StereoSiTE: A framework to spatially and quantitatively profile the cellular neighborhood organized iTME

Xing Liu[1*], Chi Qu[2,3,1*], Chuandong Liu[1*], Na Zhu[1*], Huaqiang Huang[1], Fei Teng[1], Caili Huang[1], Bingying Luo[1], Xuanzhu Liu[1], Yisong Xu[2,3,1], Min Xie[2,3,1], Feng Xi[2,3,1], Mei Li[1], Liang Wu[1,2,3], Yuxiang Li[1], Ao Chen[2,3,1#], Xun Xu[1#], Sha Liao[2,3,1#], Jiajun Zhang[2,3,1#]

[1.] BGI-Shenzhen, Shenzhen 518083, China

[2.] BGI Research-Southwest, BGI, Chongqing 401329, China

[3.] JFL-BGI STOmics Center, Jinfeng Laboratory, Chongqing 401329, China

* ：Theses authors contributed equally to this study.

#: Co-corresponding author

Correspondence to:

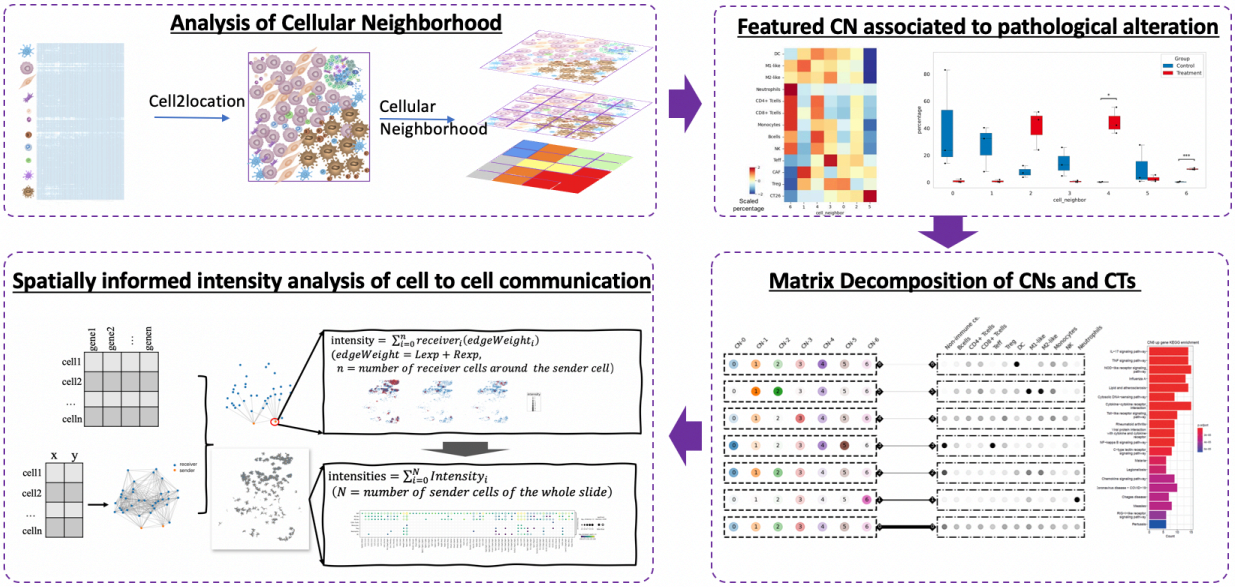
Dr. Jiajun ZHANG

BGI-Shenzhen, Shenzhen 518083, China

Email: zhangjiajun1@genomics.cn

Highlight

A framework based on space nearest neighbor graph and gene expression profile to spatially resolve iTME and to quantitatively define cell-cell communication intensity (StereoSiTE)



Abstract

Spatial Transcriptomics (ST) technology enables systematic depiction of regional milieu of a tissue, like tumor immuno-microenvironment (iTME). However, a powerful algorithmic framework to dissect spatially resolved niches, and to quantitatively evaluate spatial cell interaction intensity will pave the ways to understand the spatial signature associated mechanism. In this study, we provide a promising framework (StereoSiTE), which is based on space nearest neighbor graph and gene expression profile to spatially resolve iTME and to quantitatively define cell-cell communication intensity. We applied StereoSiTE to dissect the iTME of xenograft model receiving immunoagonist treatment, 7 distinct cellular neighborhoods (CN) were identified, and each CN was considered as the functional unit with exclusive cell type (CT) composition. Further deconvolving the joint matrix covering CNs and CTs indicated the importance of neutrophils in CN6, which was confirmed by pathway enrichment analysis. What's more, analysis of interaction intensity indicated that the recruited neutrophils preserved tumor protection activity through paired IL-1β/IL-1R after immunoagonist treatment exclusively in CN6. This evidence provided a new possible vision of tumor immune evasion orchestrated by neutrophils. StereoSiTE is believed to be a promising framework of mapping iTME niches using spatial transcriptomics, which could be utilized to spatially reveal tumoribiology mechanisms.

Key words: Stereo-seq, iTME, Cellular neighborhood, Cell to cell communication intensity, Xenograft tumor model

Introduction

Dissecting and deciphering the mechanisms by which genes bio-functionally orchestrates complex heterogeneity across inter- and intra- tumors is unprecedentedly crucial for improving treatment response of tumor. Over the past decade, the target of cancer therapy has gradually moved from tumor cells to other cell types in the tumor microenvironment(1). The important role of cell types in the immune tumor microenvironment (iTME) in tumor progression is increasingly emphasized. The iTME is a complex system of tumor cells, stromal cells, immune cells and extracellular components that rely on a high degree of spatial organization and interaction of cell types to exert tumor suppressive or promotive effects(2, 3). The iTME has been iteratively elucidated to propel malignant progression across cancer types. Mechanisms underlying in TME-induced intratumoral heterogeneity are deemed as guaranteed frontiers for understanding tumoribiology(4-6). Therefore, dissecting the coordinated activities of various cell types in iTME will advance our understanding of the mechanisms of tumor progression.

Unfortunately, trade-off between "-omics" tools generated comprehensive molecular measurement and "-targeted" tools conferred low-throughput but highly resolved cellular investigation in situ is a long-lasting debate, which, thereby, hinders our understanding in biology(7). However, this historical conundrum is being up to termination under the recent development of spatial transcriptomic (ST) technologies(8). The recent developments in the spatial transcriptome (ST) provide opportunities to better elucidate physiological and pathological progress covering cell-type definition(9-11), cellular interaction(7, 12), molecular signaling(13, 14), and hence facilitated us to spatially map the microenvironment landscape. However, there is still lacking a useful and systematic analytical framework for identifying important iTME units, locating major cell types, and delineating the spatial organization and orchestration of cell types in iTME functional units under tumor progression or response of therapy.

As most of the published algorithms utilize spatial information and gene expression profile by deep learning and graph neural networks model(15), whose performance relies on training data leading to poor repeatability and practicality. Some accessible algorithms used the spatial coordinate to construct space nearest neighbor graph without taking the gene expression profile into consideration(16, 17). Although these algorithms can be used to evaluate distance between cells, without valuable information on gene expression profile, researchers can hardly address scientific questions with reasonable conclusion from the spatial transcriptomics data. The stereoSiTE constructs space nearest neighbor graph based on spatial

coordinate and assigns the edge with weight based on ligand-receptor expression. Then, we established two key formulas, which can be used to compute the spatial cell interaction intensity (SCII). Besides, to quantitatively profile the cellular neighborhood organized iTME, StereoSiTE integrates the analysis of distinguishing different cellular neighborhood (CN) regions based on cell composition resolved by deconvolution. The combination of SCII with CN will pave the ways for exploring significant scientific discovery.

Here, we present StereoSiTE, an analytical framework for comprehensive depiction of iTME multicellular communities from the spatial transcriptome. Our approach combines spatial coordinate based spatial nearest neighbor maps and gene expression profiles of cell types to elucidate iTME multicellular communities, and to evaluate the quantitative intensity of cellular interactions from spatially resolved gene expression data. To demonstrate the utility of this framework, we applied StereoSiTE to xenograft models receiving immunoagonist treatment. We then defined 7 CNs across 6 xenograft tumor samples. Importantly, CN6 was significantly altered after treatment of immunoagonist, which was characterized by enrichment of neutrophils. We also observed that the recruited neutrophils preserved tumor protection activity through paired IL-1β/IL-1R after immunoagonist treatment exclusively in CN6 by analysis of interaction intensity. This evidence provided a new possible vision of tumor immune evasion orchestrated by neutrophils. These results underscore the necessity of analytical framework, StereoSiTE, to dissect the immune tumor microenvironment by identifying cellular neighborhoods and elucidating cell interaction under spatial organization.

Results

Framework based on space nearest neighbor graph and gene expression profile

While the spatial transcriptomics data at single-cell resolution was available(18), there is urgent need to develop a data analysis solution or algorithms to unearth the spatial signature hidden under tumoribiology. We designed a framework for spatial transcriptomics data analysis, which can fully utilize the spatial coordinate and gene expression profile (Fig 1A). The framework contains two key components. First, we cluster the entire data matrix into specific functional units based on their different cellular neighborhood (CN). Then, the spatial cell interaction intensity was computed by space nearest neighbor graph which combined the spatial coordinate and ligand-receptor expression information (Fig 1B& C). Here, we can collect spatially resolved transcriptomic data from tumor tissue by Stereo-seq. As registration of the spatial gene expression profile with the nuclear staining, which records the cell nuclear morphology, data matrix at resolution of single cell was obtained after cell segmentation

processing. Then the cell types will be defined based on their gene expression profile by Cell2location(19) with reference from publications(20). One the other side, without taking the cell morphology into consideration, we binned data can be obtained by binned 100 × 100 DNB bins into one analyzed window with size of 50x50um. We assigned captured mRNA into specific binned 100, which is a square of specific size, for example bin100 means a square with side length 100 (the unit of length is capture site). We resolved the cell composition of every bin100 which contains about 25 cells by deconvolution (Cell2location)(19), which can be considered as specific cellular neighborhood associated with iTME (Fig 1C). At last, we can perform integrated analysis of the CN distribution and the annotation at single cell resolution. CNs of interest were featured by associating to pathological alteration. Then data matrix at single cell resolution in featured CNs could be extracted for computing the spatial cell interaction intensity (SCII) to explore how intercellularly interacts and biologically function (Fig 1A). To compute SCII, we firstly constructed the cell neighborhood graph based on the spatial location of every cell, then assigned edge with weight based on ligand-receptor gene expression of the corresponding sender and receiver cells. The key formulas to compute intensity and intensities are listed in Fig 1B. Spatial cell interaction intensity of every sender cell with its paired receiver cells is calculated by summing of the edge weight between them. And the spatial cell interaction intensities of the entire data matrix or specific CN region equals to the sum of all connected edge weight. Therefore, we constructed this framework called StereoSite, which was a powerful algorithmic tool to **S**patially and quantitatively profile the cellular neighborhood organized **iT**ME.

StereoSiTE precisely and distinguishingly measures cell-to-cell communication in specific cellular neighborhood organized iTME region.

To verify that StereoSiTE could identify distinct cellular neighborhood organized iTME, we applied this framework to a representative Stereo-seq data in single-cell resolution of sample from xenograft model. As shown in Fig 2A, the whole slice showed specific neighborhood enrichment pattern. We performed clustering analysis and obtained seven distinct cellular neighborhood organized iTME regions based on different cell composition. The Fig 2B illustrated that each CN region had specific neighborhood enrichment and cell composition. For example, CN1, CN3 and CN4 dominantly contained Treg cells, neutrophils and Teff cells, respectively. On the other hand, there was more self-neighborhood enrichment of different immune cells in CN1 than other CNs, while neighborhood enrichment of neutrophils in CN3 and neighborhood enrichment of Teff in CN4 were observed. As each CN

showed distinct feature, it was reasonable to group samples based on CNs and to further investigate cell-cell communication in the CN region of interest. This framework can help us to find some new scientific observation, which could be possibly vailed by analysis of whole section. Next, we computed the SCII between non-immune cells with all immune cells in CN4 region, and presented the results with dot plot (Fig 2C). The size of dot corresponded with SCII score, and the color represented ligand-receptor expression. To estimate the performance of SCII in assessing the strength of intercellular communication, we computed the p values of ligand-receptor pairs by CellPhoneDB, which was re-implemented in squidpy (Fig 2D). Then we compared SCII intensity with p values of the paired ligand-receptor. Fig 2E illustrated that SCII could find a greater number of cell-to-cell communication events. Meanwhile, SCII covered most of the interaction pairs found by permutation test method. From the comparison of the Mif and Cd74 induced cell-to-cell communication results which were highlighted by red square in Fig 2C and Fig 2D, we found that SCII had robust power to precisely and distinguishingly measure cell-to-cell communication activities. And differences can be obviously observed by comparing space cell interaction intensity while the p values from different groups were hardly comparable (Fig 2G). To confirm that the discrimination between different cell pairs was true, we plotted the spatial cell interaction intensity in situ (Fig 2F). Very strong interaction between non-immune cells with Teff and relatively weak interaction between non-immune cells with Treg, Monocytes and DC in CN4 region can be observed, which was missing in permutation test method. The interactions between non-immune cells with M1-like, M2-like and Neutrophils were different, which was consistent with SCII (Fig 2F& 2G). At last, there was no interaction between non-immune cells and B cells, which had similar observation in both of the two methods.

Profiling of tumor microenvironment using spatial transcriptomics

To apply designed framework to address iTME associated research questions, we introduced xenograft models (Fig 3A) with treatment of immune agonist (STING agonist). BALB/c mice were subcutaneously injected with colon cancer-originated CT26 cells and respectively treated with vehicle or DMXAA 14 days after tumor transplantation. Samples were harvested after 24 hours treatment and Stereo-Seq was subsequently performed (Fig 3A). Here, we collected spatially resolved transcriptomic data from xenograft tumor tissues by Stereo-seq, a spatial sequencing technology with the subcellular resolution of 500nm**.** Data matrix at resolution of single cell was obtained after cell segmentation processing based on nuclear staining. By cell2location induced deconvolution of spatial transcriptomic matrix based

on reference previously reported(20), we identified and validated 12 distinct cell types (Fig 3B), including 6 of lymphoid-lineage, 5 of myeloid-lineage and 1 of non-immune cluster. With population proportion of cell types across samples (Fig 3C), we noticed the different compositions of immune cells across samples. Further comparison in quantitative analysis indicated a divergent variation of cell numbers in different groups (Fig 3D), this may result from varied treatment regimens, where necrosis attributed to reducing cell numbers in treatment group (Fig 3E). One of advantages of spatially resolved sequencing is to allow in situ mapping of cell types (Fig 3F). Notably, control groups had higher frequency of M2-like macrophages, while, treatment groups possessed higher frequencies of neutrophils. Interestingly, we synchronously observed a location preference (Fig 3F) of M2-like macrophages in control groups, where they frequently located and associated with tumor cells and of neutrophils in treatment groups, where they tended to cluster around necrosis niches (Fig 3E & 3G). However, methods interrogating the correlation between specific bioactivities and the corresponding spatial preference were rarely exploited. To validate our hypothesis of spatial preference between different cell types, we applied analysis of CN in following section.

Organizing TME-associated cellular neighborhood

The immune tumor microenvironment (iTME) heterogeneity prevails both intra- and inter- tumor, which is intrinsically attributed by varied cell organizations in each spatially compartmented niche. The most possible extension to interrogating and elucidating iTME of xenografts, in short of distinct histological characteristics, is to visualize tissues with CNs. We firstly clustered windows labelling all samples and identified 7 distinct CNs co-currently existent in both groups (Fig 4A). Interestingly, constitution of cell types in each CNs significantly ranged across the cohort, this in return suggested that different immune cells preferentially co-localize and associate with certain cell types in compartmented iTME niche. We therefore entitled each CN based on their dominant cell proportions (Fig 4B). We next calculated the frequencies of CNs in different groups (Fig 4C) and observed a distinct correspondence of CN4 (lymphoid cell lead) and CN6 (neutrophils enriched) elevation in treatment group (Fig 4D), which was aligned with the treatment background as expected. To spatially exhibit and evaluate CNs (Fig 4E), we sought to assess CN6 in situ since neutrophils were acquiescently recruited by chemokine motivation but not persistent in targeted tissues (21-23). We therefore orthotopically projected CN6 to adjacent H&E staining to probe the putative distribution pattern of this neutrophils dominant iTME niche (Fig 4F). An obvious

trend of CN6 co-localizing around necrosis edges comparing to other CNs was observed, which potentially highlighted specific bioactivities exserted by tumor cells after receiving immunoagonist therapy. Altogether, we herein introduced CNs to spatially organize different iTME niches and exclusively exhibited distinct biological characteristics across divergent units.

Deconvolving iTME regarding to individual CNs

The CNs we herein organized provided a new dimension in evaluating iTME other than the CTs, where spatial information of cell neighborhood in tissue was included. We believed that we could shape the iTME more comprehensively by synchronous evaluating CTs and CNs. We therefore considered, for each sample, the matrix co-currently covering CNs and CTs should be utilized, instead of individually evaluating CNs and CTs, to elucidate how iTME differs between different groups. Tensor technology was hence induced(24). By decomposing the module matrix, we determined that dissecting each sample with 7 CN modules and 7 CT modules could furthest retain information and synchronously exhibit biological features across different samples (Fig 5A). Subsequent analysis indicated a potent coupling between CN6 and neutrophils which was further enhanced in treatment group (Fig 5B). The result suggests a prompt and activated crosstalk of neutrophils with other cells in CN6 in response to immunoagonist treatment. We next decided to look into CN6 and deeply shape the iTME niche. With KEGG analysis in CN6 of treatment set comparing to control set, we identified a potent recruitment activity of neutrophils in this milieu, as upregulated IL-17 signaling, TNF signaling and chemokine signaling pathways. Moreover, tumor associated inflammatory response was also enhanced (Fig 5C). In accordance, GO analysis and GSEA were both performed, and the result agreed with conclusion from KEGG analysis (Fig 5D&5E). In addition, cell-cell communication in this unit by SCII indicated a potent interaction between neutrophils and tumor cells with IL-1β/IL-1R overexpression (Fig 5F) and in situ intensity visualization of this pair revealed the frequent crosstalk around necrosis niches (Fig 5G), which was consistent with the biological feature of pro-tumor neutrophils (Fig 3B) and suggested a tumor-protecting function of neutrophils in CN6(22, 25). Altogether, we shaped the iTME niche by deconvolving the biological event undergoing in CN6 and indicated the recruitment of neutrophils with accompanied tumor protection activities after receiving immunoagonist treatment.

Discussion

Spatially organized iTME and its disorganized manner can be manifested as pathological disease. Therefore, to decode spatial signature of iTME is crucial to find the clinicopathological association for understanding mechanism(8). Alterations in tissue organization and cellular morphology set up the foundation of diagnosis of cancer, and provide gold criteria in pathology(26). It indicates the irreplaceable role of spatial information in understanding the initial and process of cancer. Pathological structures like invasive front, buddings and Tertiary Lymphoid Structures (TLSs) have been studied or used in pathology to indicate risk of metastasis, relapse and survival(8). Ji et al.(12) reported that tumor-specific keratinocytes (TSKs) at leading edges of squamous cell cancer associated to tumorigenesis. TSKs function like a hub to interact with nearby immune cells to form a reactive neighbor. Different cell types build up and remodel the iTME by crosstalk with cells in their neighborhood(3). Grunwald et al.(27) discovered "subTMEs" in human pancreatic cancer guided by pathological annotation. Immuno-activities of subTMEs were phenotyped by celltype components and differentiation status of cancer associated fibroblasts. Result of this paper suggested that reactive unit in iTME can be defined as a group of cells which interact within the neighbor, which has more relevant to pathological disease. This conclusion agreed with the concept in our framework, in which, cellular neighborhood could be featured and associated to alteration in tumor. Key cellular components in targetable structure have more accurate prognostic power.

Emerging evidences show cell neighborhoods could be the important iTME units with different cell compositions to indicate distinct biological processes. Schurch et al.(3) reported nine conserved cellular neighborhoods at invasive front of the colorectal cancer. CNs were characterized by the main cell types, and intra-CN communications orchestrated the antitumor activities. Afterwards, more studies have discussed CNs in multiple disease(28-30). It has been indicated that CNs might spatially reflect TME organization. Cell to cell communication in CN possibly leads to resultant phenotypes and participants in reprogramming regional environment. In our result, we found a specific CN (CN6) associated to inflammatory reaction like IL-17 signaling pathway. In the meantime, we found CN6 mainly composed of neutrophils whose biological activity is aligned with function of CN6. With indication of H&E, the position of CN6 lies on the edge of necrosis area. Recent cancer research highlighted importance of differentiation of neutrophils including high heterogeneity and plasticity(31). In our model, the neutrophils seem to be polarized into protumoral N2(21), which performs the protumoural activities. Therefore, our framework utilized the spatial information of cellular neighborhood

to target important CN groups associated to pathological condition. In current on-going studies, our group has been further proving the molecular mechanism of neutrophils regulating tumoral cells in CN6 with wet experimental evidence.

Spatial multi-omic techniques promises new modalities to obtain spatially informed omic data (8). Comparing to other omic methods, spatial transcriptomic technologies can generate high-throughput data at resolution of subcellular level(18). The state-of-art sequencing method enables researchers to obtain the genuine landscape of TME without unbalanced loss of cells caused by dissociation(32). With multi-dimensional and high throughput of gene expression, it requires powerful algorithmic toolchain to efficiently associate specific cell types to molecular activities, biological processes and clinical manifestation to address scientific questions. In this paper, we developed a framework called StereoSiTE to identify cellular neighborhood unit harboring disease associated spatial features, and to further decipher underlying biological function of each CN. In addition, we offer a new algorithm in this framework, which can qualitatively measure the interaction intensity of paired ligand and receptor with spatially confined cell to cell communication.

In our framework, we defined the cellular neighborhood with portion of CTs. However, recognizing cell identity in spatial transcriptome is extremely hard due to the relatively high dropout rate in gene capturing and low resolution. Deconvolution is a common solution to deduce the cell composition in each analytical window based on expression profiles. We used Cell2location with reference of single-cell RNA sequencing to perform annotation. Although this method has been designed for deconvolution, we also generated a single cell resolution data matrix in Stereo-seq data. The result showed purity and specificity of cell markers with cell2location in single cell spatial data could be satisfactory. After testing references from different publications, we found that quality of reference guaranteed the success of annotation in spatial sequencing data. Therefore, we recommended that testing the reference quality beforehand. Therefore, cell2location is a recommended tool in annotation of spatial data at different resolution in our framework.

Recent advent in spatially resolved omic technologies promises the investigation of authentic cell to cell communication. Therefore, our framework adds distance filter at cut-off of 200um(33). It is proposed that the longest acting distance of paired ligand-receptor is 200um. With in situ mapping of each gene expression, we add the distance filter in this framework when calculating ligand-receptor to reduce the false positive signal. In addition, we developed a novel algorithm to calculate the interaction intensity of paired ligand-receptor based on gene

expression. Our framework offers a new paradigm to quantitatively compare cell interaction intensity and to avoid false positive interaction by referring to physical distance on tissue.

We also aware that the function of StereoSiTE has been limited as follows. In our approach, the results of deconvolution annotation by cell2location have proved their accuracy, but it requires the users to provide a high-quality single-cell reference dataset suitable for their study design. Moreover, another key analytic compartment of StereoSiTE is spatial cell interaction intensity (SCII) which is a novel method to analyze cell to cell communication by integrating gene expression and spatial coordinate information. In this paper, the single-cell resolution data is used for SCII analysis. Due to our experimental model with high cell density, there is a certain drop-out rate of cells resulted from inaccurate identification of cell segmentation. In addition, we used the database from cellphoneDB for SCII analysis of mouse samples, databases covering other species will be included in the future. Larger sample sizes are needed to make conclusive biological explanations of the differences in cellular neighborhoods and cell interactions observed between the treatment and control groups.

With fast development of spatial transcriptomic technology, we're granted with high-throughput data and spatially resolved information when investigating the intricate network of tumoribiology. Here, we developed a comprehensive analytic framework, StereoSiTE, which contributes to spatially resolved iTME by organizing different iTME niches and dissecting detailed biological process within each individual CN. In the xenograft model, we identified a distinct CN that phenotypically and quantitatively altered after immunoagonist treatment. Subsequent deconvolution of this niche indicated a neutrophil leading bioactivity. Further enrichment analysis and spatial cell interaction intensity (SCII) analysis in return revealed a tumor exerted defensive action under immunoagonist exposure, which was by recruiting pro-tumor neutrophils to this very region. Altogether, our approach of mapping iTME niches using spatial transcriptomics could be utilized to spatially reveal tumoribiology mechanisms inherent in orthotopic tissues.

Conflict of Interest

The authors report no conflicts of interest in this work.

Acknowledgments

The authors would like to acknowledge China National GeneBank. The authors would like to acknowledge Ms. Meisong Yang (BGI Research-Shenzhen) for her assistance in Stereo-seq experiments

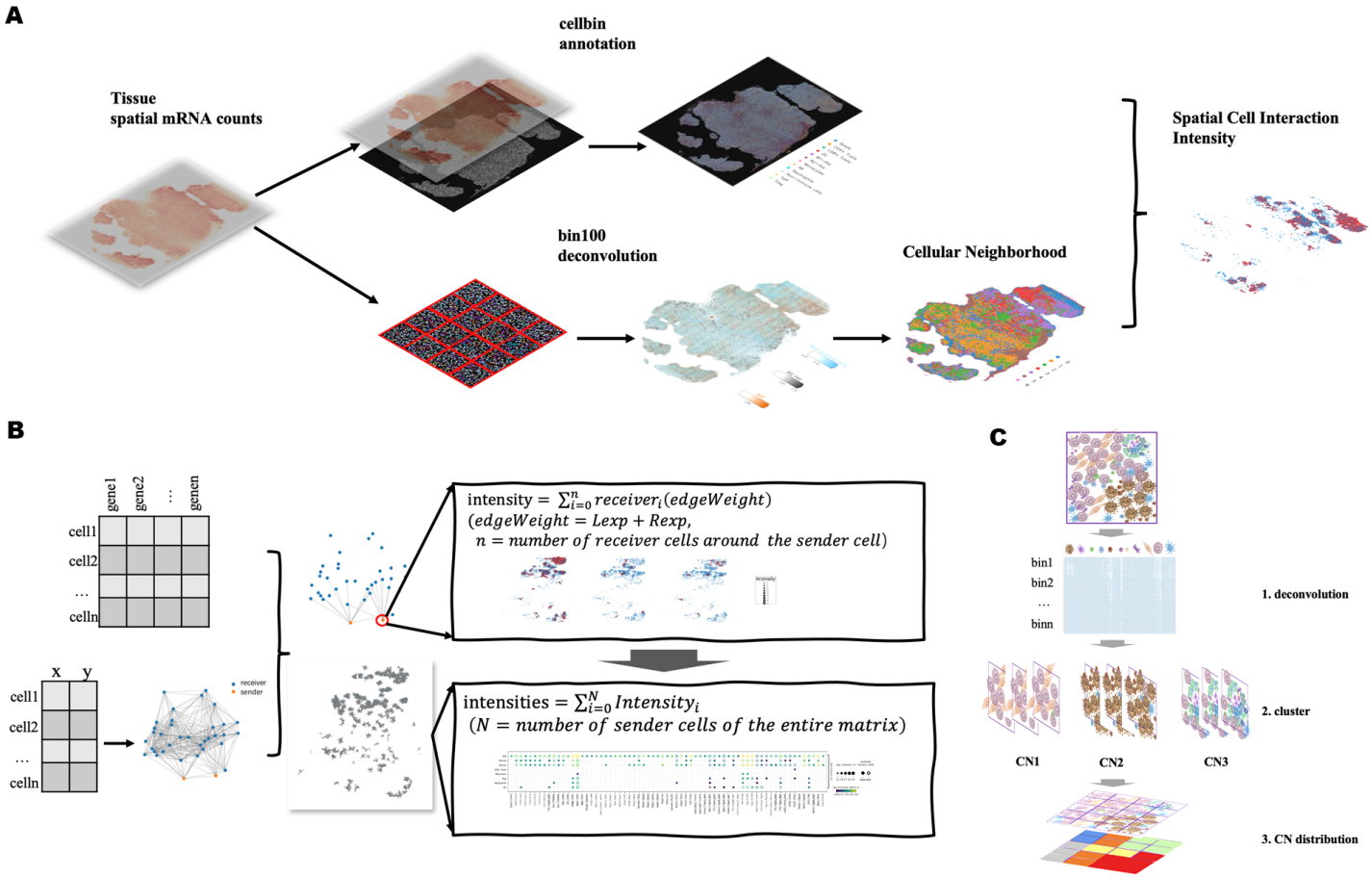
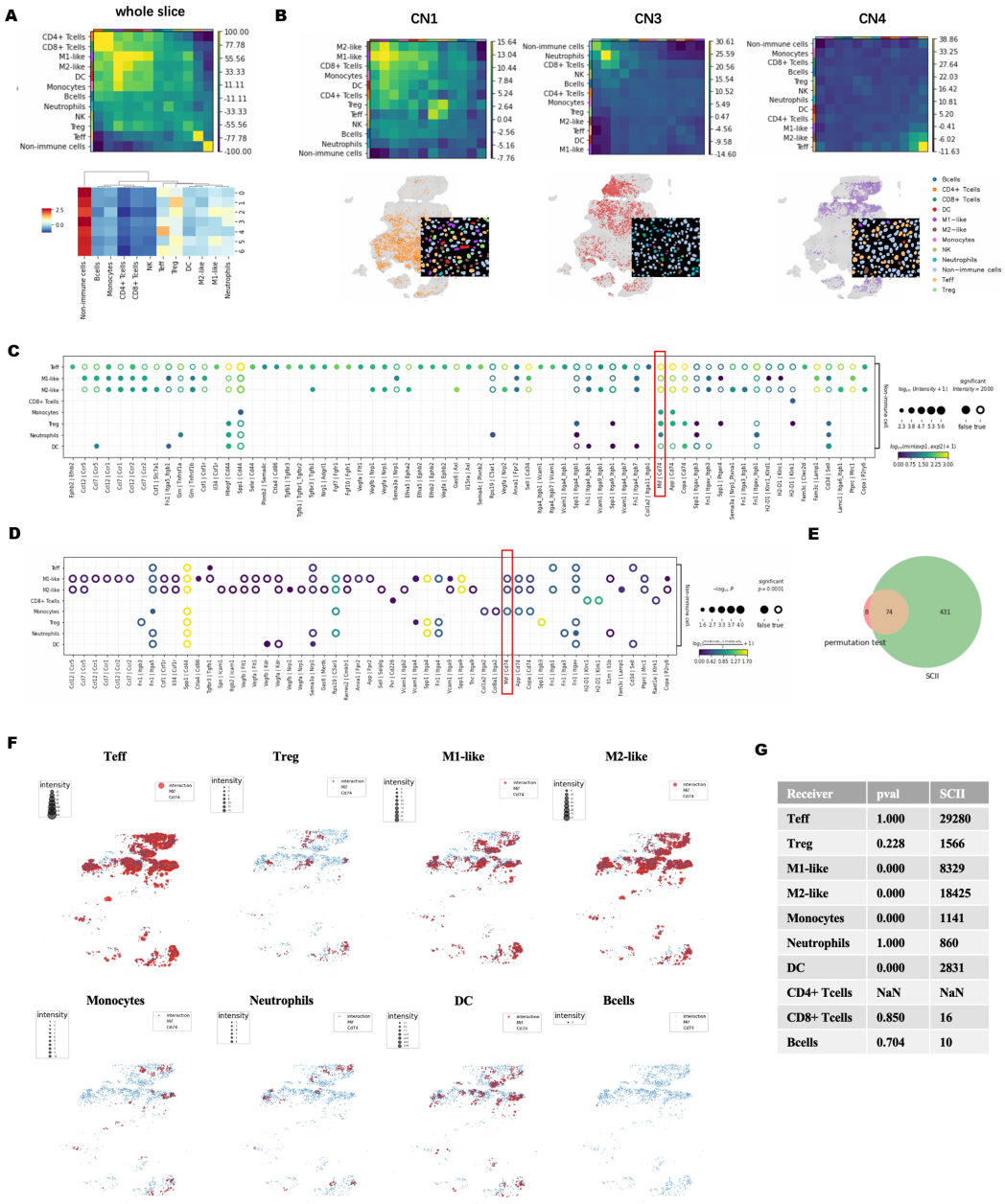


Figure 1. StereoSiTE, framework of SCII and CN A. Workflow sketch of stereoSiTE, showed data analysis process from the raw spatial transcriptomic data to spatial cell interaction intensity of specific CN region. **B.**Algorithm principle of SCII. Construct space nearest neighbor graph based on cell spatial location and assign the gene expression as weight to connected edges. Then compute intensity and intensities from the graph. **C.**Principle of CN. After binned the spatial gene expression matrix with specific window and resolving the cell composition by deconvolution, the bins were clustered to different cellular neighborhoods. Finally, show the CN distribution in situ space.



Receiver	pval	SCII
Teff	1.000	29280
Treg	0.228	1566
M1-like	0.000	8329
M2-like	0.000	18425
Monocytes	0.000	1141
Neutrophils	1.000	860
DC	0.000	2831
CD4+ Tcells	NaN	NaN
CD8+ Tcells	0.850	16
Bcells	0.704	10

Figure 2. StereoSiTE precisely and distinguishingly measure cell-to-cell communication in specific cellular neighborhood organized iTME region. A. Cell neighborhood enrichment pattern of the whole slice and cell composition of seven distinct CNs. **B.**Cell neighborhood enrichment pattern of CN1 CN3 and CN4 and their spatial distribution on the slice. The squares with black background showed the real cell composition of corresponding CN. **C.**Dot plot presented the SCII result. The size of dot represented spatial cell interaction intensity while the shade of color represented ligand-receptor expression high or low. **D.**Dot plot presented the cell-to-cell communication result computed by squidpy. The size of dot represented p values while the shade of color represented mean expression of ligand-receptor. **E.**Venn diagram between result of SCII and squidpy. **F.**Interaction intensity of non-immune cells that express Mif with different immune cells that express Cd74 in situ. The red dots represented intensity, blue dots represented Mif expression of non-immune cells and orange dots represented Cd74 expression of immune cells. **G.**The table listed the p values computed by squidpy and the spatial cell interaction intensity computed by SCII. All the corresponding sender cell was non-immune cells.

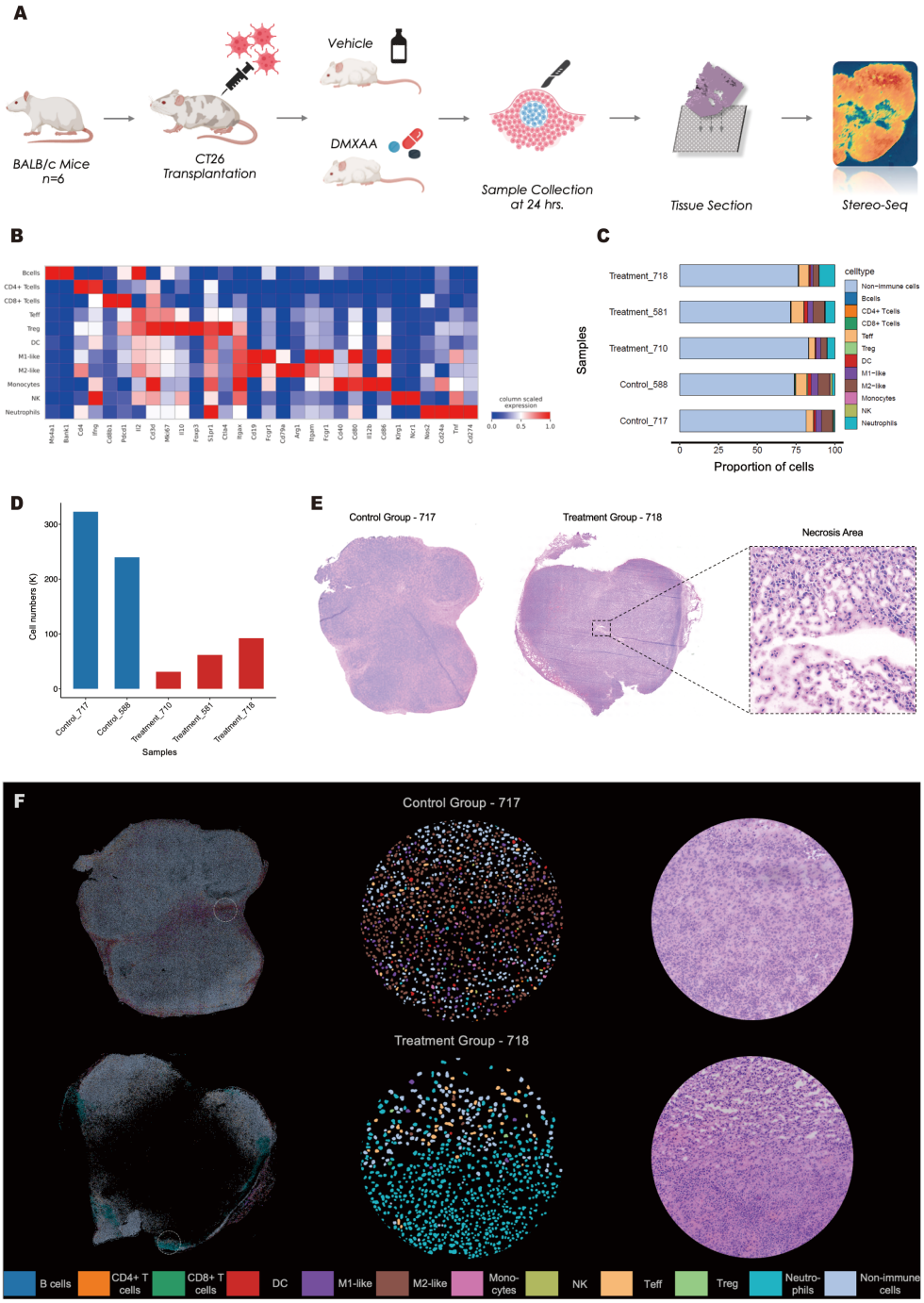


Figure 3. Spatial Transcriptomic mapping of xenograft model in situ. A. Flow chart of xenograft model construction and stereo-seq data management. **B.**Heatmap of transcriptional markers expression of 11 annotated immune cell types at resolution of single cell. **C.**Cell proportion analysis of indicated samples at single cell resolution. **D.**Cell number analysis of different samples at single cell resolution (blue stands for control group; red stands for treatment group). **E.**Adjacent H&E staining of sample 717 from control group (left) and sample 718 from treatment group (middle). H&E staining on the (right) side exhibited necrosis area enlarged from the black blank marked site. **F.**In situ visualization of annotated cell types using Stereo-seq data (Left); Enlarged image of white circle marked site exhibiting distinct cell distribution across samples (Middle); adjacent H&E staining of white circle marked site objectively displaying cell distribution (Right) .

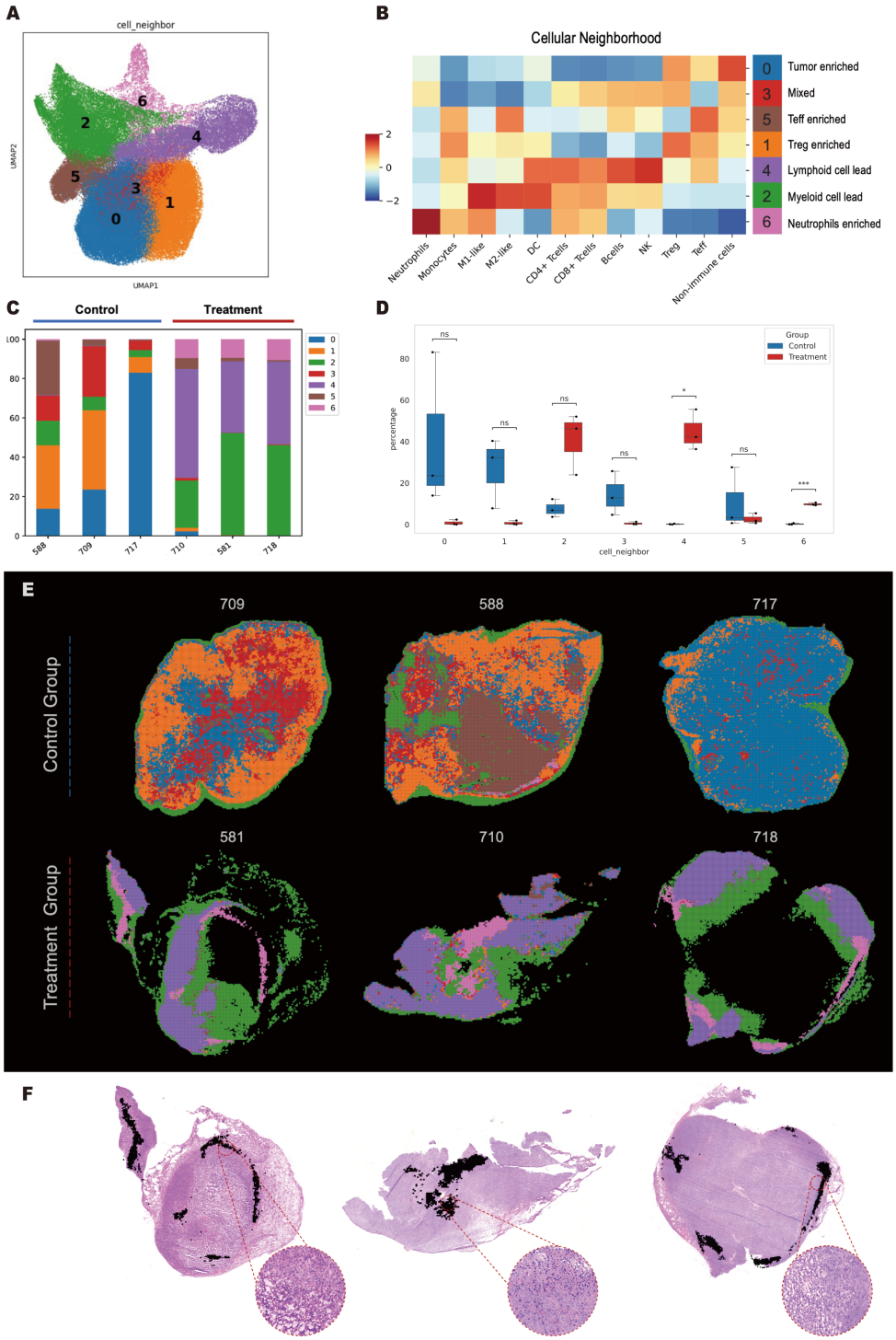


Figure 4. Organization of TME-associated cellular neighborhood. A.UMAP exhibiting the distribution of the identified 7 CN clusters. **B.**Heatmap indicating varied cell composition in different CNs (left) and confetti labeling the corresponding title for each CN (right). **C.**CN frequencies in different samples from control and treatment group, respectively. **D.**Box plot indicating the statistical variation of CN frequency across groups. **E.**In situ distribution of CNs in different groups. **F.**CN6 section (black lump indicated) on adjacent H&E staining of samples from treatment group, with enlarged image of red circle marked site displaying highly resolved H&E staining.

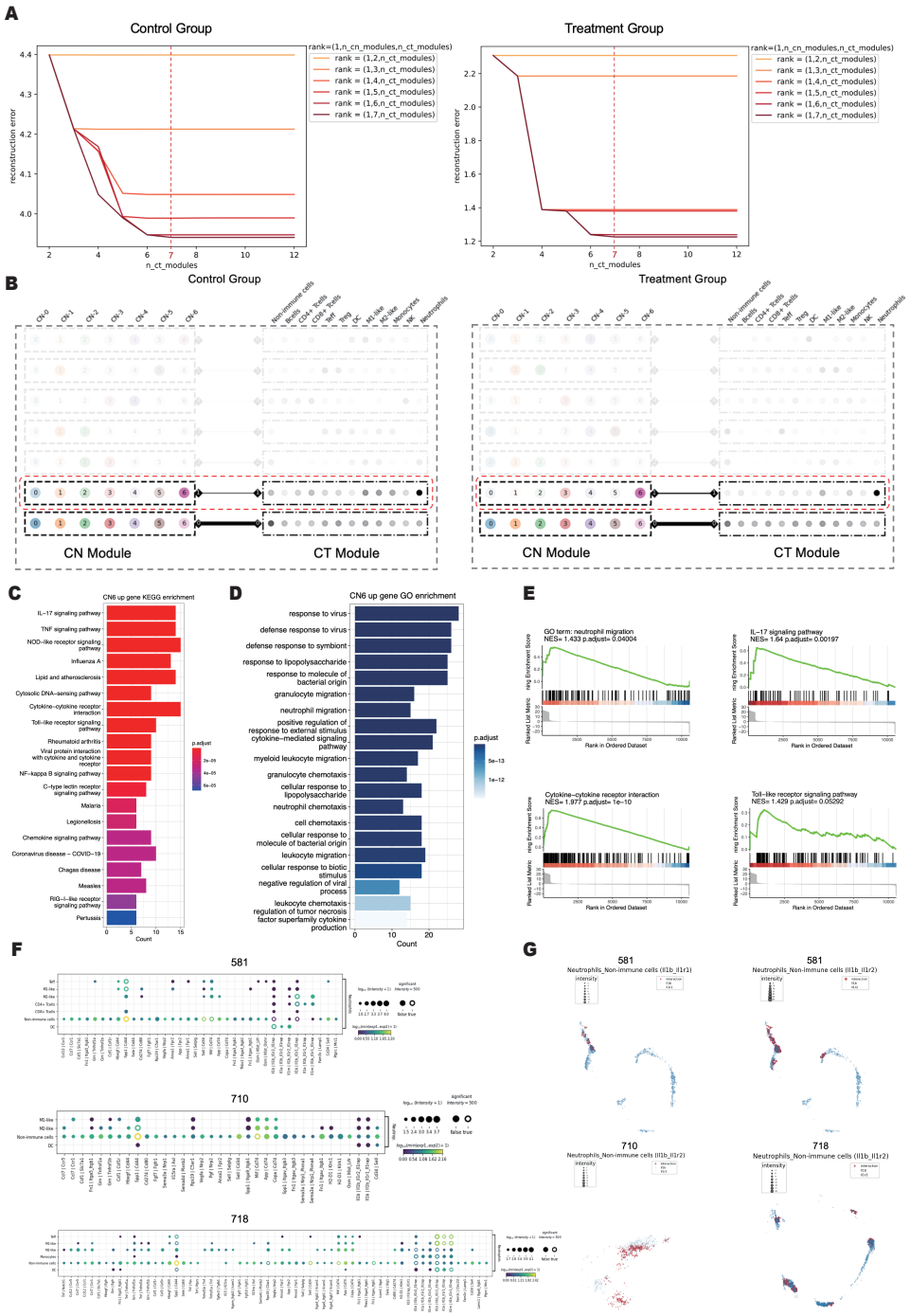


Figure 5. Deconvolution of CN of interest. A. Rank selection of Tucker tensor decomposition to stratify CN modules and CT modules. Tensor decomposition loss in different CN modules (different colors) or CT modules numbers (x axis). **B.**Decomposition results for both groups. The crosstalk extent of indicated CN and CT was represented by weight of the line linking them and weight lines under empirically recognized threshold were concealed. **C.**KEGG analysis of CN6 in control group comparing to treatment group. **D.**GO analysis of CN6 in control group comparing to treatment group. **E.**GSEA analysis validation with result obtained from KEGG and GO analysis. **F.**SCII calculation of potential ligand-receptor interaction between neutrophils and other cell types in CN6. **G.**In situ visualization of IL-1β/IL-1R between neutrophils and non-immune cells to exhibit the crosstalk coordinates.

Method and materials

Mice and cell lines

Female BALB/c mice at 6 weeks of age were obtained from GemPharmatech Co., Ltd. All mice were housed in a specific pathogen–free animal facility at GemPharmatech Co., Ltd. CT26 colon cancer cells were purchased from ATCC. Cells were cultured at 37 ºC under 5% CO2 in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin.

Xenograft tumor models and treatment

We implanted 5×10^5 cells/100 μL CT26 cells into the right flanks of BALB/c mice. When the tumors reached volume of 250-300 mm^3, we performed intratumoral injections of the STING agonist (0.5 mg/50ul/mouse, DMXAA, Vadimezan). Mice in the control group were intratumorally injected with the same volume of PBS, accordingly. Xenograft tumor samples were collected 24hours after treatment and embedded by OCT on dry ice.

Stereo-seq library preparation and sequencing

Tissue processing

Two consecutive cyro sections of 10 μm were prepared. One section was attached to glass slide and stained by H&E staining following previous protocol. The second section was adhered to the Stereo-seq chip surface and incubated at 37℃ for 3-5 minutes. Then, the sections were fixed in methanol and incubated for 40 minutes at -20℃. Stereo-seq library preparation and sequencing followed previous published protocol(18).

In situ reverse transcription

Prepared section was processed according to the Stereo-seq Transcriptomics Set User Manual (STOmics) and all reagents were from the Stereo-seq Transcriptomics T kit and Stereo-seq Library Preparation kit (STOmics). Briefly, after washed with PR rinse buffer, tissue sections placed on the chip were permeabilized at 37°C for 10 minutes. RNA released from the permeabilized tissue and captured by the DNB was reverse transcribed overnight at 42°C. After reverse transcription, tissue sections were digested with Tissue Removal buffer at 55°C for 30 minutes. The resulting cDNA was then amplified.

Amplification

The collected cDNAs were amplified with KAPA HiFi Hotstart Ready Mix (Roche, KK2602) with 0.8 μM cDNA-PCR primer. PCR reactions were performed in sequential steps as incubation at 95℃ for 5 minutes, 15 cycles at 98℃ for 20 seconds, 58℃ for 20 seconds, 72℃ for 3 minutes and a final incubation at 72℃ for 5 minutes.

Library construction and sequencing

The concentrations of the resulting PCR products were quantified by Qubit™ dsDNA Assay Kit (Thermo, Q32854). A total of 20 ng of DNA were then fragmented. Fragmented products were amplified. PCR products were purified using the AMPure XP Beads (0.6× and 0.15×), used for DNB generation and finally sequenced on MGI DNBSEQ-T7 sequencer.

Data analysis

Raw sequencing data analysis

Fastq files were generated by MGI SEQ-2000 sequencer. ssDNA image stitching, tissue cut, cell segmentation, gene expression register, and genome mapping, gene count were performed by online analysis Platform: Stereo Analysis Platform (SAP, https://uat.stomics.tech/sap/researchProject/index.html). Expression profile matrix was divided into non-overlapping bins covering an area of 100 × 100 DNBs (bin100) for further cellular neighborhood construction and functional enrichment analysis. Gene expression matrix for each cell were obtained according to spatial coordinates of cell segmentation result and corrected by Sklearn (34)gaussian mixture model. Then anndata data structure were constructed by Scanpy (35)in python 3.9 for further analysis. The binned spots with total counts less than 1000 and cells with total counts less than 100 were filtered.

Celltype annotation

We used a mouse colon cancer cell line CT26 single cell gene expression dataset (20)as a reference to deconvolute a mixture of 11 immune cell types and non-immune cells in our Stereo-seq data by Cell2location(19) with hyperparameter N_cells_per_location= 1, detection_alpha=20. The cell type with maximum abundance was assign to each cell, then cell types frequency was calculated and visualized by R package ggplot2 (36).

Cellular neighborhood construction

The tissues were binned to side-by-side windows, each with an area of 100 × 100 DNBs (bin100). Then Cell2location deconvoluted all cell types abundance of each window. According to cell composition matrix, the windows of all samples were subsequently clustered to 7 cellular neighborhoods (CNs) (3) by using KNN graph and Leiden with n_neighbors = 20, resolution=0.3. Therefore, the windows with similar cell types composition were gather together to form a microenvironment. For each CN, all windows cell types abundance were summed to calculate percentage and visualized by Python package Seaborn.

Tensor decomposition

For each group, we constructed a tensor with 3×7×12 dimensions (3 samples, 7 CNs and 12 cell types). We performed Non-negative Tucker decomposition by Tensorly (37)Python package. By calculating the decomposition losses of different combinations of the number of CN modules and CT modules, we select the suitable rank in the elbow point to performed non-negative tensor decomposition (Fig 5A). The visualization way of decomposition result refers to Schürch's article (3).

Functional enrichment analysis

We performed differential expression analysis on CN6 between treatment set and control set using the scanpy.tl.rank_genes_groups (35) function. Genes were retained when abs(logfoldchanges)>log2(1.5) and pvals_adj<0.05. The KEGG enrichment analysis, gene ontology enrichment analysis and GSEA were employed to dissect the biological function of CN6 using functions of R package ClusterProfiler (38). The top 20 significantly enriched pathways of KEGG enrichment analysis and gene ontology enrichment analysis were displayed as barplot, respectively. The interesting pathways of GSEA were shown through the gseaplot2() function.

Spatial cell interaction intensity

Firstly, we constructed the space nearest neighbor graph based on spatial coordinate of all cells, and cell pairs with distance less than 200um were connected by an edge. Then we summed up the ligand and receptor gene expression counts of the corresponding sender and receiver cell as weight which was assigned to the edge. Next, we computed the spatial cell interaction intensity of every sender cell with its linked receiver cells by summing of the edge

weight between them, defined as

$$intensity = \sum_{i=0}^{n} receiver_i(edgeWeight)$$

$$(edgeWeight = Lexp + Rexp,\ \ n = number\ of\ receiver\ cells\ around\ \ the\ sender\ cell) \qquad (1)$$

Besides, the intensities of the entire data matrix between specific sender cell type with receiver cell type equaled to the sum of all sender cell's intensity, which meant the sum of all connected edge weight. The formula below presented the computation rule.

$$intensities = \sum_{i=0}^{N} Intensity_i$$

$$(N = number\ of\ sender\ cells\ of\ the\ entire\ matrix) \qquad (2)$$

As to the complex ligand or receptor which was composed of several subunits, we selected the molecule with minimal expression to compute the SCII.

STAR

Software and Algorithms	
Stereo Analysis Platform (SAP)	https://uat.stomics.tech/sap/researchProject/index.html
Python 3.9	https://www.python.org/
Numpy 1.22.4	https://numpy.org/
Pandas 1.5.1	https://pandas.pydata.org/
Sklearn 1.0.1	https://scikit-learn.org/
Cell2location 0.1	https://cell2location.readthedocs.io/en/latest/
Scanpy 1.9.1	https://scanpy.readthedocs.io/en/stable/index.html
Tensorly 0.7.0	http://tensorly.org/stable/index.html
Seaborn 0.11.2	https://seaborn.pydata.org/index.html
R 4.2.1	https://www.r-project.org/
ggplot2 3.4.0	https://ggplot2.tidyverse.org/
ClusterProfiler 4.6.0	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html

1. Topalian SL, Drake CG, and Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell.* 2015;27(4):450-61.
2. Jackson HW, Fischer JR, Zanotelli VRT, Ali HR, Mechera R, Soysal SD, et al. The single-cell pathology

landscape of breast cancer. *Nature.* 2020;578(7796):615-20.
3. Schurch CM, Bhate SS, Barlow GL, Phillips DJ, Noti L, Zlobec I, et al. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell.* 2020;183(3):838.
4. Altorki NK, Markowitz GJ, Gao D, Port JL, Saxena A, Stiles B, et al. The lung microenvironment: an important regulator of tumour growth and metastasis. *Nat Rev Cancer.* 2019;19(1):9-31.
5. Ceelen W, Ramsay RG, Narasimhan V, Heriot AG, and De Wever O. Targeting the Tumor Microenvironment in Colorectal Peritoneal Metastases. *Trends Cancer.* 2020;6(3):236-46.
6. Ho WJ, Jaffee EM, and Zheng L. The tumour microenvironment in pancreatic cancer - clinical challenges and opportunities. *Nat Rev Clin Oncol.* 2020;17(9):527-40.
7. Chen WT, Lu A, Craessaerts K, Pavie B, Sala Frigerio C, Corthout N, et al. Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. *Cell.* 2020;182(4):976-91 e19.
8. Zhang J, Yin J, Heng Y, Xie K, Chen A, Amit I, et al. Spatiotemporal Omics-Refining the landscape of precision medicine. *Life Medicine.* 2022:lnac053.
9. Moffitt JR, Bambah-Mukku D, Eichhorn SW, Vaughn E, Shekhar K, Perez JD, et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science.* 2018;362(6416).
10. Zhang M, Eichhorn SW, Zingg B, Yao Z, Cotter K, Zeng H, et al. Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature.* 2021;598(7879):137-43.
11. Wei X, Fu S, Li H, Liu Y, Wang S, Feng W, et al. Single-cell Stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. *Science.* 2022;377(6610):eabp9444.
12. Ji AL, Rubin AJ, Thrane K, Jiang S, Reynolds DL, Meyers RM, et al. Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. *Cell.* 2020;182(6):1661-2.
13. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer discovery.* 2022;12(1):134-53.
14. Hunter MV, Moncada R, Weiss JM, Yanai I, and White RM. Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. *Nat Commun.* 2021;12(1):6278.
15. Yuan Y, and Bar-Joseph Z. GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data. *Genome Biol.* 2020;21(1):300.
16. Palla G, Spitzer H, Klein M, Fischer D, Schaar AC, Kuemmerle LB, et al. Squidpy: a scalable framework for spatial omics analysis. *Nature Methods.* 2022;19(2):171-8.
17. Dries R, Zhu Q, Dong R, Eng CL, Li H, Liu K, et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* 2021;22(1):78.
18. Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell.* 2022;185(10):1777-92 e21.
19. Kleshchevnikov V, Shmatko A, Dann E, Aivazidis A, King HW, Li T, et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nature Biotechnology.* 2022;40(5):661-71.
20. Ravirala D, Pei G, Zhao Z, and Zhang X. Comprehensive characterization of tumor immune landscape following oncolytic virotherapy by single-cell RNA sequencing. *Cancer Immunol Immunother.* 2022;71(6):1479-95.
21. Shaul ME, and Fridlender ZG. Tumour-associated neutrophils in patients with cancer. *Nature Reviews Clinical Oncology.* 2019;16(10):601-20.
22. Jaillon S, Ponzetta A, Di Mitri D, Santoni A, Bonecchi R, and Mantovani A. Neutrophil diversity and plasticity in tumour progression and therapy. *Nature Reviews Cancer.* 2020;20(9):485-503.
23. Coffelt SB, Wellenstein MD, and de Visser KE. Neutrophils in cancer: neutral no more. *Nature Reviews Cancer.* 2016;16(7):431-46.
24. Hore V, Viñuela A, Buil A, Knight J, McCarthy MI, Small K, et al. Tensor decomposition for multiple-tissue gene expression experiments. *Nature Genetics.* 2016;48(9):1094-100.
25. Di Mitri D, Toso A, Chen JJ, Sarti M, Pinton S, Jost TR, et al. Tumour-infiltrating Gr-1+ myeloid cells antagonize senescence in cancer. *Nature.* 2014;515(7525):134-7.
26. Kumar V, Abbas AK, and Aster JC. *Robbins and Cotran pathologic basis of disease.* Philadelphia, PA: Elsevier/Saunders; 2015.
27. Grunwald BT, Devisme A, Andrieux G, Vyas F, Aliar K, McCloskey CW, et al. Spatially confined sub-tumor microenvironments in pancreatic cancer. *Cell.* 2021;184(22):5577-92 e18.
28. Tian L, Chen F, and Macosko EZ. The expanding vistas of spatial transcriptomics. *Nat Biotechnol.* 2022.
29. Bove A, Gradeci D, Fujita Y, Banerjee S, Charras G, and Lowe AR. Local cellular neighborhood controls proliferation in cell competition. *Molecular Biology of the Cell.* 2017;28(23):3215-28.
30. Reshef YA, Rumker L, Kang JB, Nathan A, Korsunsky I, Asgari S, et al. Co-varying neighborhood analysis identifies cell populations associated with phenotypes of interest from single-cell transcriptomics.

Nature Biotechnology. 2022;40(3):355-63.
31. Hedrick CC, and Malanchi I. Neutrophils in cancer: heterogeneous and multifaceted. *Nature reviews Immunology.* 2022;22(3):173-87.
32. Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med.* 2018;24(8):1277-89.
33. Longo SK, Guo MG, Ji AL, and Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nature Reviews Genetics.* 2021;22(10):627-44.
34. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research.* 2011;12:2825-30.
35. Wolf FA, Angerer P, and Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* 2018;19(1):15.
36. Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York; 2016.
37. Kossaifi J, Panagakis Y, Anandkumar A, and Pantic M. TensorLy: Tensor Learning in Python. *Journal of Machine Learning Research.* 2019;20:1-6.
38. Yu G, Wang L-G, Han Y, and He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology.* 2012;16(5):284-7.