

RESEARCH ARTICLE

# scSDNE: A semi-supervised method for inferring cell-cell interactions based on graph embedding

Chenchen Jia<sup>1</sup>, Haiyun Wang<sup>2</sup>, Jianping Zhao<sup>1, \*</sup>, Junfeng Xia<sup>1, 3, \*</sup>, Chunhou Zheng<sup>3</sup>

**1** College of Mathematics and System Sciences, Xinjiang University, Urumqi, China, **2** School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, China, **3** School of Physical Science and Information Technology, Anhui University, Hefei, China

\* [jpzha@xju.edu.cn](mailto:jpzha@xju.edu.cn) (JZ); [jfxia@ahu.edu.cn](mailto:jfxia@ahu.edu.cn) (JX)



## Abstract

As a fundamental characteristic of multicellular organisms, cell-cell communication is achieved through ligand-receptor (L-R) interactions, enabling the exchange of information and revealing the diversity of biological processes and cellular functions. To gain a comprehensive understanding of these complex interaction mechanisms, we constructed a manually curated L-R interaction database and developed a semi-supervised graph embedding model called scSDNE for inferring cell-cell interactions mediated by L-R interactions. scSDNE model utilizes the power of deep learning to map genes from interacting cells into a shared latent space, allowing for a nuanced representation of their relationships. Leveraging the prior information provided by database, scSDNE can infer significant L-R pairs involved in intercellular communication. Experiments on real single-cell RNA sequencing (scRNA-seq) datasets demonstrate that our method detects interactions with a high degree of reliability compared with other methods. More importantly, the model integrates gene regulation information within cells to enhance the accuracy and biological interpretability of the inferences. Our method provides a more comprehensive view of cell-cell interactions, offering new insights into complex intercellular communication.

## OPEN ACCESS

**Citation:** Jia C, Wang H, Zhao J, Xia J, Zheng C (2025) scSDNE: A semi-supervised method for inferring cell-cell interactions based on graph embedding. PLoS Comput Biol 21(5): e1013027. <https://doi.org/10.1371/journal.pcbi.1013027>

**Editor:** Stacey D. Finley, University of Southern California, UNITED STATES OF AMERICA

**Received:** September 12, 2024

**Accepted:** April 4, 2025

**Published:** May 7, 2025

**Copyright:** © 2025 Jia et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** All relevant data and code are available at <https://github.com/jcc457/scSDNE>

**Funding:** JZ is funded by National Natural Science Foundation of China (62362062), Talent Program of Xinjiang Autonomous Region-Youth Outstanding Talent and Youth

## Author summary

In multicellular organisms, effective communication between cells is essential for coordinating various biological processes, primarily mediated through ligand-receptor interactions. However, most existing methods inadequately consider the impact of intracellular gene regulatory networks, resulting in insufficient capture of global information regarding intercellular interactions and limiting understanding of cellular communication mechanisms. To address this limitation, we propose scSDNE, a semi-supervised graph embedding model designed to infer crucial ligand-receptor pairs that facilitate intercellular communication. By leveraging a manually curated ligand-receptor interaction database and advanced

Innovative Talent (2023TSYCCX0104) and Multimodal Major Chronic Disease Prevention and Control Science and Engineering Research Project (MCD-2023-1-15). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

deep learning techniques, scSDNE effectively maps the gene expression data of interacting cells into a shared latent space, thereby enhancing the representation of their relationships. Experimental results demonstrate that scSDNE reliably detects intercellular interactions and improves the biological interpretability of inference results by integrating gene regulatory information. In conclusion, scSDNE can provide valuable insights for complex intercellular communication and understanding disease pathogenesis.

## Introduction

Every cell in multicellular organisms exists within a complex signaling environment [1]. Through intercellular communication and coordination, they collectively accomplish various intricate biological tasks, from early development to tissue and organ maturation [2,3]. Intercellular communication relies extensively on the interaction between ligands expressed by sending cells and homologous receptors expressed by receiving cells [4,5]. Therefore, accurately identifying and analyzing ligand-receptor (L-R) interactions is essential for understanding cell behavior and responses to neighboring cells [1]. Single-cell RNA sequencing (scRNA-seq) is a powerful technique for exploring tissue heterogeneity [6], providing unprecedented resolution to reveal cell diversity and facilitate the exploration of cell-to-cell communication and interactions [7].

Recently, various methods have been developed to infer intercellular communication from scRNA-seq data, such as CellCall [3], CellChat [8], CellPhoneDB [9], NATMI [10], and NicheNet [11]. These methods primarily focus on the expression strength and specificity of L-R pairs. However, this expression-dependent approach has limitations, including the inability to detect certain stable or low-abundance receptor transcripts and the failure to comprehensively reflect changes in intracellular signaling pathways [12]. To utilize information within the receiving cell for constructing communication networks, some methods attempt to integrate insights from L-R interactions and intracellular signaling pathways. For instance, CellCall combines cell-to-cell communication with intracellular transcription factor expression to analyze receptor pathway changes. CINS [13] identifies key ligands and targets involved in interactions between cell types by employing a regression model based on a matrix of ligand-target interactions. Nevertheless, they fail to account for intracellular gene interactions, which may lead to missed biologically significant interactions.

On the other hand, the absence of spatial information in scRNA-seq data restricts its ability to study cell communication in tissues with clear spatial structures [14]. As a result, several approaches have emerged based on spatial transcriptome data. For example, HoloNet [15] infers spatial communication networks across defined regions using a multi-view network approach. Other methods, such as SpaOTsc [16], stLearn [17], and COMMOT [18], integrate scRNA-seq data with spatial imaging data to ascertain the spatial positioning of cells within tissues, thereby inferring

cell-cell communication [16–19]. Although spatial transcriptomics offers a new perspective for analyzing cellular communication, its limited resolution and small sample sizes [20] present challenges in thoroughly studying intercellular interactions. In comparison, scRNA-seq is a more mature technique that offers comprehensive information on cell types and gene expression data, facilitating a better understanding of cell-cell interactions and key signaling pathways.

In this context, we developed scSDNE, a semi-supervised graph embedding model designed to infer L-R pairs while incorporating intracellular gene regulatory information. The model takes gene expression data from interacting cells as input and integrates the L-R interaction database (LRdb) to identify significant intercellular communications. scSDNE constructs a weighted adjacency matrix that combines gene regulatory and L-R interaction information. It then employs Structure Deep Network Embedding (SDNE) [21] to capture both first-order and second-order similarities of the network, effectively embedding genes into a shared latent space. By analyzing the distances between L-R genes in the latent space, significant L-R pairs mediating communication between interacting cells can be inferred. Furthermore, experimental results on human atopic dermatitis (AD), gastric cancer, and hepatocellular carcinoma (HCC), demonstrate the model's efficacy in elucidating the overall characteristics of cell communication.

## Results

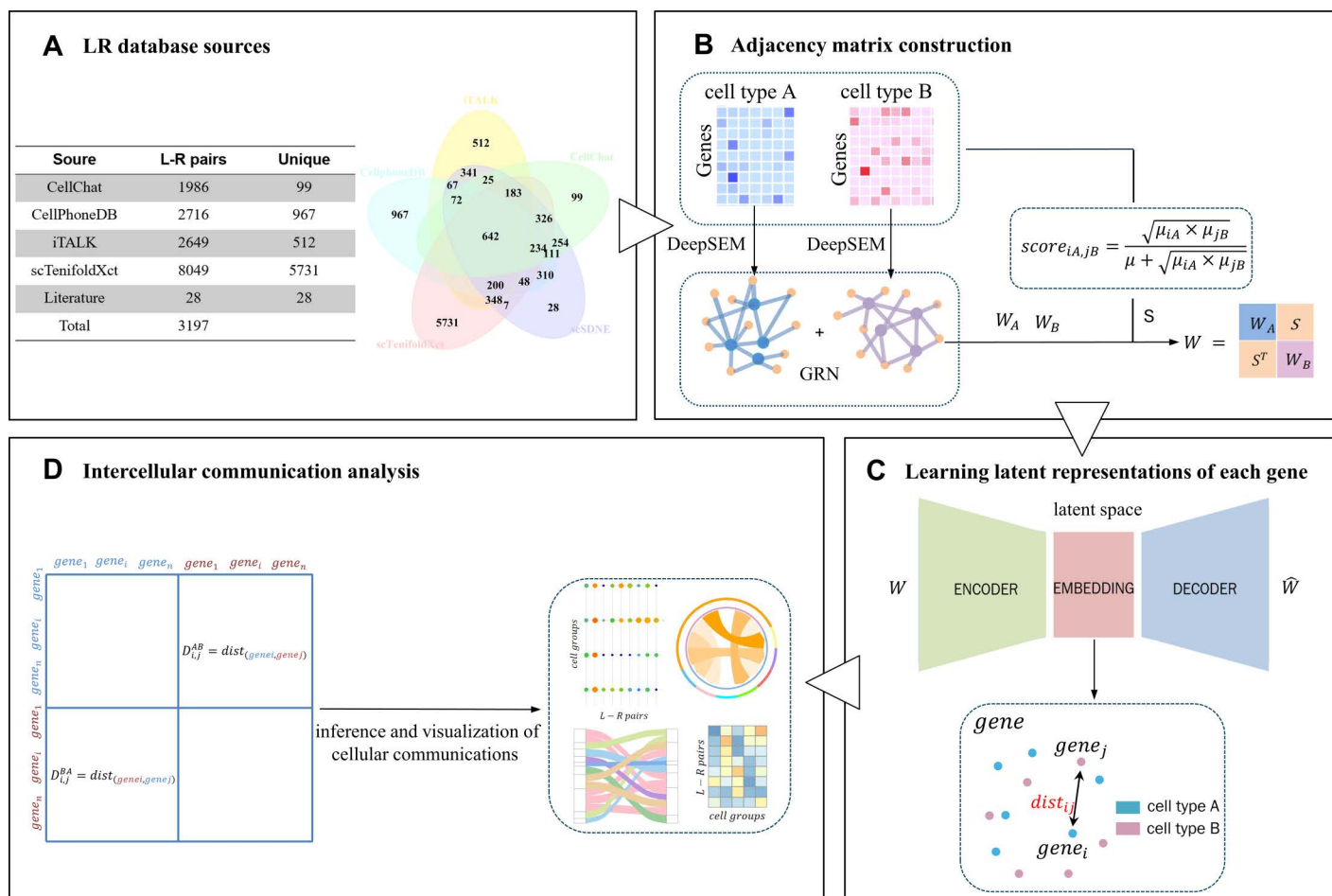
The core algorithm and the underlying intercellular communication model of scSDNE are depicted in Fig 1. To construct the joint similarity matrix, scSDNE first utilizes a gene expression matrix with annotated cell types to calculate the gene regulatory network adjacency matrix for interacting cell types separately. Thereafter, interaction scores between cell types are computed based on a defined scoring mechanism. This joint similarity matrix  $W$  serves as a weighted adjacency matrix input for the graph embedding model SDNE, enabling the extraction of gene coordinate representations within a low-dimensional latent space. Finally, by analyzing the distances between genes in this latent space, significant L-R pairs between interacting cell types are inferred in conjunction with the previously compiled LRdb.

### Inferring cell-cell communications between the human lesional skin cells

AD is a common inflammatory skin disease characterized by a complex pathogenesis involving immune cells and epidermal abnormalities [22]. Previous studies have identified specific skin cell expression of chemokines from fibroblasts to immune cells [23], revealing a potential role in fibroblasts transmitting to immune cells. Therefore, we applied scSDNE to lesional skin scRNA-seq data from AD patients to detect cell communication. The inflammatory skin datasets encompass twelve distinct cell sub-types, including four fibroblast sub-types (APOE + FIB, FBN1 + FIB, COL11A + FIB, and Inflam. FIB), four dendritic cell sub-types (cDC1, cDC2, LC and Inflam. DC), as well as four T cell sub-types (TC, Inflam. TC, CD40LG + TC and NKT). Given that the hallmark of AD lesions involves unique interactions between Inflam. FIB and immune cells [22], we focused on analyzing Inflam. FIB as sending cells and their interactions with other immune cells.

The interactions identified when Inflam. FIB act as sending cells concentrate on 3 key ligand genes: CCL19, CXCL12, and THBS2 (Fig 2A, 2C, and 2D). Specifically, CCL19 plays a crucial role in lymphocyte recirculation, homing, and migration to secondary lymphoid organs. The significant interaction between CCL19 and its receptor CCR7 highlights the regulatory effects of Inflam. FIB on lymphocytes and modulates type 2 inflammation involving Th2 cells, a CD4 + T cell subset associated with inflammation [24]. Furthermore, CXCL12 and its receptor CXCR4 exhibit high activity in lesional skin. Notably, a study by Sun et al. [25] observed that in AD, CXCL12 acts as a chemokine to recruit inflammatory cells, facilitating their migration regions by binding to CXCR4, thereby activating these cells to release inflammatory cytokines and accelerating the progression of inflammation [26].

By examining the roles of ligands THBS2 and THBS1, we found that they effectively inhibit angiogenesis through their interaction with the receptor CD36 [27, 28]. Moreover, interleukin-13 (IL-13) is closely associated with disease severity [29] and demonstrates a positive correlation with inflammatory processes. In Fig 2B, the expression levels of IL-13 are significantly lower in the third sample compared to the other samples, while expression of THBS2 is markedly elevated



**Fig 1. Overview of scSDNE.** (A) Overview of the LRdb. (B) Construction of adjacency matrix (including adjacency matrices of gene regulatory networks and crosstalk score matrices of cell types). (C) Learning of latent representations for each gene pair through graph embedding. (D) Detection and visualization of significant L-R pairs based on the distances observed in the latent space.

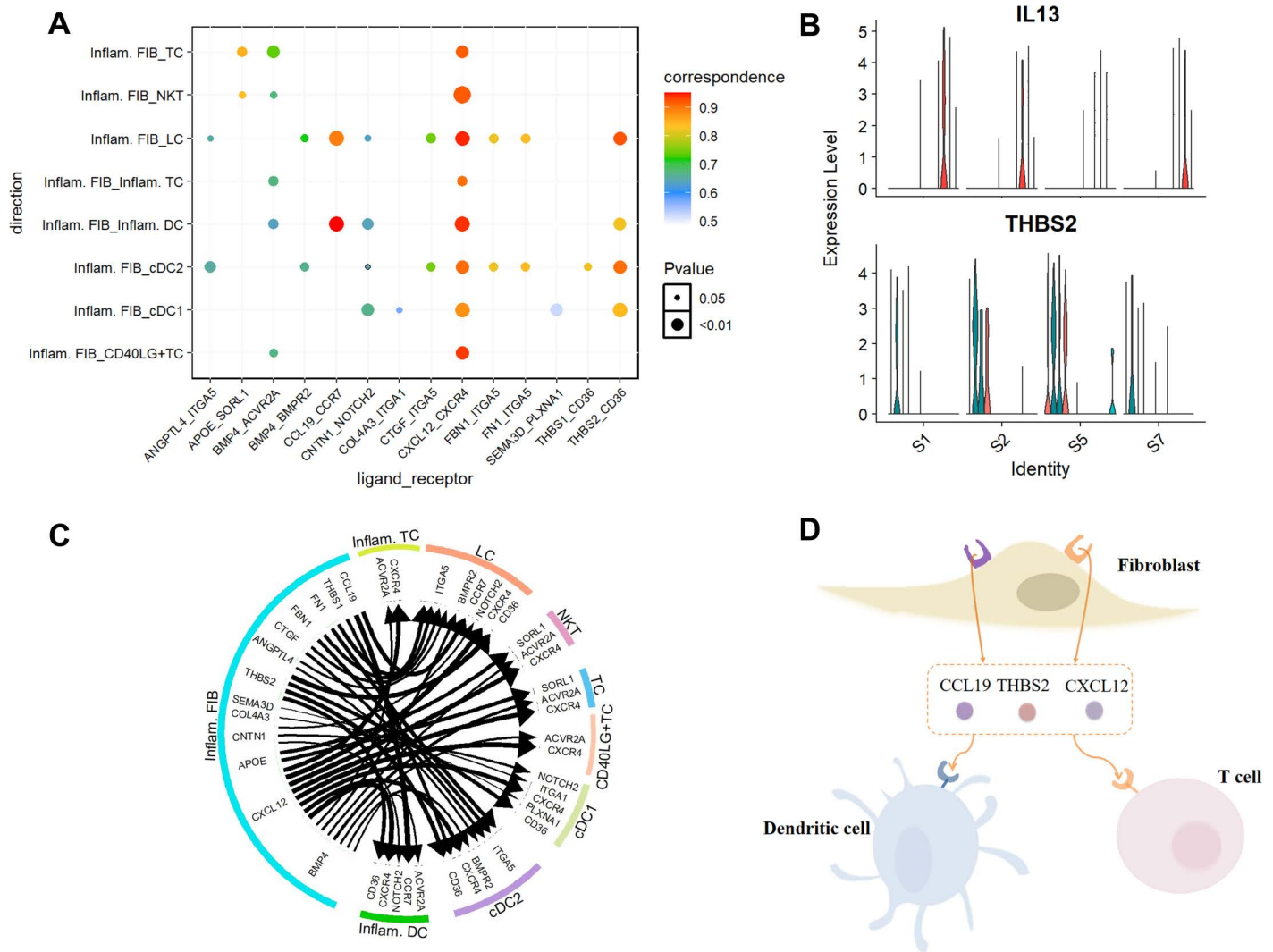
<https://doi.org/10.1371/journal.pcbi.1013027.g001>

in this sample. This finding suggests a potential relationship between the down-regulation of IL-13 and the up-regulation of THBS2 expression under specific conditions, subsequently influencing the inflammatory processes and angiogenesis associated with AD. Therefore, further exploration of the interaction mechanism between IL-13 and THBS2 is essential for enhancing our understanding of the pathophysiological processes in AD and for the development of novel therapeutic strategies.

### Revealing cellular crosstalk in tumor-adjacent tissues and HCC tissues

HCC, one of the most malignant cancers, exhibits a highly heterogeneous complex ecosystem characterized by complex intercellular communication between different cell types [30]. After processing data from tumor-adjacent and primary tumor tissues of 5 HCC patients in GSE149614, we identified a total of 8 cell types: T cells, B cells, NK cells, plasma cells, hepatocytes (malignant cells in HCC), dendritic cells (DC), Mono/Macro, and mast cells (Fig 3A) [31,32].

Upon further analysis, we noted a significant reduction in the proportion of T cells and NK cells within tumor tissues decreased compared to adjacent tissues, whereas Mono/Macro and malignant cells constituted a higher proportion of



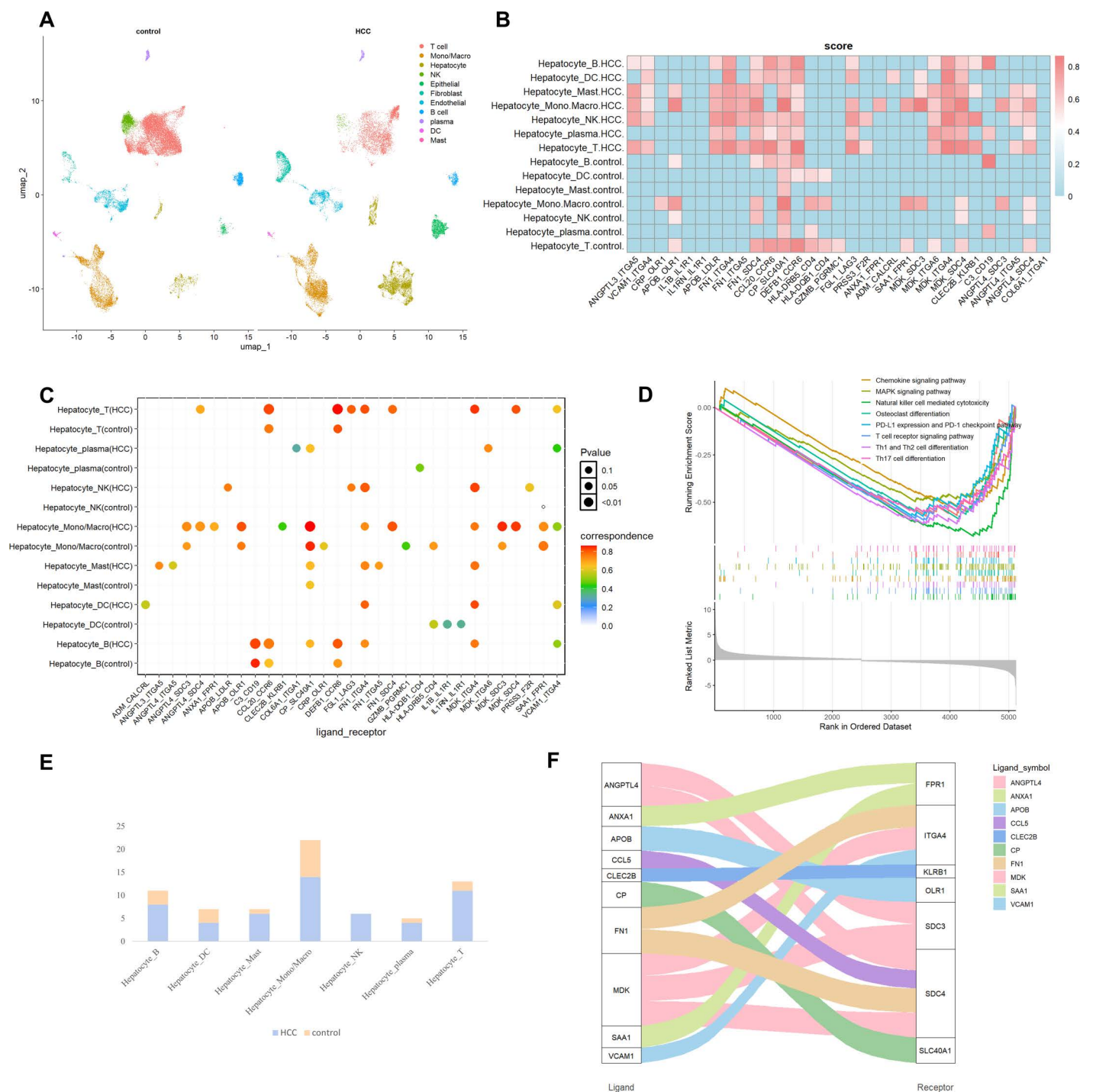
**Fig 2. Identification of intercellular communication in diseased human skin.** (A) Dot plot displays the predicted interactions between Inflam. FIB and the specified immune cell types. The color of the points reflects the communication probability, while the size of the points represents the calculated p-value. Blank areas signify a communication probability of zero. (B) Violin plot shows the expression levels of genes IL-13 and THBS2 across samples. (C) Circos plot depicts intercellular communication from Inflam. FIB to other cell types. The arrow points from the ligands in the sending cells to the receptors in the receiving cells. The thickness of the line and the size of the arrow reflect the expression of the ligands and receptors, respectively. (D) Illustration of representative ligands from Inflam. FIB to other cell types.

<https://doi.org/10.1371/journal.pcbi.1013027.g002>

the immune component in the primary tumor. This finding reveals that as HCC progresses, plenty of immune cells are recruited to the liver and interact closely with stromal cells to jointly construct an active immune micro-environment, which is of great significance for the occurrence and development of liver cancer [31]. To delve into mechanisms of interaction between the liver and immune cells, we selected highly variable genes and applied the scSDNE for in-depth analysis.

Intercellular interactions mediated by ligands FN1 and MDK, along with their receptors ITGA4, ITGA5, and ITGA6, have been extensively documented in HCC tissues, but rarely identified in normal tissues (Fig 3B and 3C). As a critical glycoprotein, FN1 has not only been confirmed in previous studies to promote metastasis, angiogenesis and proliferation of cancer cells, but also can bind to ITGA5 and ITGB1 to trigger the recruitment and activation of signaling pathway-related





**Fig 3. Comparison analysis of intercellular communication between HCC and tumor-adjacent tissues.** (A) Overview of the cell clusters derived from scRNA-seq data of tumor-adjacent and HCC tissues (UMAP). (B) Heatmap illustrates intercellular communication from hepatocytes to other immune cell types (normalized score greater than 0.5). (C) Dot plot shows the predicted interactions between hepatocyte and the specified immune cell types in both HCC and tumor-adjacent tissues. (D) Enrichment analysis of the KEGG pathways. (E) Comparison of the number of significant L-R pairs from hepatocytes to immune cell types. (F) Sankey plot represents intercellular communication from hepatocytes to Mono/Macro in HCC where the thickness of the connecting bands reflects the intensity of L-R interactions.

<https://doi.org/10.1371/journal.pcbi.1013027.g003>

proteins, including FAK/Src complex, which exerts a profound impact on tumor progression [33–35]. It has been reported that MDK interacts with the ITGA4 and ITGA6 receptors, activating a series of downstream signaling cascades that significantly enhance cancer cell growth, migration, metastasis, and angiogenesis, thereby further exacerbating the progression of HCC [36]. Notably, FGL1-LAG3 is highly expressed in HCC, particularly in T and NK cells. Multiple studies have indicated that FGL1 binding to LAG3 impairs T cell function and promotes immune escape, thereby suppressing immune responses [37–39]. This observation aligns with Gene Set Enrichment Analysis (GSEA) results suggesting the down-regulation of T/NK cell-related immune pathways in HCC, underscoring the critical role of FGL1-LAG3 in immune evasion in HCC (Fig 3D).

As illustrated in Fig 3E, there is extensive intercellular communication between hepatocytes/malignant cells and immune cells, especially Mono/Macro, with this interaction being markedly more pronounced in HCC. This phenomenon prompts us to further investigate the interaction between malignant cells and Mono/Macro within tumor tissues. Fig 3F indicates that malignant cells primarily communicate with Mono/Macro through receptor SDC4, particularly involving L-R pairs such as ANGPTL4-SDC4, FD1-SDC4, and MDK-SDC4, which are absent in adjacent tumor tissues. This suggests that SDC4, as a key endogenous membrane receptor, may exert a crucial role in tumor initiation and development by modulating cellular adhesion and migration in various cancers through its interactions with ligands [40–42]. Additionally, scSDNE successfully detected that receptor ITGA4 in Mono/Macro forms another important set of L-R pairs with the ligands FN1, MDK, VCAM1 expressed by malignant cells (Fig 3F). Notably, VCAM1 is strongly associated with malignant tumor development, as confirmed by multiple studies [43,44]. These observations further underscore the accuracy of scSDNE.

### Analysis of cell communication in the human lymph node microenvironment

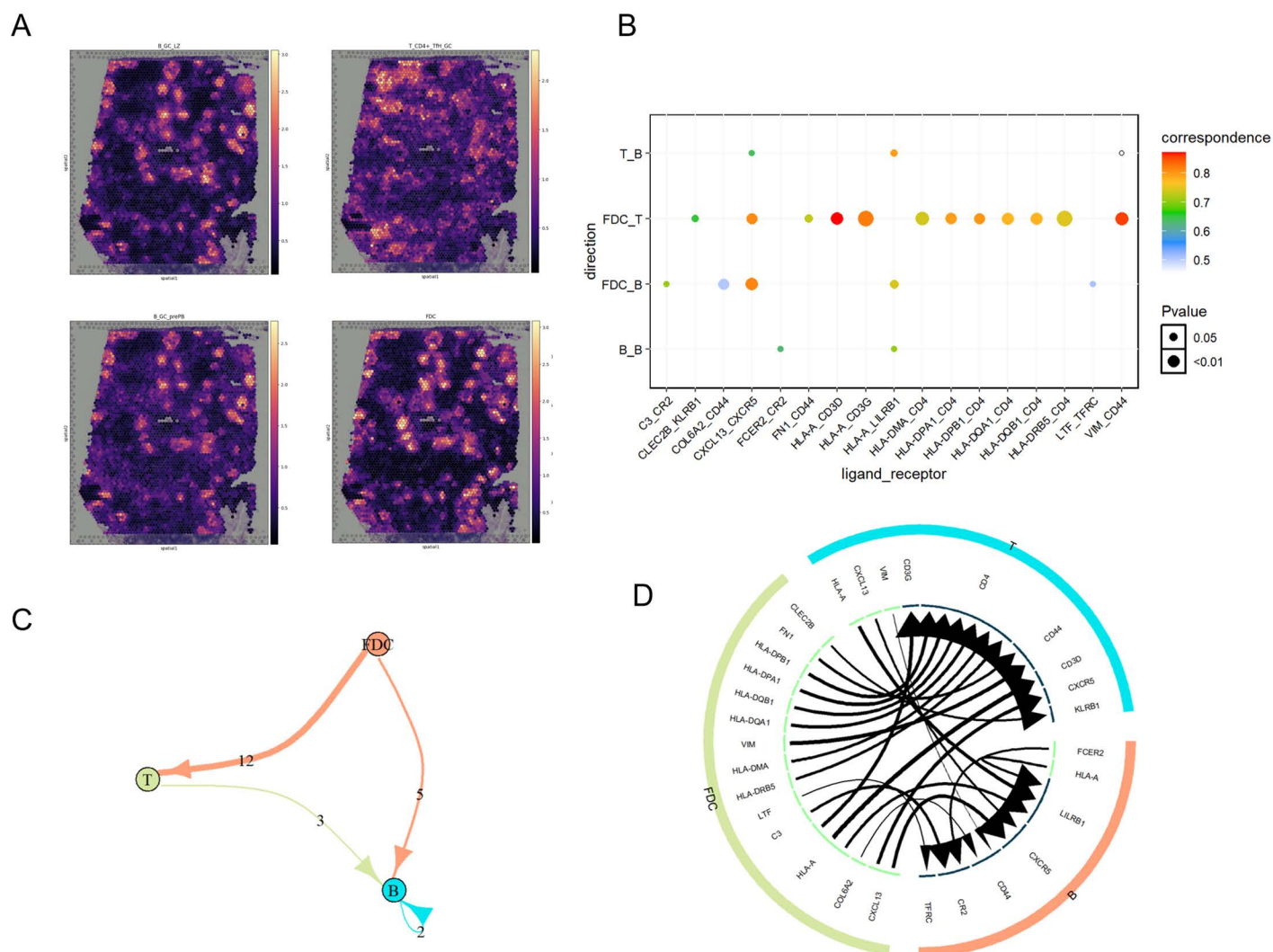
The human lymph node is characterized by a dynamic micro-environment with many spatially intermingled cell populations [45]. We used cell2location to integrate the Visium human lymph node datasets from 10x Genomics with the scRNA-seq datasets from human secondary lymphoid organs. According to the analysis by cell2location, in the germinal center (GC) light zone, B cells are selected by T follicular helper (Tfh) cells and follicular dendritic cells (FDCs) to differentiate into antibody-producing plasma cells [45]. Therefore, we chose to study B cells, T cells, and FDCs enriched in the GCs to investigate the cell communication during the B cell differentiation process (Fig 4A).

The results show that the number of L-R interactions between FDCs and T cells is most significant, with the majority mediated by genes encoding MHC II molecules (Fig 4B and 4D). The study indicated that There are significant interactions between human MHC II molecules encoded by the HLA-DP, HLA-DQ, and HLA-DR genes on FDCs and T cells [46]. CD4+ T cells can bind to antigens on MHC class II molecules, thereby activating B cells, macrophages, and other T cells [47,48]. This process not only enhances the activity of other immune cells but also promotes the development of memory T cells and B cells.

Additionally, the interaction between CXCL13 and CXCR5 among cell types in this region was generally significant (Fig 4B and 4C). The study demonstrated that CXCL13 plays a crucial role in coordinating cell migration within different regions of secondary lymphoid organs [49], primarily through its receptor CXCR5. CXCL13-CXCR5 guides B cell migration to lymphoid follicles, promotes the formation of lymphoid follicles, and participates in B cell and T cell-mediated immune responses. In the light zone, B cells from the dark zone can acquire antigens from FDCs, assisting in the formation of Tfh cells, and selectively re-enter the dark zone or exit the germinal center by differentiating into long-lived plasma cells and memory B cells [49, 50].

### Application of scSDNE in gastric cancer TME and comparative analysis with various methods

The tumor micro-environment (TME) in gastric cancer plays a critical role in regulating tumor progression through complex cellular interactions [51,52]. Therefore, we applied scSDNE to analyze the scRNA-seq datasets of gastric cancer. After processing the superficial and deep tumor invasion data from five patients in GSE167297, a total of 10 different cell types



**Fig 4. A case study on the application of scSDNE in the human lymph node microenvironment.** (A) Spatial plots show cell abundance (color intensity) for the specified cell types. (B) Dot plot shows the predicted interactions between the immune cell types. (C) The edge width represents the strength of intercellular communication between the cell types. (D) Circos plot. The arrow points from the ligands in the sending cells to the receptors in the receiving cells.

<https://doi.org/10.1371/journal.pcbi.1013027.g004>

were identified (S1 and S2 Figs). Cancer-associated fibroblasts (CAFs) are a crucial component of the TME in gastric cancer, contributing significantly to the recruitment of immune-suppressive cells and facilitating immune evasion [53–55]. Therefore, our analysis primarily focuses on the interactions of fibroblasts, which serve as sender cells, with other cell types within the TME.

scSDNE detected 76 L-R pairs between fibroblasts and immune cells (S3 Fig). Among these, collagen family genes and FN1 highly expressed in CAFs interact with the CD44 receptor on immune cells, driving cancer progression [56–59]. Additionally, laminin encoded by LAMA4, LAMB1, and LAMB2 mediate adhesion through Integrin binding—specifically the Integrin  $\alpha6\beta4$ , composed of integrin  $\alpha6$  (ITGA6) and Integrin  $\beta4$  (ITGB4)—is suggested to be essential for tumor development and progression, promoting pro-cancer signaling pathways [60].



As shown in [Fig 5A](#), macrophages serve as the primary recipients of signals from fibroblasts. We compared the accuracy and effectiveness of scSDNE with several established methods (CellChat, CellPhoneDB, iTALK [61], and scTenifoldXct) on these two cell types. As shown in [Fig 5B](#), using their respective databases, scSDNE identified 14 unique communications that were not detected by any other method, while CellPhoneDB, iTALK and scTenifoldXct identified 6, 17, and 23 particular communications; CellChat did not identify any unique communications. Significantly, the predictions from scTenifoldXct displayed considerable divergence from those of other methods, likely due to the reduced overlap of its L-R database with others, thereby highlighting the critical impact of the selected database on prediction outcomes ([S4 Fig](#)).

To facilitate a fair comparison of the predictive performance across these methods, we utilized a unified LR database as a benchmark for further analysis. scSDNE, CellChat, CellPhoneDB, iTALK and scTenifoldXct identified 35, 15, 35, 34, and 26 intercellular communications from fibroblasts to macrophages, respectively. [Fig 5C](#) demonstrates that when all methods utilized the same database, the consistency of the detected L-R pairs significantly improved. Among the methods, CellChat and scTenifoldXct identified fewer L-R pairs compared to others and did not reveal any specific cell-cell communications. Importantly, scSDNE not only demonstrated superiority in the number of identified L-R pairs, but also showed a considerable overlap in detection results with those of other methods.

Furthermore, using relevant literature reports as evaluation criterion, we compared the predictive performance of these methods using ROC curves ([Fig 5D](#)). The results indicated that scSDNE achieved the highest area under the ROC curve (AUC = 0.82), indicating its comprehensiveness and accuracy in capturing cell-cell interactions.

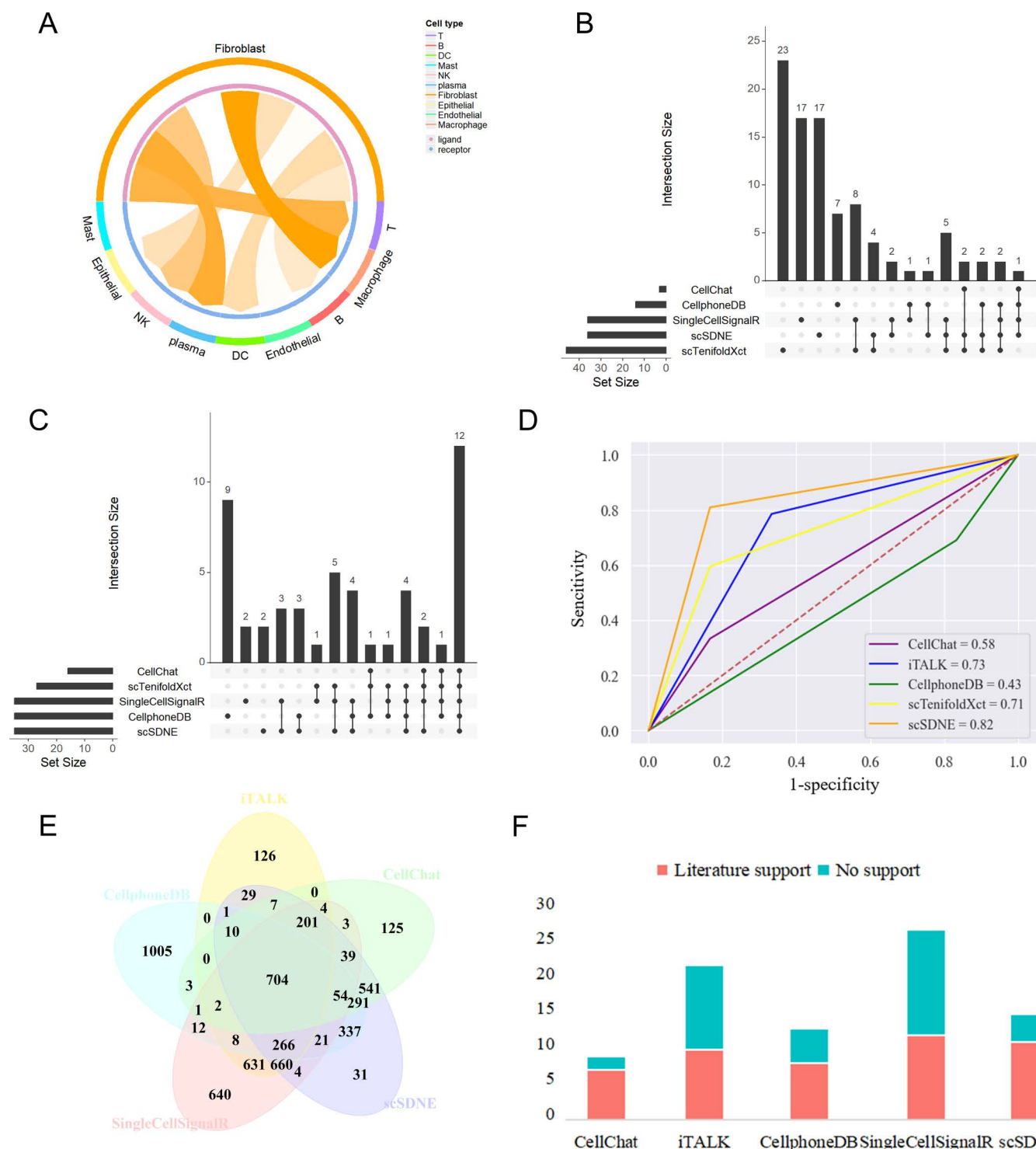
To further assess effectiveness, we utilized scSDNE in conjunction with five different databases (scSDNE, CellChat, CellPhoneDB, iTALK, and SingleCellSignalR) to analyze the interactions of fibroblasts with Mono/Macro in HCC. [Fig 5E](#) shows that there are overlapping regions among these five databases. When using different databases, the number of literature-supported L-R pairs detected by scSDNE was similar, indicating that its predictive results have high stability and consistency ([Fig 5F](#)).

## Discussion

In fact, GRNs encompass most known and unknown gene regulatory information, including L-R secretion and the activities of downstream regulatory factors [62]. To this end, we introduced scSDNE, a model that integrates intercellular L-R interactions with regulatory information among intracellular genes to infer and analyze cell-cell communication. scSDNE embeds genes into a shared latent space, calculating the distances between all gene pairs, which allows for the identification of L-R pairs that may have been overlooked by other methods but are nonetheless of significant investigative relevance. Importantly, given that most L-R pairs in LRdb overlap with other databases, LRdb is regarded as a reliable resource.

Case studies on scRNA-seq data from human lesional AD skin, HCC and gastric cancer demonstrate that scSDNE effectively infers intercellular communication, exhibiting superior accuracy compared to other tools. A case study focusing on lesional AD skin revealed notable signal crosstalk among immune cells. Many L-R pairs inferred from Inflam. FIB, such as CXCL12-CXCR4 and CCL19-CCR7, have been confirmed in multiple studies as primarily mediators of cell recruitment to inflammatory regions, significantly promoting inflammation exacerbation and development [24–26]. Additionally, we identified that the THBS2-CD36 interaction may influence the pathological process of AD through the inhibition of angiogenesis, meriting further investigation.

Analysis of the HCC datasets unveiled a complex interaction network between hepatocytes and immune cells, highlighting the key roles of FN1, MDK, and their receptors ITGA4, ITGA5, ITGA6. This analysis revealed significant L-R interactions between malignant cells and Mono/Macro, with the receptor SDC4 playing a crucial role. Furthermore, we observed notable metabolic differences between tumor tissues and adjacent tissues with significant interactions involving FGL1-LAG3 being negatively correlated with T cell subset numbers, indicating down-regulation of immune pathways and highlighting their immunosuppressive effects in liver cancer [37–39].



**Fig 5. Comparison of the performance of scSDNE with CellChat, CellPhoneDB, iTALK and scTenifoldXct on the scRNA-seq dataset of gastric cancer (from fibroblasts to macrophages).** (A) Circos plot depicts the strength of intercellular communication from fibroblasts to other cell types in TME. (B) UpSetR plot illustrates the results from the five tools utilizing their respective LR database. The horizontal bar graph in the lower left corner represents the total number of L-R pairs detected by each method. The intercellular communication results obtained by the different methods are represented by multiple black dots and connecting lines, and the number of intersections displayed in the bar graph above. (C) UpSetR plot shows the results from the five tools utilizing their respective LR database. The horizontal bar graph in the lower left corner represents the total number of L-R pairs detected by each method. The intercellular communication results obtained by the different methods are represented by multiple black dots and connecting lines, and the number of intersections displayed in the bar graph above.

from the five tools using a common LRdb. (D) ROC curves plot depicts the performance of the five methods in assessing intercellular communication. (E) Overlap analysis of the LR database. (F) Comparison of literature support rates for scSDNE using different databases.

<https://doi.org/10.1371/journal.pcbi.1013027.g005>

cell2location combines Visium and scRNA-seq data, allowing scSDNE to deeply analyze intercellular communication within the human lymph node micro-environment. The study reveals key interaction mechanisms between germinal center B cells, T cells, and FDCs. The results indicate that the MHC class II molecule-mediated interactions between FDCs and T cells are the most significant, promoting the activation of B cells, macrophages, and T cells. Furthermore, the CXCL13-CXCR5 signaling pathway plays a crucial role in regulating B cell migration and lymph follicle formation, further supporting the collaborative function of immune cells [49,50].

Through scSDNE analysis, we also elucidated L-R interactions between fibroblasts and immune cells—particularly macrophages—in gastric cancer, where interactions of CD44 with collagen family genes and FN1 emerged as key mechanisms for information exchange. The substantial enrichment of multiple signaling pathways further underscores the significance of these key L-R pairs in the pathological mechanisms of gastric cancer.

scSDNE represents a semi-supervised graph embedding approach designed to infer intercellular communication. By effectively leveraging gene expression regulation information derived from scRNA-seq data, it accurately identifies L-R pairs that are biologically significant between interacting cells. Although scSDNE performs excellently in inferring intercellular communication, its requirement for users to manually select sending and receiving cells somewhat limits its applicability. To overcome this limitation, future research will focus on developing automated algorithms to reduce reliance on user intervention, thereby enhancing the method's applicability and efficiency in complex biological environments. Additionally, integrating spatial transcriptomics data to obtain micro-environment information will enable scSDNE to predict intercellular interactions at the spatial tissue level. In summary, scSDNE, as an effective tool, has broad application prospects for analyzing intercellular communication at the single-cell level, and its future development will open new possibilities for biomedical research.

## Methods

### Data processing

The input of scSDNE is a gene expression matrix with annotated cell types. In scRNA-seq datasets, low-quality cells are filtered out based on metrics such as gene counts and the ratios of RNA and mitochondrial genes per cell. The data is then normalized using the `NormalizeData` function from Seurat [63]. To integrate data from different samples, batch effect correction is performed using the R package “Harmony” to address technical differences between samples. Subsequently, PCA and UMAP [64] are employed for data dimensional reduction and visualization. Cell type annotation is accomplished by identifying marker genes for each cell cluster. Finally, the analysis focuses on highly variable genes to highlight gene features that are most representative of distinct cell states.

### Crosstalk score calculation

Assume that gene  $i$  has an average expression level of  $u_{iA}$  in cell type A, and gene  $j$  has an average expression level of  $u_{jB}$  in cell type B. According to the scoring method described in SingleCellSignalR [65], the interaction score between gene  $i$  in cell type A and gene  $j$  in cell type B is denoted as follows:

$$score_{iA,jB} = \frac{\sqrt{u_{iA}u_{jB}}}{c + \sqrt{u_{iA}u_{jB}}} \quad (1)$$

where  $c = \text{mean}(C)$  and  $C$  is the count matrix for cells A and B. Hence, the interaction score matrix  $S \in \mathbb{R}^{n \times n}$  consists of  $score_{iA,jB}$  between cell types A and B, with all elements in  $S$  ranging from  $[0, 1)$ .

## Gene regulatory network construction

Given the gene expression matrices  $X_A$  and  $X_B$  for two cell types, A and B, DeepSEM [66] can infer the gene regulatory networks (GRNs) within each cell type. DeepSEM is a deep generative model that jointly infers GRNs and the biological representation of single-cell RNA sequencing data. Its framework is built on a beta-variational auto-encoder (beta-VAE), where the weights of the encoder and decoder functions directly represent the GRN adjacency matrix.  $W_A$  and  $W_B$  represent GRN adjacency matrices, with the absolute value of an element indicating the likelihood that gene  $i$  regulates gene  $j$ . Since the GRN constructed by DeepSEM is a directed graph, we require the likelihood of a regulatory relationship between two genes, necessitating that the constructed matrix  $W$  be symmetric and each element non-negative. To ensure symmetric, we define  $W_{ij} = W_{ji} = \max(W_{ij}, W_{ji})$ .

Given the significant differences between the edge weights of the GRN and the interaction scores, a scaling factor  $\mu$  is introduced to ensure that the regulatory information of genes and the crosstalk scores are comparable. This adjustment allows their contributions in the combined similarity matrix  $X$  to be roughly equal, preventing bias towards one measure during the embedding process. Referencing scTenifoldXct [67], the scaling factor  $\mu$  is set as follows:

$$\mu = \frac{\sum_{i,j} (W_A)_{ij} + \sum_{i,j} (W_B)_{ij}}{2 \sum_{i,j} \text{score}_{iA,jB}} \quad (2)$$

Since matrices  $W_A$  and  $W_B$  are relatively sparse, it is essential to identify the positions of the non-zero elements in both  $W_A$  and  $W_B$  to utilize them more effectively. Then, the corresponding elements in the crosstalk score matrix  $S$  are summed to obtain  $s$ . Consequently, let

$$\mu = \frac{\sum_{i,j} (W_A)_{ij} + \sum_{i,j} (W_B)_{ij}}{2s}, \quad W'_A = \frac{W_A}{\mu}, \quad W'_B = \frac{W_B}{\mu},$$

and  $X$  construct the joint similarity matrix as follows:

$$X = \begin{bmatrix} W'_A & S \\ S^T & W'_B \end{bmatrix}.$$

## Learning latent representations of each gene

Using the constructed weighted adjacency matrix  $X$  as input, the SDNE is employed to reconstruct and obtain the coordinates of genes from both the sender and receiver cells within a unified latent space. SDNE is a semi-supervised deep learning model that simultaneously optimizes both first-order and second-order similarities. First-order similarity refers to the local similarity between pairs of vertices, where the similarity is proportional to the edge weights and serves as supervised information to capture the local structure of the network. In contrast, second-order similarity describes the relationships between vertices and their neighborhoods, providing unsupervised information to capture the global structure of the network.

The loss function of scSDNE can be expressed as follows:

$$\mathcal{L}_{mix} = \mathcal{L}_{2nd} + \mathcal{L}_{1st} + \mathcal{L}_{reg} = \left\| \hat{X} - X \right\|_2^2 + 2tr(Y^T L Y) + \frac{1}{2} \sum_{k=1}^K \left( \left\| W^{(K)} \right\|_F^2 + \left\| \hat{W}^{(K)} \right\|_F^2 \right) \quad (3)$$

where  $\mathcal{L}_{reg}$  represents the regularization term to prevent over-fitting, and  $W^{(K)}$ ,  $\hat{W}^{(K)}$  denote the weight matrices of the  $K$ -th layer of the auto-encoder.  $L$  is the Laplacian matrix, where  $L = D - X$ ,  $D \in \mathbb{R}^{n \times n}$  are diagonal matrices with diagonal elements  $D_{i,i} = \sum_j X_{i,j}$ . To avoid mapping all instances to zero, an additional constraint  $Y^T D Y = I_n$  is added, with  $I_n \in \mathbb{R}^{n \times n}$



being an identity matrix. To satisfy these constraints, the optimization method proposed by Nguyen et al. [68] is employed. This method computes Riemannian gradients by projecting the gradient onto the tangent space of the Stiefel manifold. These gradients are then employed to update the parameters, ensuring that the output of the optimization problem remains constrained within Stiefel manifold.

The SDNE model consists of an encoder and a decoder, each featuring multiple stacked linear layers, with a Sigmoid activation function applied after each layer to facilitate the network's learning of non-linear data mappings. The encoder transforms the adjacency matrix, which contains interaction scores between cell types and gene regulatory relationships, into a hidden representation. Subsequently, the decoder reconstructs this representation to yield the output adjacency matrix. By minimizing the loss function, scSDNE effectively captures the local structural information of nodes, particularly emphasizing multi-order neighborhood relationships.

### Determining the statistical significance of interactions between cell types

The obtained  $Y = \begin{bmatrix} Y_A \\ Y_B \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$  by minimizing the loss function represents the embedding of all genes in a low-

dimensional latent space. Where  $Y_A$  and  $Y_B \in \mathbb{R}^{n \times n}$  are low-dimensional representations of genes in cell types A and B, respectively. Non-parametric tests are then used to determine the statistical significance of L-R pairs. Under the null hypothesis that there are no LR-mediated interactions between gene pairs, important gene pairs among all combinations are identified. The Euclidean distance across cell types for each paired gene in cell types A and B is calculated. After collecting the distances of gene pairs excluding L-R in the LRdb, the null hypothesis distribution  $dist_{null}$  can be obtained. Gene pairs that are not present in the database but appear in the combinations are considered less likely to be LR-mediated interacting gene pairs. Next, the percentiles of L-R pairs under the null hypothesis distribution are calculated as follows:

$$p - value = \frac{\text{percentile}(dist_{null}, dist_i)}{100} \quad (4)$$

and set it as the original p-value. The threshold for all datasets is set at 0.05.

### Supporting information

**S1 Fig. UMAP plot of scRNA-seq data.** UMAP plot shows ten cell types from 19865 cells across 5 patients in the GSE167297 datasets.

(TIF)

**S2 Fig. Clustering heatmap.** Heatmap depicts the expression levels of marker gene in indicated cell types, with cell types displayed at the top and corresponding marker genes listed at the bottom.

(TIF)

**S3 Fig. Dot plot of cell-cell interactions.** Dot plot shows the predicted interactions between fibroblasts and other cell types.

(TIF)

**S4 Fig. Wayne diagram of the 5 L-R pairs database.** Overlap analysis of the LR database.

(TIF)

### Author contributions

**Conceptualization:** Chenchen Jia.

**Formal analysis:** Haiyun Wang, Jianping Zhao.

**Investigation:** Haiyun Wang.

**Methodology:** Chenchen Jia.

**Project administration:** Jianping Zhao, Chunhou Zheng.

**Resources:** Jianping Zhao.

**Supervision:** Haiyun Wang, Jianping Zhao, Junfeng Xia, Chunhou Zheng.

**Visualization:** Chenchen Jia.

**Writing – original draft:** Chenchen Jia.

**Writing – review & editing:** Jianping Zhao, Junfeng Xia.

## References

1. Ma F, Zhang S, Song L, Wang B, Wei L, Zhang F. Applications and analytical tools of cell communication based on ligand-receptor interactions at single cell level. *Cell Biosci.* 2021 Jul 3;11(1):121. <https://doi.org/10.1186/s13578-021-00635-z> PMID: [34217372](#)
2. Peng L, Wang F, Wang Z, Tan J, Huang L, Tian X, et al. Cell-cell communication inference and analysis in the tumour microenvironments from single-cell transcriptomics: data resources and computational strategies. *Brief Bioinform.* 2022 Jul 18;23(4). <https://doi.org/10.1093/bib/bbac234> PMID: [35753695](#)
3. Zhang Y, Liu T, Hu X, Wang M, Wang J, Zou B, et al. CellCall: integrating paired ligand-receptor and transcription factor activities for cell-cell communication. *Nucleic Acids Res.* 2021;49(15):8520–8534. <https://doi.org/10.1093/nar/gkab638> PMID: [34331449](#)
4. Ramilowski JA, Goldberg T, Harshbarger J, Kloppmann E, Lizio M, Satagopam VP, et al. A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat Commun.* 2015;6:7866. <https://doi.org/10.1038/ncomms8866> PMID: [26198319](#)
5. Armingol E, Officer A, Harismendy O, Lewis NE. Deciphering cell-cell interactions and communication from gene expression. *Nat Rev Genet.* 2021 Feb;22(2):71–88. <https://doi.org/10.1038/s41576-020-00292-x> PMID: [33168968](#)
6. Almet AA, Cang Z, Jin S, Nie Q. The landscape of cell-cell communication through single-cell transcriptomics. *Curr Opin Syst Biol.* 2021 Jun;26:12–23. <https://doi.org/10.1016/j.coisb.2021.03.007> PMID: [33969247](#); PMCID: PMC8104132
7. Cheng C, Chen W, Jin H, Chen X. A review of single-cell RNA-seq annotation, integration, and cell-cell communication. *Cells.* 2023 Jul 30;12(15):1970. <https://doi.org/10.3390/cells12151970> PMID: [37566049](#); PMCID: PMC10417635
8. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun.* 2021 Feb 17;12(1):1088. <https://doi.org/10.1038/s41467-021-21246-9> PMID: [33597522](#); PMCID: PMC7889871
9. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc.* 2020 Apr;15(4):1484–1506. <https://doi.org/10.1038/s41596-020-0292-x> PMID: [32103204](#)
10. Hou R, Denisenko E, Ong HT, Ramilowski JA, Forrest ARR. Predicting cell-to-cell communication networks using NATMI. *Nat Commun.* 2020 Oct 6;11(1):5011. <https://doi.org/10.1038/s41467-020-18873-z> PMID: [33024107](#)
11. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods.* 2020;17(2):159–162. <https://doi.org/10.1038/s41592-019-0667-5> PMID: [31819264](#)
12. Hu Y, Peng T, Gao L, Tan K. CytoTalk: De novo construction of signal transduction networks using single-cell transcriptomic data. *Sci Adv.* 2021 Apr 14;7(16):eabf1356. <https://doi.org/10.1126/sciadv.abf1356> PMID: [33853780](#); PMCID: PMC8046375
13. Yuan Y, Cosme C Jr, Adams TS, Schupp J, Sakamoto K, Xylourgidis N, et al. CINS: Cell Interaction Network inference from Single cell expression data. *PLoS Comput Biol.* 2022 Sep 12;18(9):e1010468. <https://doi.org/10.1371/journal.pcbi.1010468> PMID: [36095011](#)
14. Dries R, Zhu Q, Dong R, Eng C-HL, Li H, Liu K, et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* 2021 Mar 8;22(1):78. <https://doi.org/10.1186/s13059-021-02286-2> PMID: [33685491](#)
15. Li H, Ma T, Hao M, Guo W, Gu J, Zhang X, et al. Decoding functional cell-cell communication events by multi-view graph learning on spatial transcriptomics. *Brief Bioinform.* 2023;24(6):bbad359. <https://doi.org/10.1093/bib/bbad359> PMID: [37824741](#)
16. Cang Z, Nie Q. Inferring spatial and signaling relationships between cells from single cell transcriptomic data. *Nat Commun.* 2020;11(1):2084. <https://doi.org/10.1038/s41467-020-15968-5> PMID: [32350282](#)
17. Pham D, Tan X, Balderson B, Xu J, Grice LF, Yoon S, et al. Robust mapping of spatiotemporal trajectories and cell-cell interactions in healthy and diseased tissues. *Nat Commun.* 2023 Nov 25;14(1):7739. <https://doi.org/10.1038/s41467-023-43120-6> PMID: [38007580](#); PMCID: PMC10676408
18. Cang Z, Zhao Y, Almet AA, Stabell A, Ramos R, Plikus MV, et al. Screening cell-cell communication in spatial transcriptomics via collective optimal transport. *Nat Methods.* 2023 Feb;20(2):218–228. <https://doi.org/10.1038/s41592-022-01728-4> PMID: [36690742](#)
19. Zhang L, Chen D, Song D, Liu X, Zhang Y, Xu X, et al. Clinical and translational values of spatial transcriptomics. *Signal Transduct Target Ther.* 2022 Apr 1;7(1):111. <https://doi.org/10.1038/s41392-022-00960-w> PMID: [35365599](#)

20. Longo SK, Guo MG, Ji AL, Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet*. 2021 Oct;22(10):627–644. <https://doi.org/10.1038/s41576-021-00370-8> PMID: [34145435](#); PMCID: PMC9888017
21. Wang D, Cui P, Zhu W. Structural deep network embedding. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2016; San Francisco, CA, USA. New York, NY: Association for Computing Machinery; 2016. p. 1225–34. <https://doi.org/10.1145/2939672.2939753>
22. He H, Suryawanshi H, Morozov P, Gay-Mimbrera J, Del Duca E, Kim HJ, et al. Single-cell transcriptome analysis of human skin identifies novel fibroblast subpopulation and enrichment of immune subsets in atopic dermatitis. *J Allergy Clin Immunol*. 2020;145(6):1615–28. <https://doi.org/10.1016/j.jaci.2020.01.042> PMID: [32035984](#)
23. Davidson S, Coles M, Thomas T, Kollias G, Ludewig B, Turley S, et al. Fibroblasts as immune regulators in infection, inflammation and cancer. *Nat Rev Immunol*. 2021 Nov;21(11):704–717. <https://doi.org/10.1038/s41577-021-00540-z> PMID: [33911232](#).
24. Yan Y, Chen R, Wang X, Hu K, Huang L, Lu M, et al. CCL19 and CCR7 Expression, Signaling Pathways, and Adjuvant Functions in Viral Infection and Prevention. *Front Cell Dev Biol*. 2019;7:212. <https://doi.org/10.3389/fcell.2019.00212> PMID: [31632965](#)
25. Sun Z, Kim JH, Kim SH, Kim HR, Zhang K, Pan Y, et al. Skin-resident natural killer T cells participate in cutaneous allergic inflammation in atopic dermatitis. *J Allergy Clin Immunol*. 2021 May;147(5):1764–77. <https://doi.org/10.1016/j.jaci.2020.11.049> PMID: [33516870](#)
26. Ward AC. Immune Factors, Immune Cells and Inflammatory Diseases. *Int J Mol Sci*. 2024 Feb 19;25(4):2417. <https://doi.org/10.3390/ijms25042417> PMID: [38397094](#); PMCID: PMC10889257
27. Jiao H, Zeng L, Zhang J, Yang S, Lou W. THBS2, a microRNA-744-5p target, modulates MMP9 expression through CUX1 in pancreatic neuroendocrine tumors. *Oncol Lett*. 2020 Mar;19(3):1683–1692. <https://doi.org/10.3892/ol.2020.11273> PMID: [32194660](#); PMCID: PMC7039111
28. Liao X, Yan S, Li J, Jiang C, Huang S, Liu S, et al. CD36 and Its Role in Regulating the Tumor Microenvironment. *Curr Oncol*. 2022 Oct 27;29(11):8133–8145. <https://doi.org/10.3390/curroncol29110642> PMID: [36354702](#)
29. Zhang Y, Jing D, Cheng J, Chen X, Shen M, Liu H. The efficacy and safety of IL-13 inhibitors in atopic dermatitis: A systematic review and meta-analysis. *Front Immunol*. 2022 Jul 27;13:923362. <https://doi.org/10.3389/fimmu.2022.923362> PMID: [35967348](#); PMCID: PMC9364267
30. Liu Y, Zhang L, Ju X, Wang S, Qie J. Single-Cell Transcriptomic Analysis Reveals Macrophage-Tumor Crosstalk in Hepatocellular Carcinoma. *Front Immunol*. 2022 Jul 25;13:955390. <https://doi.org/10.3389/fimmu.2022.955390> PMID: [35958556](#)
31. Zhou P-Y, Zhou C, Gan W, Tang Z, Sun B-Y, Huang J-L, et al. Single-cell and spatial architecture of primary liver cancer. *Commun Biol*. 2023 Nov 20;6(1):1181. <https://doi.org/10.1038/s42003-023-05455-0> PMID: [37985711](#)
32. Li C, Chen J, Li Y, Wu B, Ye Z, Tian X, et al. 6-Phosphogluconolactonase Promotes Hepatocellular Carcinogenesis by Activating Pentose Phosphate Pathway. *Front Cell Dev Biol*. 2021 Oct 26;9:753196. <https://doi.org/10.3389/fcell.2021.753196> PMID: [34765603](#); PMCID: PMC8576403
33. Xiao B, Li G, Gulizeba H, Liu H, Sima X, Zhou T, et al. Choline metabolism reprogramming mediates an immunosuppressive microenvironment in non-small cell lung cancer (NSCLC) by promoting tumor-associated macrophage functional polarization and endothelial cell proliferation. *J Transl Med*. 2024 May 10;22(1):442. <https://doi.org/10.1186/s12967-024-05242-3> PMID: [38730286](#)
34. Xu X, Shen L, Li W, Liu X, Yang P, Cai J. ITGA5 promotes tumor angiogenesis in cervical cancer. *Cancer Med*. 2023 May;12(10):11983–11999. <https://doi.org/10.1002/cam4.5873> PMID: [36999964](#); PMCID: PMC10242342
35. Qin Z, Zhou C. HOXA13 promotes gastric cancer progression partially via the FN1-mediated FAK/Src axis. *Exp Hematol Oncol*. 2022;11(1):7. <https://doi.org/10.1186/s40164-022-00260-7> PMID: [35197128](#)
36. Filippou PS, Karagiannis GS, Constantinidou A. Midkine (MDK) growth factor: a key player in cancer progression and a promising therapeutic target. *Oncogene*. 2020 Mar;39(10):2040–54. <https://doi.org/10.1038/s41388-019-1124-8> PMID: [31801970](#)
37. Guo M, Yuan F, Qi F, Sun J, Rao Q, Zhao Z, et al. Expression and clinical significance of LAG-3, FGL1, PD-L1 and CD8+T cells in hepatocellular carcinoma using multiplex quantitative analysis. *J Transl Med*. 2020 Aug 6;18(1):306. <https://doi.org/10.1186/s12967-020-02469-8> PMID: [32762721](#)
38. Shi A-P, Tang X-Y, Xiong Y-L, Zheng K-F, Liu Y-J, Shi X-G, et al. Immune Checkpoint LAG3 and Its Ligand FGL1 in Cancer. *Front Immunol*. 2022 Jan 17;12:785091. <https://doi.org/10.3389/fimmu.2021.785091> PMID: [35111155](#)
39. Sun L-L, Zhao L-N, Sun J, Yuan H-F, Wang Y-F, Hou C-Y, et al. Inhibition of USP7 enhances CD8+ T cell activity in liver cancer by suppressing PRDM1-mediated FGL1 upregulation. *Acta Pharmacol Sin*. 2024;45(8):1686–1700. <https://doi.org/10.1038/s41401-024-01263-2> PMID: [38589688](#)
40. Yang H, Liu Y, Zhao MM, Guo Q, Zheng XK, Liu D, et al. Therapeutic potential of targeting membrane-spanning proteoglycan SDC4 in hepatocellular carcinoma. *Cell Death Dis*. 2021 May 14;12(5):492. <https://doi.org/10.1038/s41419-021-03780-y> PMID: [33990545](#); PMCID: PMC8121893
41. Zhu Y, Zheng D, Lei L, Cai K, Xie H, Zheng J, et al. High expression of syndecan-4 is related to clinicopathological features and poor prognosis of pancreatic adenocarcinoma. *BMC Cancer*. 2022 Oct 5;22(1):1042. <https://doi.org/10.1186/s12885-022-10128-y> PMID: [36199068](#); PMCID: PMC9533499
42. Keller-Pinter A, Gyulai-Nagy S, Becsky D, Dux L, Roivo L. Syndecan-4 in Tumor Cell Motility. *Cancers (Basel)*. 2021 Jul 1;13(13):3322. <https://doi.org/10.3390/cancers13133322> PMID: [34282767](#); PMCID: PMC8268284
43. Zhang D, Bi J, Liang Q, Wang S, Zhang L, Han F, et al. VCAM1 Promotes Tumor Cell Invasion and Metastasis by Inducing EMT and Transendothelial Migration in Colorectal Cancer. *Front Oncol*. 2020 Jul 23;10:1066. <https://doi.org/10.3389/fonc.2020.01066> PMID: [32793471](#)
44. Shen J, Zhai J, You Q, Zhang G, He M, Yao X, et al. Cancer-associated fibroblasts-derived VCAM1 induced by H. pylori infection facilitates tumor invasion in gastric cancer. *Oncogene*. 2020;39(14):2961–74. <https://doi.org/10.1038/s41388-020-1197-4> PMID: [32034307](#)

45. Kleshchevnikov V, Shmatko A, Dann E, Aivazidis A, King HW, Li T, et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat Biotechnol.* 2022 May;40(5):661–671. <https://doi.org/10.1038/s41587-021-01139-4> PMID: 35027729.
46. Dileepan T, Malhotra D, Kotov DI, Kolawole EM, Krueger PD, Evavold BD, Jenkins MK. MHC class II tetramers engineered for enhanced binding to CD4 improve detection of antigen-specific T cells. *Nat Biotechnol.* 2021;39(8):943–948. <https://doi.org/10.1038/s41587-021-00893-9> PMID: 33941928
47. Wang RF. The role of MHC class II-restricted tumor antigens and CD4+ T cells in antitumor immunity. *Trends Immunol.* 2001;22(5):269–76. [https://doi.org/10.1016/s1471-4906\(01\)01896-8](https://doi.org/10.1016/s1471-4906(01)01896-8) PMID: 11323286
48. Roche PA, Furuta K. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat Rev Immunol.* 2015 Apr;15(4):203–16. <https://doi.org/10.1038/nri3818> PMID: 25720354; PMCID: PMC6314495
49. Wang B, Wang M, Ao D, Wei X. CXCL13-CXCR5 axis: Regulation in inflammatory diseases and cancer. *Biochim Biophys Acta Rev Cancer.* 2022 Sep;1877(5):188799. <https://doi.org/10.1016/j.bbcan.2022.188799> PMID: 36103908
50. Jenh CH, Cox MA, Hipkin W, Lu T, Pugliese-Sivo C, Gonsiorek W, Chou CC, Narula SK, Zavodny PJ. Human B cell-attracting chemokine 1 (BCA-1; CXCL13) is an agonist for the human CXCR3 receptor. *Cytokine.* 2001 Aug 7;15(3):113–21. <https://doi.org/10.1006/cyto.2001.0923> PMID: 11554781.
51. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. *Cancer Immunol Res.* 2019;7(5):737–750. <https://doi.org/10.1158/2326-6066.CIR-18-0436> PMID: 30842092
52. Sun K, Xu R, Ma F, Yang N, Li Y, Sun X, et al. scRNA-seq of gastric tumor shows complex intercellular interaction with an alternative T cell exhaustion trajectory. *Nat Commun.* 2022 Aug 23;13(1):4943. <https://doi.org/10.1038/s41467-022-32627-z> PMID: 35999201
53. Al-Bzour NN, Al-Bzour AN, Ababneh OE, Al-Jezawi MM, Saeed A, Saeed A. Cancer-Associated Fibroblasts in Gastrointestinal Cancers: Unveiling Their Dynamic Roles in the Tumor Microenvironment. *Int J Mol Sci.* 2023 Nov 19;24(22):16505. <https://doi.org/10.3390/ijms242216505> PMID: 38003695; PMCID: PMC10671196
54. Locati M, Curtale G, Mantovani A. Diversity, Mechanisms, and Significance of Macrophage Plasticity. *Annu Rev Pathol.* 2020 Jan 24;15:123–147. <https://doi.org/10.1146/annurev-pathmechdis-012418-012718> PMID: 31530089
55. Li X, Sun Z, Peng G, Xiao Y, Guo J, Wu B, et al. Single-cell RNA sequencing reveals a pro-invasive cancer-associated fibroblast subgroup associated with poor clinical outcomes in patients with gastric cancer. *Theranostics.* 2022 Jan 1;12(2):620–638. <https://doi.org/10.7150/thno.60540> PMID: 34976204; PMCID: PMC8692898
56. Necula L, Matei L, Dragu D, Pitica I, Neagu A, Bleotu C, et al. Collagen Family as Promising Biomarkers and Therapeutic Targets in Cancer. *Int J Mol Sci.* 2022 Oct 17;23(20):12415. <https://doi.org/10.3390/ijms232012415> PMID: 36293285; PMCID: PMC9604126
57. San Antonio JD, Jacenko O, Fertala A, Orgel JPRO. Collagen Structure-Function Mapping Informs Applications for Regenerative Medicine. *Bioengineering (Basel).* 2020 Dec 29;8(1):3. <https://doi.org/10.3390/bioengineering8010003> PMID: 33383610
58. Sun X, Li K, Hase M, Zha R, Feng Y, Li B-Y, et al. Suppression of breast cancer-associated bone loss with osteoblast proteomes via Hsp90ab1/ moesin-mediated inhibition of TGFβ/FN1/CD44 signaling. *Theranostics.* 2022 Jan 1;12(2):929–43. <https://doi.org/10.7150/thno.66148> PMID: 34976221. Erratum in: *Theranostics.* 2023 Jan 1;13(1):16–19. <https://doi.org/10.7150/thno.79085> PMID: 34976221
59. Gao F, Zhang G, Liu Y, He Y, Sheng Y, Sun X, et al. Activation of CD44 signaling in leader cells induced by tumor-associated macrophages drives collective detachment in luminal breast carcinomas. *Cell Death Dis.* 2022;13(6):540. <https://doi.org/10.1038/s41419-022-04986-4> PMID: 35680853
60. Belkin AM, Stepp MA. Integrins as receptors for laminins. *Microsc Res Tech.* 2000 Nov 1;51(3):280–301. [https://doi.org/10.1002/1097-0029\(20001101\)51:3<280::AID-JEMT7>3.0.CO;2-O](https://doi.org/10.1002/1097-0029(20001101)51:3<280::AID-JEMT7>3.0.CO;2-O) PMID: 11054877.
61. Wang Y, Wang R, Zhang S, Song S, Jiang C, Han G, et al. iTALK: an R Package to Characterize and Illustrate Intercellular Communication. *bioRxiv.* 2019:507871. <https://doi.org/10.1101/507871>
62. Liu Y, Zhang Y, Chang X, Liu X. MDIC3: Matrix decomposition to infer cell-cell communication. *Patterns (N Y).* 2024 Jan 11;5(2):100911. <https://doi.org/10.1016/j.patter.2023.100911> PMID: 38370122
63. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive integration of single-cell data. *Cell.* 2019 Jun 13;177(7):1888–902.e21. <https://doi.org/10.1016/j.cell.2019.05.031> PMID: 31178118
64. Dorrity MW, Saunders LM, Queitsch C, Fields S, Trapnell C. Dimensionality reduction by UMAP to visualize physical and genetic interactions. *Nat Commun.* 2020 Mar 24;11(1):1537. <https://doi.org/10.1038/s41467-020-15351-4> PMID: 32210240; PMCID: PMC7093466
65. Cabello-Aguilar S, Alame M, Kon-Sun-Tack F, Fau C, Lacroix M, Colinge J. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res.* 2020 Jun 4;48(10):e55. <https://doi.org/10.1093/nar/gkaa183> PMID: 32196115
66. Shu H, Zhou J, Lian Q, Li H, Zhao D, Zeng J, et al. Modeling gene regulatory networks using neural network architectures. *Nat Comput Sci.* 2021 Jul;1(7):491–501. <https://doi.org/10.1038/s43588-021-00099-8> PMID: 38217125
67. Yang Y, Li G, Zhong Y, Xu Q, Lin Y-T, Roman-Vicharra C, et al. scTenifoldXct: A semi-supervised method for predicting cell-cell interactions and mapping cellular communication graphs. *Cell Syst.* 2023;14(4):302–11.e4. <https://doi.org/10.1016/j.cels.2023.01.004> PMID: 36787742
68. Nguyen ND, Huang J, Wang D. A deep manifold-regularized learning model for improving phenotype prediction from multi-modal data. *Nat Comput Sci.* 2022 Jan;2(1):38–46. <https://doi.org/10.1038/s43588-021-00185-x> PMID: 35480297; PMCID: PMC9038085