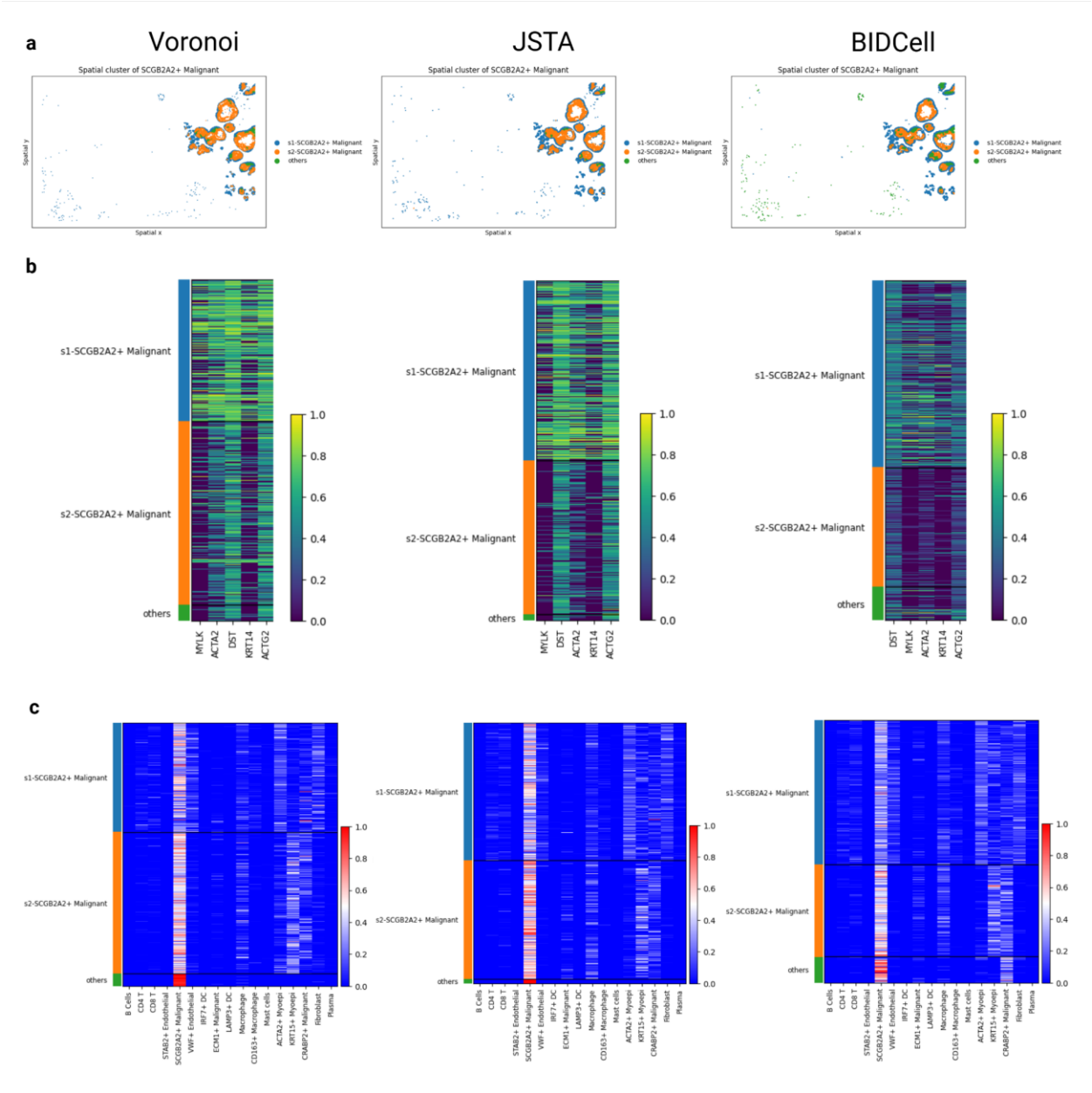
Supplementary Fig. 4 Composition of sender cell types influencing each receiver cell type in the (a) mouse PMC, (b) SEA-AD, (c) NSCLC, and (d) breast cancer datasets.



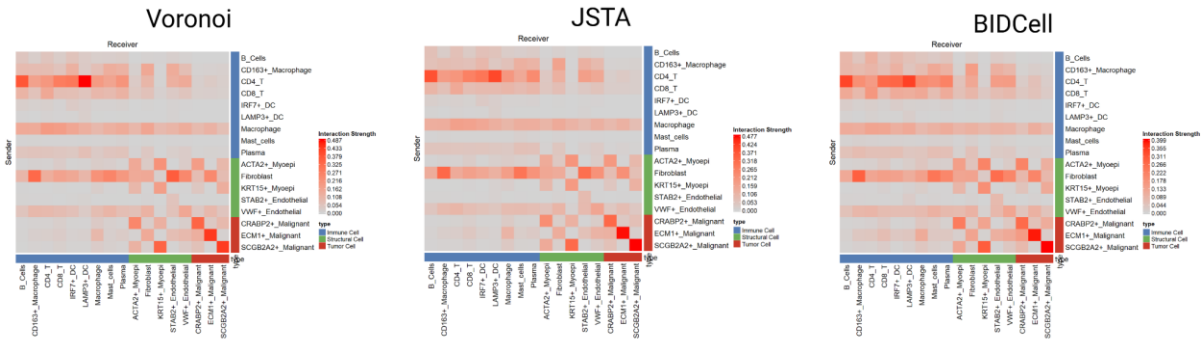
Supplementary Fig. 5 Comparison of clustering results of macrophages across data processed by different cell segmentation methods. a, Spatial distribution of subgroups. **b,** Heatmap of the top differentially expressed genes in subgroups. **c,** Neighboring cell type influence on subgroups, across three cell segmentation methods: Voronoi (left), JSTA (middle), and BIDCell (right).



Supplementary Fig. 6 Comparison of clustering results of fibroblasts across data processed by different cell segmentation methods. a, Spatial distribution of subgroups. **b,** Heatmap of the top differentially expressed genes in subgroups. **c,** Neighboring cell type influence on subgroups, across three cell segmentation methods: Voronoi (left), JSTA (middle), and BIDCell (right).



Supplementary Fig. 7 Comparison of clustering results of SCGB2A2+ malignant cells across data processed by different cell segmentation methods. a, Spatial distribution of subgroups. **b,** Heatmap of the top differentially expressed genes in subgroups. **c,** Neighboring cell type influence on subgroups, across three cell segmentation methods: Voronoi (left), JSTA (middle), and BIDCell (right).



Supplementary Fig. 8 Composition of sender cell types influencing each receiver cell type in the breast cancer data processed by three cell segmentation methods: Voronoi (left), JSTA (middle), and BIDCell (right).

References

1. Armingol, E., Baghdassarian, H.M. & Lewis, N.E. The diversification of methods for studying cell-cell interactions and communication. *Nat Rev Genet* **25**, 381–400 (2024).
2. Armingol, E. *et al.* Context-aware deconvolution of cell-cell communication with Tensor-cell2cell. *Nat Commun* **13**, 3665 (2022).
3. Graeber, T.G. & Eisenberg, D. Bioinformatic identification of potential autocrine signaling loops in cancers from gene expression profiles. *Nat Genet* **29**, 295–300 (2001).
4. Liu, Z., Sun, D. & Wang, C. Evaluation of cell-cell interaction methods by integrating single-cell RNA sequencing data with spatial information. *Genome Biol* **23**, 218 (2022).
5. Wang, T., Li, B., Nelson, C.E. & Nabavi, S. Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinformatics* **20**, 40 (2019).
6. Zhao, W., Johnston, K.G., Ren, H., Xu, X. & Nie, Q. Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat. *Nat Commun* **14**, 1128 (2023).
7. Arnol, D., Schapiro, D., Bodenmiller, B., Saez-Rodriguez, J. & Stegle, O. Modeling Cell-Cell Interactions from Spatial Molecular Data with Spatial Variance Component Analysis. *Cell Rep* **29**, 202–211 e6 (2019).
8. Cang, Z. *et al.* Screening cell-cell communication in spatial transcriptomics via collective optimal transport. *Nat Methods* **20**, 218–228 (2023).
9. Dries, R. *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol* **22**, 78 (2021).
10. Li, H. *et al.* Decoding functional cell-cell communication events by multi-view graph learning on spatial transcriptomics. *Brief Bioinform* **24**(2023).

11. Li, Z., Wang, T., Liu, P. & Huang, Y. SpatialDM for rapid identification of spatially co-expressed ligand-receptor and revealing cell-cell communication patterns. *Nat Commun* **14**, 3995 (2023).
12. Pham, D. *et al.* Robust mapping of spatiotemporal trajectories and cell-cell interactions in healthy and diseased tissues. *Nat Commun* **14**, 7739 (2023).
13. Schrod, S. *et al.* Spatial Cellular Networks from omics data with SpaCeNet. *Genome Res* **34**, 1371–1383 (2024).
14. Shao, X. *et al.* Knowledge-graph-based cell-cell communication inference for spatially resolved transcriptomic data with SpaTalk. *Nat Commun* **13**, 4429 (2022).
15. So, E., Hayat, S., Nair, S.K., Wang, B. & Haibe-Kains, B. GraphComm: A Graph-based Deep Learning Method to Predict Cell-Cell Communication in single-cell RNAseq data. *bioRxiv* (2024).
16. Wang, H. *et al.* SOAPy: a Python package to dissect spatial architecture, dynamics, and communication. *Genome Biol* **26**, 80 (2025).
17. Wu, D., Gaskins, J.T., Sekula, M. & Datta, S. Inferring Cell-Cell Communications from Spatially Resolved Transcriptomics Data Using a Bayesian Tweedie Model. *Genes (Basel)* **14**(2023).
18. Yang, W. *et al.* Deciphering cell-cell communication at single-cell resolution for spatial transcriptomics with subgraph-based graph attention network. *Nat Commun* **15**, 7101 (2024).
19. Browaeys, R. *et al.* MultiNicheNet: a flexible framework for differential cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. *bioRxiv* (2023).
20. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* **17**, 159–162 (2020).
21. Efremova, M., Vento-Tormo, M., Teichmann, S.A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484–1506 (2020).
22. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, 1088 (2021).
23. Jin, S., Plikus, M.V. & Nie, Q. CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. *Nat Protoc* **20**, 180–219 (2025).
24. Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S. & Zhuang, X. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
25. He, S. *et al.* High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat Biotechnol* **40**, 1794–1806 (2022).
26. Janesick, A. *et al.* High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat Commun* **14**, 8353 (2023).
27. Garcia-Alonso, L. *et al.* Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. *Nat Genet* **53**, 1698–1711 (2021).
28. Liu, J. *et al.* CytoSignal Detects Locations and Dynamics of Ligand-Receptor Signaling at Cellular Resolution from Spatial Transcriptomic Data. *bioRxiv* (2024).
29. Yan, L., Cheng, J., Nie, Q. & Sun, X. Dissecting multilayer cell-cell communications with signaling feedback loops from spatial transcriptomics data. *Genome Res* **35**, 1400–1414 (2025).
30. Fischer, D.S., Schaar, A.C. & Theis, F.J. Modeling intercellular communication in tissues using spatial graphs of cells. *Nat Biotechnol* **41**, 332–336 (2023).

31. Wang, H. *et al.* CGCom: a framework for inferring Cell-cell Communication based on Graph Neural Network. *bioRxiv* (2023).
32. Yuan, Y. & Bar-Joseph, Z. GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data. *Genome Biol* **21**, 300 (2020).
33. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598**, 137–143 (2021).
34. Lanjakornsiripan, D. *et al.* Layer-specific morphological and molecular differences in neocortical astrocytes and their dependence on neuronal layers. *Nat Commun* **9**, 1623 (2018).
35. Ben Haim, L. & Rowitch, D.H. Functional diversity of astrocytes in neural circuit regulation. *Nat Rev Neurosci* **18**, 31–41 (2017).
36. Batiuk, M.Y. *et al.* Identification of region-specific astrocyte subtypes at single cell resolution. *Nat Commun* **11**, 1220 (2020).
37. Bayraktar, O.A. *et al.* Astrocyte layers in the mammalian cerebral cortex revealed by a single-cell in situ transcriptomic map. *Nat Neurosci* **23**, 500–509 (2020).
38. Hodge, R.D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
39. O'Dea, M.R. & Hasel, P. Are we there yet? Exploring astrocyte heterogeneity one cell at a time. *Glia* **73**, 619–631 (2025).
40. Mathys, H. *et al.* Single-cell multiregion dissection of Alzheimer's disease. *Nature* **632**, 858–868 (2024).
41. Singhal, V. *et al.* BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nat Genet* **56**, 431–441 (2024).
42. Molina-Gonzalez, I. *et al.* Astrocyte-oligodendrocyte interaction regulates central nervous system regeneration. *Nat Commun* **14**, 3372 (2023).
43. Cubelos, B. *et al.* Cux-2 controls the proliferation of neuronal intermediate precursors of the cortical subventricular zone. *Cereb Cortex* **18**, 1758–70 (2008).
44. Nagata, I., Keilhauer, G. & Schachner, M. Neuronal influence on antigenic marker profile, cell shape and proliferation of cultured astrocytes obtained by microdissection of distinct layers from the early postnatal mouse cerebellum. *Brain Res* **389**, 217–32 (1986).
45. Gabitto, M.I. *et al.* Integrated multimodal cell atlas of Alzheimer's disease. *Nat Neurosci* **27**, 2366–2383 (2024).
46. Gao, C., Jiang, J., Tan, Y. & Chen, S. Microglia in neurodegenerative diseases: mechanism and potential therapeutic targets. *Signal Transduct Target Ther* **8**, 359 (2023).
47. Leng, F. & Edison, P. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here? *Nat Rev Neurol* **17**, 157–172 (2021).
48. Leng, K. *et al.* Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nat Neurosci* **24**, 276–287 (2021).
49. Stogsdill, J.A. *et al.* Pyramidal neuron subtype diversity governs microglia states in the neocortex. *Nature* **608**, 750–756 (2022).
50. Depp, C., Doman, J.L., Hingerl, M., Xia, J. & Stevens, B. Microglia transcriptional states and their functional significance: Context drives diversity. *Immunity* **58**, 1052–1067 (2025).
51. Fang, R. *et al.* Conservation and divergence of cortical cell organization in human and mouse revealed by MERFISH. *Science* **377**, 56–62 (2022).
52. Suzuki, T. *et al.* CUX2 deficiency causes facilitation of excitatory synaptic transmission onto hippocampus and increased seizure susceptibility to kainate. *Sci Rep* **12**, 6505 (2022).

53. Fujita, Y., Nakanishi, T., Ueno, M., Itohara, S. & Yamashita, T. Netrin-G1 Regulates Microglial Accumulation along Axons and Supports the Survival of Layer V Neurons in the Postnatal Mouse Brain. *Cell Rep* **31**, 107580 (2020).
54. Yaguchi, K., Nishimura-Akiyoshi, S., Kuroki, S., Onodera, T. & Itohara, S. Identification of transcriptional regulatory elements for Ntng1 and Ntng2 genes in mice. *Mol Brain* **7**, 19 (2014).
55. McNamara, N.B. *et al.* Microglia regulate central nervous system myelin growth and integrity. *Nature* **613**, 120–129 (2023).
56. Whalley, K. Neuronal neighbours tune microglial identity. *Nat Rev Neurosci* **23**, 582–583 (2022).
57. Bryan, K.J. *et al.* Expression of CD74 is increased in neurofibrillary tangles in Alzheimer's disease. *Mol Neurodegener* **3**, 13 (2008).
58. Su, H., Na, N., Zhang, X. & Zhao, Y. The biological function and significance of CD74 in immune diseases. *Inflamm Res* **66**, 209–216 (2017).
59. Olah, M. *et al.* Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nat Commun* **11**, 6129 (2020).
60. Bill, R. *et al.* CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. *Science* **381**, 515–524 (2023).
61. Yang, Q. *et al.* Single-Cell RNA Sequencing Reveals the Heterogeneity of Tumor-Associated Macrophage in Non-Small Cell Lung Cancer and Differences Between Sexes. *Front Immunol* **12**, 756722 (2021).
62. Lichtenberger, B.M. *et al.* Autocrine VEGF signaling synergizes with EGFR in tumor cells to promote epithelial cancer development. *Cell* **140**, 268–79 (2010).
63. Sharma, S.V., Bell, D.W., Settleman, J. & Haber, D.A. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* **7**, 169–81 (2007).
64. Lazaratos, A.M., Annis, M.G. & Siegel, P.M. GPNMB: a potent inducer of immunosuppression in cancer. *Oncogene* **41**, 4573–4590 (2022).
65. Saha, S.K. *et al.* KRT19 directly interacts with beta-catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties. *Oncogene* **36**, 332–349 (2017).
66. Astsaturov, I. *et al.* Synthetic lethal screen of an EGFR-centered network to improve targeted therapies. *Sci Signal* **3**, ra67 (2010).
67. Xu, S. *et al.* Using clusterProfiler to characterize multiomics data. *Nat Protoc* **19**, 3292–3320 (2024).
68. DeLeon-Pennell, K.Y., Barker, T.H. & Lindsey, M.L. Fibroblasts: The arbiters of extracellular matrix remodeling. *Matrix Biol* **91-92**, 1–7 (2020).
69. Lai, Y. *et al.* Interleukin-8 induces the endothelial cell migration through the activation of phosphoinositide 3-kinase-Rac1/RhoA pathway. *Int J Biol Sci* **7**, 782–91 (2011).
70. Denisenko, T.V., Pivnyuk, A.D. & Zhivotovsky, B. p53-Autophagy-Metastasis Link. *Cancers (Basel)* **10**(2018).
71. Jung, H., Seong, H.A. & Ha, H. Critical role of cysteine residue 81 of macrophage migration inhibitory factor (MIF) in MIF-induced inhibition of p53 activity. *J Biol Chem* **283**, 20383–96 (2008).
72. Pavlyukov, M.S. *et al.* Apoptotic Cell-Derived Extracellular Vesicles Promote Malignancy of Glioblastoma Via Intercellular Transfer of Splicing Factors. *Cancer Cell* **34**, 119–135 e10 (2018).

73. Chen, F. *et al.* Extracellular vesicle-packaged HIF-1alpha-stabilizing lncRNA from tumour-associated macrophages regulates aerobic glycolysis of breast cancer cells. *Nat Cell Biol* **21**, 498–510 (2019).
74. Horrevorts, S.K. *et al.* Glycan-Modified Apoptotic Melanoma-Derived Extracellular Vesicles as Antigen Source for Anti-Tumor Vaccination. *Cancers (Basel)* **11**(2019).
75. He, Y. *et al.* FOXA1 overexpression suppresses interferon signaling and immune response in cancer. *J Clin Invest* **131**(2021).
76. Schultz, S. *et al.* Progression-specific genes identified in microdissected formalin-fixed and paraffin-embedded tissue containing matched ductal carcinoma in situ and invasive ductal breast cancers. *BMC Med Genomics* **11**, 80 (2018).
77. Takahashi, Y., Kitadai, Y., Bucana, C.D., Cleary, K.R. & Ellis, L.M. Expression of vascular endothelial growth factor and its receptor, KDR, correlates with vascularity, metastasis, and proliferation of human colon cancer. *Cancer Res* **55**, 3964–8 (1995).
78. Hanley, C.J., Henriet, E., Sirka, O.K., Thomas, G.J. & Ewald, A.J. Tumor-Resident Stromal Cells Promote Breast Cancer Invasion through Regulation of the Basal Phenotype. *Mol Cancer Res* **18**, 1615–1622 (2020).
79. Ma, L.H. *et al.* Graph Inductive Biases in Transformers without Message Passing. *Proceedings of Machine Learning Research* **202**, 23321–23337 (2023).
80. Littman, R. *et al.* Joint cell segmentation and cell type annotation for spatial transcriptomics. *Mol Syst Biol* **17**, e10108 (2021).
81. Fu, X. *et al.* BIDCell: Biologically-informed self-supervised learning for segmentation of subcellular spatial transcriptomics data. *Nat Commun* **15**, 509 (2024).
82. Traag, V.A., Waltman, L. & van Eck, N.J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* **9**, 5233 (2019).
83. Wolf, F.A., Angerer, P. & Theis, F.J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15 (2018).
84. Lin, Y. *et al.* scClassify: sample size estimation and multiscale classification of cells using single and multiple reference. *Mol Syst Biol* **16**, e9389 (2020).