

RESEARCH

Open Access



Characterizing tertiary lymphoid structures associated single-cell atlas in breast cancer patients

Xiaokai Fan^{1,2†}, Daqin Feng^{1†}, Donggui Wei^{1†}, Anqi Li¹, Fangyi Wei¹, Shufang Deng¹, Muling Shen¹, Congzhi Qin¹, Yongjia Yu^{1*} and Lun Liang^{1*}

Abstract

The tertiary lymphoid structure (TLS) is recognized as a potential prognosis factor for breast cancer and is strongly associated with response to immunotherapy. Inducing TLS neogenesis can enhance the immunogenicity of tumors and improve the efficacy of immunotherapy. However, our understanding of TLS associated region at the single-cell level remains limited. Therefore, we employed high-resolution techniques, including single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics (ST), and a TLS-specific signature to investigate TLS associated regions in breast cancer. We identified eighteen cell types within the TLS associated regions and calculated differential expression genes by comparing TLS associated regions with other areas. Notably, macrophages in the TLS associated regions exhibit lineage transformation, shifting from facilitators of immune activation to supporters of tumor cell growth. In terms of cell–cell communication within the TLS associated regions, KRT86⁺ CD8⁺ T cells, HISTIH4C⁺ cycling CD8⁺ T cells, IFNG⁺ CD8⁺ T cells, and IGKV3-20⁺ B cells demonstrate strong interactions with other cells. Additionally, we found that APOD⁺ fibroblast and CCL21⁺ fibroblast primarily recruit T and B cells through the CXCL12–CXCR4 ligand–receptor signaling pathway. We also validate these findings in four independent breast cancer datasets, which include one cell-level resolution dataset from the 10×Xenium platform and three spot-level datasets from the 10×Visium platform.

Keywords Tertiary lymphoid structure, Cell components, Lineage trajectory, Cell–cell communication

Introduction

Breast cancer can be classified into three clinically relevant subtypes: luminal (characterized by the expression of estrogen receptor [ER] and/or progesterone receptor [PR]), human epidermal growth factor receptor-2 positive (HER2+), and triple-negative breast cancer (TNBC), which lacks the expression of ER, PR, and HER2 [1]. Benefit for the development of diverse treatment methods, most early breast cancer patients experience prolonged survival following treatment [2]. However, for patients with highly malignant breast cancer, more effective treatment options are still needed [3, 4].

Recently, immunotherapy has been shown to extend survival in various types of cancers [5, 6]. According to

[†]Xiaokai Fan, Daqin Feng and Donggui Wei have contributed equally to this work.

*Correspondence:

Yongjia Yu

yyjfish@126.com

Lun Liang

lianglun@sr.gxmu.edu.cn

¹ Department of Neurosurgery, The First Affiliated Hospital, Guangxi Medical University, Nanning, China

² Center for Intelligent Medicine Research, Greater Bay Area Institute of Precision Medicine (Guangzhou), School of Life Sciences, Fudan University, Shanghai, China



previous studies, patients with higher levels of lymphocyte infiltration tend to benefit more from adjuvant or neoadjuvant chemotherapy and immunotherapy, resulting in longer progression-free survival (PFS) and overall survival (OS) [7]. However, breast cancer is not generally considered a highly immunogenic disease, among the subtypes, only HER2⁺ and TNBC patients exhibit higher levels of lymphocyte infiltration and are more likely to benefit from immunotherapy [8]. Therefore, the immune therapy outcomes vary significantly.

Tertiary lymphoid structures (TLSs) have emerged as novel predictors and facilitators of anti-tumor immune activity [9]. Notably, TLSs enhance immune activity within the tumor microenvironment and are strongly correlated with clinical benefits in breast cancer patients [10]. Additionally, TLSs promote lymphocyte infiltration and facilitate T-cell activation by enhancing tumor

antigen presentation [11]. Over the past decades, TLSs have been primarily studied using immunohistochemistry (IHC)-based technologies, which provide only limited information. Thus, there is a pressing need for using high-resolution approaches to better understand TLS regions. By integrating scRNA-seq and spatial transcriptomics (ST), it is possible to obtain high-resolution data on the cells and tissues within TLSs.

In this study, we leveraged ST and scRNA-seq datasets from Wu et al. to investigate the cellular components, macrophage transformations, and cell-cell communications within TLS associated region [12]. And to validate findings in other independent datasets, which include one cell-level resolution dataset from the 10×Xenium platform and three spot-level datasets from the 10×Visium platform, Fig. 1a indicates the workflow of this study. This study provides foundational knowledge of TLS associated

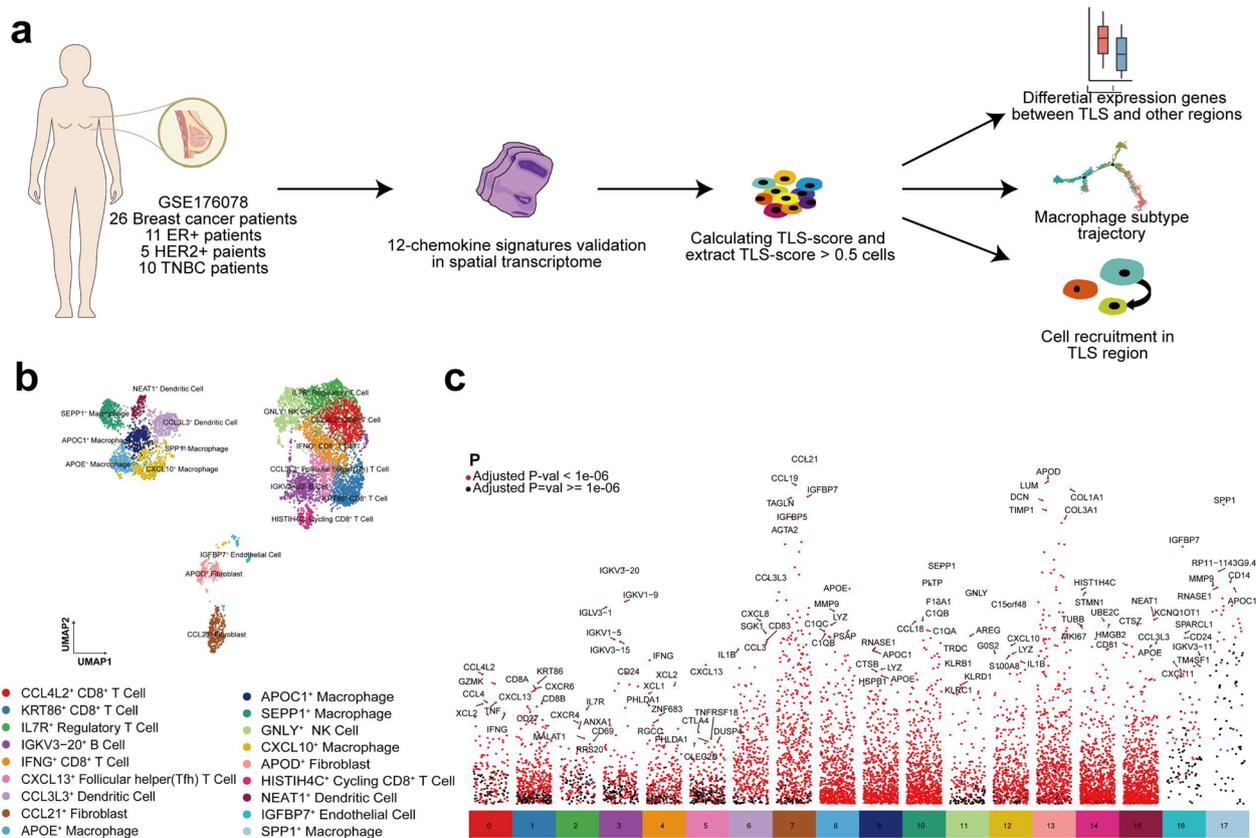


Fig. 1 The workflow of the study and major cell types in the TLS associated regions. **a** First, we downloaded the processed data of 26 scRNA-seq datasets and four spatial transcriptome datasets of breast cancer patients. Next, we validated 12-chemokines signature reported in a previous study, which has the ability to predict TLS regions in breast cancer. We then used this signature to calculate the TLS score of cells in these 26 scRNA-seq datasets, extracting high TLS-score cells for further analysis. Then, we annotated these extracted cells, calculated the differential expression genes between high TLS-score cells and low TLS-score cells, analyzed the macrophage lineage trajectory in the TLS associated regions, and cell-cell communications in the TLS associated region. **b** The UMAP plot presents the major cell types in the TLS associated regions, which includes four types of CD8⁺ T cells, five types of macrophages, two types of dendritic cells, fibroblasts, and five other cell types. **c** The highest expression genes in each cell type, with gene labels displaying the top six average log₂ fold change (log₂FC) genes for each cell type

region that may contribute to the development of novel immune therapeutic strategies or serve as an adjunct to enhance the efficacy of existing immunotherapies.

Materials and methods

Data acquisition

The scRNA-seq datasets of 26 breast cancer samples (GSE176078) were obtained from the GEO database (<https://www.ncbi.nlm.nih.gov/geo>). This dataset includes samples from 11 ER⁺, 5 HER2⁺, and 10 TNBC patients [13]. Corresponding spatial transcriptomics data for four patients were retrieved from the Zenodo data repository (<https://doi.org/https://doi.org/10.5281/zenodo.4739739>). Additionally, Xenium breast cancer patient data were acquired from the following link: <https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast>. The data from three breast cancer samples using the Visium platform are available at these links or through GEO accession: <https://www.10xgenomics.com/datasets/human-breast-cancer-block-a-section-1-1-standard-1-0-0>, <https://www.10xgenomics.com/datasets/human-breast-cancer-block-a-section-2-1-standard-1-0-0>, GSM6177599. The RNA sequencing datasets of TCGA breast cancer patients were extracted from the University of California Santa Cruz (UCSC) Xena (<https://xena.ucsc.edu/>) database, along with clinical information. This dataset comprised 1,211 tumor samples and 113 normal samples. After filtering out unpaired and duplicate data, we retained 1,076 tumor samples.

Quality control of single-cell RNA sequencing and spatial transcriptome data

Quality control was performed on the 26 single-cell RNA sequencing datasets using Seurat (version 4.0.1). Cells with fewer than 500 unique molecular identifiers (UMIs) and those with mitochondrial gene percentages exceeding 20% were filtered out. To eliminate batch effects among patients and normalize the data, we employed the CCA and the IntegrateData function in Seurat. For the Xenium spatial transcriptome data, genes with total counts below 10 and cells with fewer than 10 counts were filtered using stLearn (version 0.4.9).

Calculate the TLS-score of each cell and each spot

We use a 12-chemokine signature through the AddModuleScore function in Seurat to calculate the TLS score for the 26 scRNA-seq datasets and four spatial transcriptome datasets. A cut-off of TLS score > 0.5 was established for further analysis of high TLS score regions. Ultimately, we identified 8,506 cells that likely represent the majority of

TLS associated cells, with three patients' breast cancer tissues showing possible high TLS infiltration.

Cell clustering and annotation of cell types

Using PCA, JackStraw, and ScoreJackStraw functions in Seurat, we identified the principal components of genes representing each cell type in the TLS. We then employed the KNN algorithm along with UMAP (Uniform Manifold Approximation and Projection) for dimensionality reduction to identify major cell groups. Specific expression genes for each cell group were calculated using the FindAllMarkers function in Seurat. Major cell types were annotated based on marker genes from the CellMarker database and our own gene collections. For the Xenium spatial transcriptome, we preprocessed the h5 file data in stLearn and fine-tuned Geneformer on our annotated 18 cell types in TLS-associated regions to facilitate accurate cell type distinction. We then used the fine-tuned Geneformer to calculate similarity between cells in the Xenium dataset and the 18 TLS-associated cell types for inference of cell identities. For cell type deconvolution of Visium platform data, we applied CARD to deconvolute the 18 cell types in spatial transcriptome data.

Differential gene expression analysis

Differential expression of genes between TLS regions and other regions was assessed using the Wilcoxon test, with a significance level defined as FDR (adjusted p-value) < 0.05. This analysis identified 61 genes that were lowly expressed and 78 genes that were highly expressed in the TLS regions.

Gene ontology enrichment analysis

We utilized ClusterProfiler [14] (version 4.14.4) to conduct gene ontology analysis of the differentially expressed genes, with significance set at p-value < 0.05.

Specific marker genes of each cell type in the TLS associated cells

Using the FindAllMarkers function in Seurat, we calculated the representative genes for each cell type. We selected the top 30 genes based on average log₂ fold change (log₂FC) for each cell type, ensuring that the selected gene expression proportion in each cell type was less than 60% while not exceeding 60% in other cell types. The remaining genes were designated as specific marker genes for each cell type.

Survival analysis of each cell type in TLS associated regions with TCGA breast cancer cohort

The ssGSEA [15] method was employed to estimate the proportion of immune cells in TCGA BRCA patients using the specific marker genes identified for each cell type. The Kaplan–Meier method was used for survival analysis, with p-values calculated via the log-rank test.

Lineage trajectory analysis of macrophages in TLS associated regions

Monocle2 [16] (version 2.12.0) was used to analyze the pseudo-lineage trajectory of macrophages in the TLS associated regions. We first extracted the macrophage cell groups and created a new `CellDataSet` object in Monocle2 with the negative binomial size as the expression family. The macrophage trajectory direction was determined based on survival analysis results. Differential genes along the lineage trajectory were selected based on average \log_2FC , p-value, and expression proportion in each cell type, with criteria of $\log_2FC > 1$, p-value < 0.0001 , and expression in at least 70% of cells of each type.

Cell–cell communication analysis

CellChat [17] (version 1.5.0) is an R-based computational tool that allows us to examine cell-to-cell communication. To analyze the communications among high TLS score cells, we first calculate the communication weight among each cell type, then assessed the communication strength across all signaling pathways in the CellChat database. For the spatial transcriptome, we employed stLearn to analyze specific ligand-receptor communication strength.

Result

Characterization of cellular components in tertiary lymphoid structure associated regions and their prognostic impact in breast cancer

TLSs are organized cellular aggregates located directly at tumor sites, resembling secondary lymphoid organs [18]. TLSs typically consist of a T cell-rich zone containing mature dendritic cells (DCs), juxtaposed with a B cell follicle exhibiting germinal center characteristics, and are surrounded by plasma cells. This structure has the potential to promote lymphocyte infiltration into tumor tissue [19]. Recent studies have shown that TLSs are associated with improved prognosis in TNBC and HER2+ breast cancer patients [20, 21].

To investigate the cellular components and cell–cell communications within TLS, we utilized Prabhakaran et al.'s 12-chemokine gene signature [22] to calculate TLS scores in ST datasets from four breast cancer patients (Supplementary Fig. 1a). By integrating the TLS scores

with the corresponding patients' H&E-stained images, we confirmed that the 12-chemokine gene signature effectively predicts TLS associated regions. This finding highlights the potential of the gene signature as a valuable tool for evaluating the cells whether belong to TLS associated regions. Therefore, we applied this gene signature to calculate TLS scores for pre-processed single cells from Wu et al.'s dataset of 26 breast cancer patients. When comparing TLS scores across the three breast cancer subtypes, we observed that TNBC and HER2+ patients had higher TLS scores than ER+ patients, which is consistent with previous studies (Supplementary Fig. 1b). Next, we extracted cells with TLS scores greater than 0.5 from the scRNA-seq datasets of the 26 breast cancer patients. We re-clustered these cells and annotated them into 18 groups based on specific cell markers (Fig. 1b, c). These groups primarily consisted of two types of fibroblasts (APOD⁺ fibroblast and CCL21⁺ fibroblast), five types of macrophage (APOC1⁺ macrophage, CXCL10⁺ macrophage, APOE⁺ macrophage, SEPP1⁺ macrophage, SPP1⁺ macrophage), six types of T cells (CCL4L2⁺ CD8⁺ T cells, KRT86⁺ CD8⁺ T cells, IL7R⁺ regulatory T cells, IFNG⁺ CD8⁺ T cells, HISTIH4C⁺ cycling CD8⁺ T cells, CXCL13⁺ Follicular helper (Tfh) T cell), and two types of dendritic cells (NEAT1⁺ dendritic cells and CCL3L3⁺ dendritic cells). The remaining cell types included IGFBP7⁺ endothelial cells, GNLY⁺ NK cells, and IGKV3-20⁺ B cells (Fig. 1b, c). Finally, we applied the ssGSEA method with specific highly expressed marker genes of these cells to infer their infiltration in TCGA-GDC breast cancer patients and analyzed the impact of these immune cells on patient survival. We found high infiltration of CXCL10⁺ macrophage, CCL21⁺ fibroblasts, CXCL13⁺ Follicular helper (Tfh) T cells, GNLY⁺ NK cells, IFNG⁺ CD8⁺ T cells, IGKV3-20⁺ B cells, IL7R⁺ regulatory T cells, KRT86⁺ CD8⁺ T cells, and NEAT1⁺ dendritic cells was significantly associated with better prognosis in breast cancer patients (Supplementary Fig. 2a-r).

Differentially expressed genes of cells in TLS associated regions compared to other regions

To gain insights into differentially expressed genes of cells in TLS associated regions compared to other regions, we calculated the differential expression of genes between TLS high-score and TLS low-score cells. Interestingly, CCL4L2, CCL3, CXCL10, CCL19, CCL5, and CCL3L3 are the top six chemokines highly expressed in the TLS associated regions, while CD24, MUCL1, and MGP are lowly expressed. Notably, CCL4L2, CCL3, and CXCL10 are key chemokines that have been linked to inflammatory responses and T cell recruitment [23, 24] (Fig. 2a). Furthermore, gene ontology (GO) enrichment analysis revealed that the highly expressed genes in

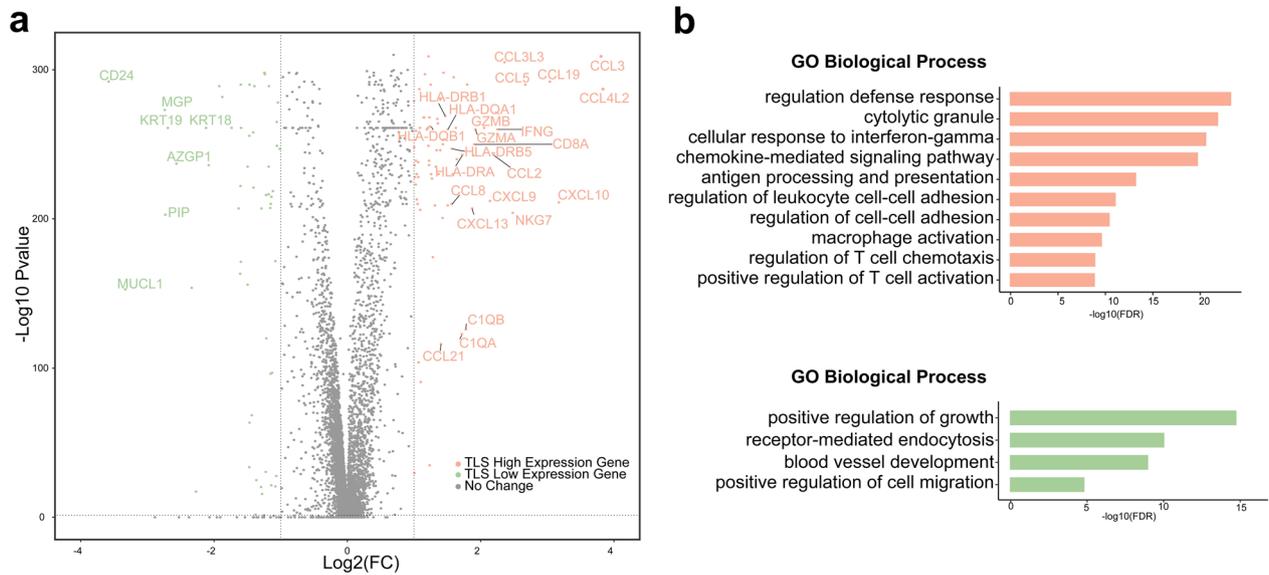


Fig. 2 The differential expression genes between TLS associated and other cells. **a** The differential expression genes between TLS associated and other cells, the green text indicates low-expression genes in TLS associated cells, while the red text indicates high expression genes in TLS associated cells. **b** Gene ontology enrichment analysis of the differentially expressed genes is presented, with the upper plot showing the enrichment of GO biological processes for high-expression genes in TLS associated cells and the lower plot displaying the enrichment for low-expression genes

TLS associated cells are predominantly involved in biological processes such as regulation of defense response, cytolytic granule activity, cellular response to interferon-gamma, chemokine-mediated signaling pathways, antigen processing and presentation, regulation of T cell chemotaxis, and positive regulation of T cell activation (Fig. 2b). In contrast, GO analysis of the low-expression genes in the TLS associated regions revealed enrichment in processes such as positive regulation of growth, positive regulation of cell migration, and blood vessel development. These findings clearly indicate that the highly expressed genes in the TLS associated regions are strongly associated with immune activity. We propose that these differentially expressed genes are key components in the formation and functional regulation of TLS associated regions. A deeper understanding of these genes could accelerate the development of strategies to induce TLS associated regions neogenesis in immune-compromised tumors, ultimately improving immunotherapy outcomes.

Macrophage transformation and functional dynamics within TLS associated regions

Survival analysis of immune cells in the TLS associated regions indicate that CXCL10⁺ macrophages are associated with a better prognosis in breast cancer compared with other types of macrophages (Supplementary Fig. 2a, c, g, q, r). Therefore, understanding the transformation of CXCL10⁺ macrophages into other macrophage subtypes is critical for uncovering the dynamic changes and the diverse functions of macrophages in the TLS associated regions. To investigate this process, we used the pseudo-time analysis method to study macrophage transformation. In the TLS associated regions, the macrophage progress through five distinct states, transitioning primarily from CXCL10⁺ macrophages and APOE⁺ macrophages to APOC1⁺ macrophages and SEPP1⁺ macrophages. Notably, SPP1⁺ macrophages are present in only a small subset of patients and appear transiently during the middle stages of the transformation process (Fig. 3a, b). The functional roles of macrophages change

(See figure on next page.)

Fig. 3 The lineage trajectory of macrophages in the TLS associated regions. **a** The transformation of macrophage lineages in the TLS can be divided into two processes, transitioning from CXCL10⁺ macrophages and APOE⁺ macrophages to APOC1⁺ macrophages and SEPP1⁺ macrophages, with SPP1⁺ macrophages predominantly present in the intermediate stages. **b** The pseudo-time plot illustrates the temporal dynamics of macrophage differentiation. **c** High-expression genes at each stage of macrophage lineage transformation are shown, along with their enrichment in biological processes. **d** The expression levels of marker genes for these macrophage types change throughout the lineage transformation process

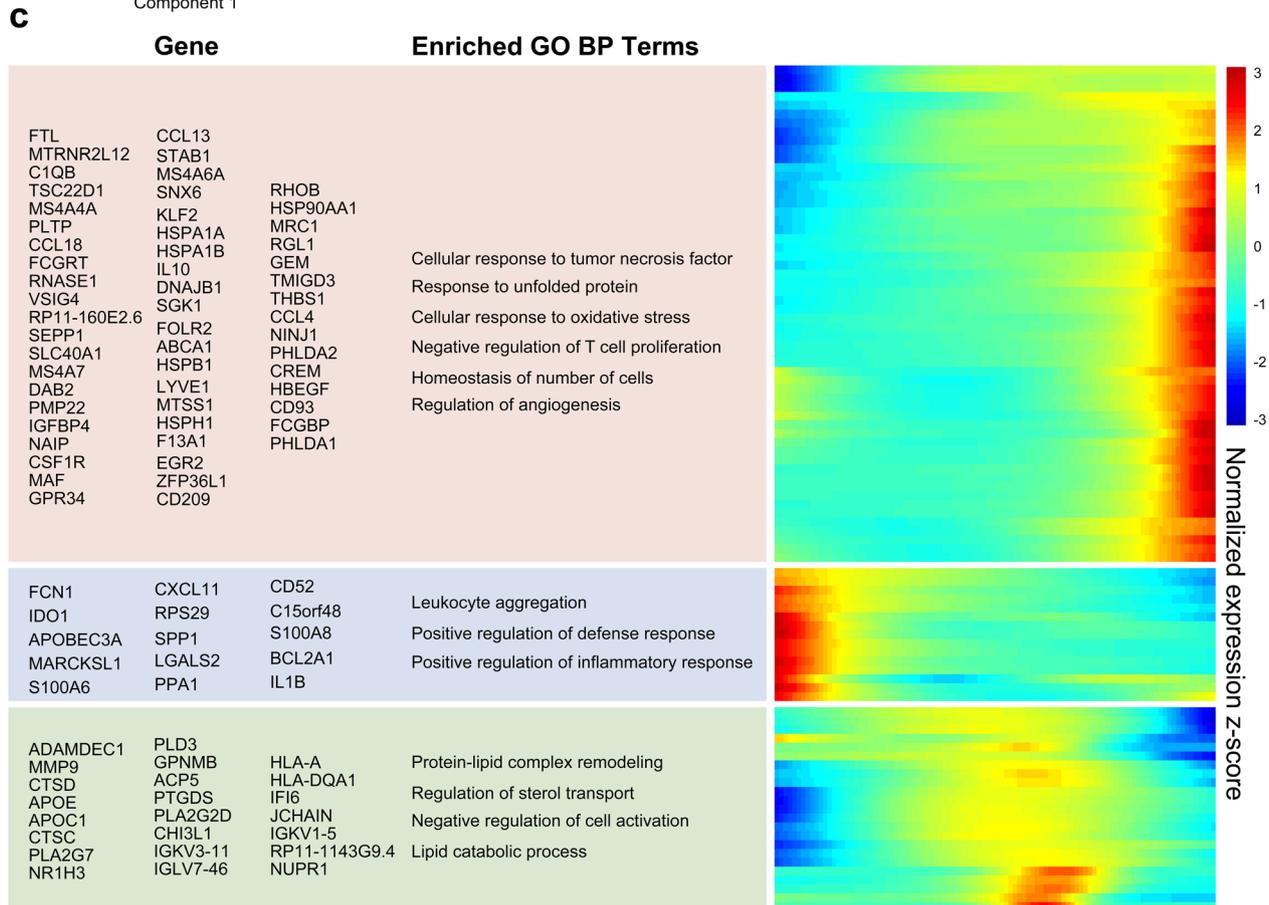
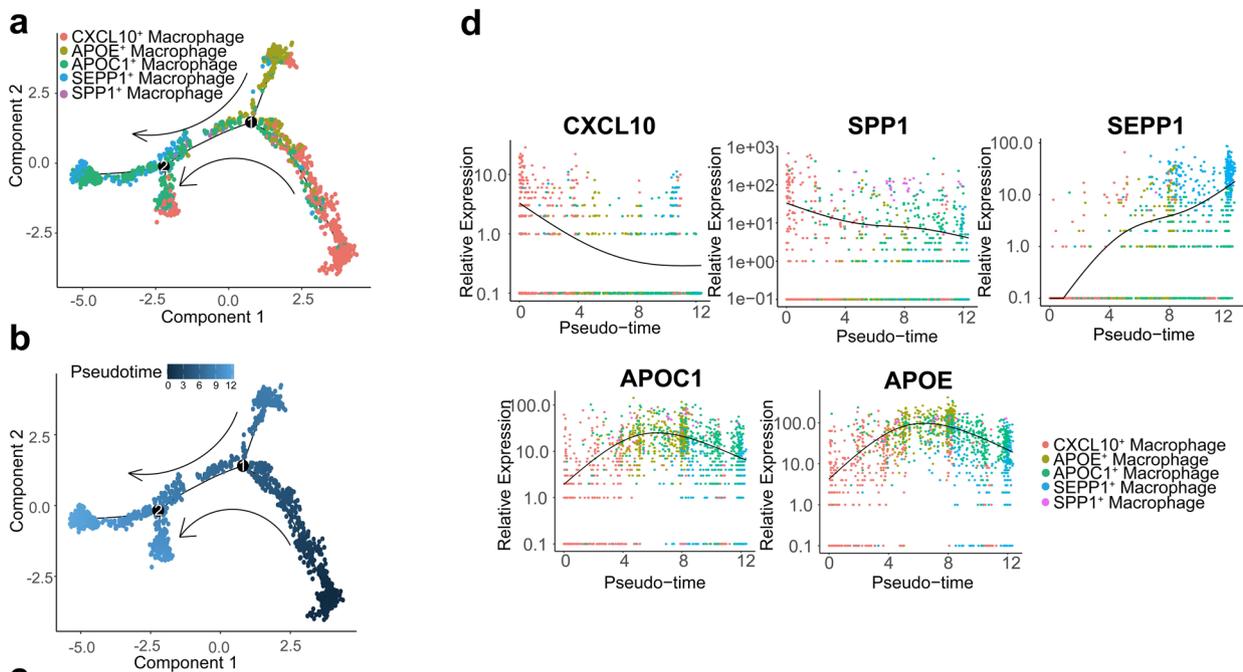


Fig. 3 (See legend on previous page.)

significantly during these state transitions. Initially, macrophages are involved in processes such as positive regulation of defense response, positive regulation of inflammatory response, and leukocyte aggregation. As they transition, their functions shift toward regulation of angiogenesis and negative regulation of T cell proliferation (Fig. 3d).

Cell–cell communication in the TLS associated regions

To investigate the communication network and mechanisms driving the recruitment of cells—particularly T and B cells—into the TLS associated regions, we used CellChat, an R-based tool for analyzing cell–cell communication. Notably, we observed that KRT86⁺ CD8⁺ T cell, HISTIH4C⁺ cycling CD8⁺ T cell, IFNG⁺ CD8⁺ T cell, and IGKV3-20⁺ B cell exhibited stronger connections with other cells, suggesting that these cells play key roles in the TLS associated regions (Supplementary Fig. 3). When analyzing the signaling pathways active in the TLS associated regions, we focused on the CXCL signaling pathway, which plays a central role in cell recruitment [23, 24], we found that CCL21⁺ fibroblasts and APOD⁺ fibroblasts were highly communication with other cell types, particularly T and B cells, which indicates that the recruitment of T cells (CCL4L2⁺ CD8⁺ T cell, KRT86⁺ CD8⁺ T cell, IFNG⁺ CD8⁺ T cell, HISTIH4C⁺ cycling CD8⁺ T cell), B cells (IGKV3-20⁺ B cell) were primarily mediated by these two fibroblast subtypes (Fig. 4a).

To further understand the main receptor–ligand interactions between CCL21⁺ fibroblast, APOD⁺ fibroblast, and T/B cells, we analyzed the signal intensity of the receptor–ligand pairs in the CXCL signaling pathway. The results revealed that the CXCL12–CXCR4 ligand–receptor pair exhibited the stronger signal intensity, which suggests that these two types of fibroblasts primarily recruit KRT86⁺ CD8⁺ T cells, IL7R⁺ regulatory T cells, IGKV3-20⁺ B cells, IFNG⁺ CD8⁺ T cells, CXCL13⁺ follicular helper (Tfh) T cells and HISTIH4C⁺ cycling CD8⁺ T cells through this ligand–receptor interaction (Fig. 4b). For further validating our findings regarding TLS-associated regions, we utilized four independent spatial transcriptomics datasets from breast cancer patients. One dataset consists of single-cell resolution sequencing data from the Xenium platform, while the other three datasets are from the Visium platform. These datasets allow us to directly obtain the spatial locations of cell types or infer their positions using deconvolution algorithms, and they also enable us to assess the strength of the CXCL12–CXCR4 ligand–receptor interaction in a spatial context. Each platform has its advantages: the Xenium platform

provides spatial distribution data at the single-cell level but measures a limited number of genes—specifically, only 313 in the dataset we used. In contrast, the Visium platform measures a large number of genes, but at a lower resolution, typically capturing multiple cells within a single spot. Combining data from these two platforms allows for better validation of our findings. To assess the similarity of these cells to the 18 types of TLS-associated cells we annotated (i.e., to infer cell types), we leveraged the strengths of a single-cell large language model, Geneformer [25], for semantic understanding cellular features. After fine-tuning Geneformer on our annotated 18 cell types in TLS-associated regions, we used this model to predict the similarity of single cells measured by the Xenium platform to the 18 types of TLS region cell types, thereby inferring their identities. We found that there are CCL21⁺/APOD⁺ high-expressing fibroblast cells (Fig. 5a–b) in the TLS regions (annotated by pathologists). Additionally, we employed stLearn to calculate the spatial communication intensity of the CXCL12–CXCR4 ligand–receptor pair (Fig. 5c). Our analysis revealed that CXCL12–CXCR4 exhibited significantly higher communication intensity in these regions compared to others, further supporting our findings. Furthermore, using the annotated single-cell data from the 18 types of TLS-associated cells, we applied deconvolution to the spots in the other three Visium datasets. We discovered that CCL21⁺/APOD⁺ high-expressing fibroblast cells are present within or near the TLS regions (Fig. 6a–c, Supplementary Fig. 4–6), and the communication intensity of the CXCL12–CXCR4 receptor–ligand pair is significantly higher than in other areas (Fig. 6a–c). These independent validation sets enhance the credibility of our main findings and provide a theoretical basis for developing novel strategies to improve the immunogenicity of low-immunogenicity breast cancer patients, ultimately enhancing tumor immunotherapy in the future.

Discussion

It has been discovered that breast cancer patients who respond to immunotherapy often have cancer tissues infiltrated by a high proportion of TLSs [26, 27]. Numerous studies have shown that high TLS infiltration is associated with improved patient prognosis [20, 21]. Therefore, further research on TLS is crucial for understanding why it performs such functions and how to induce it to promote immunotherapy. Considering that TLS, as a functional structural region, is difficult to analyze with conventional techniques, and that high-precision analysis of TLS is urgently needed, we

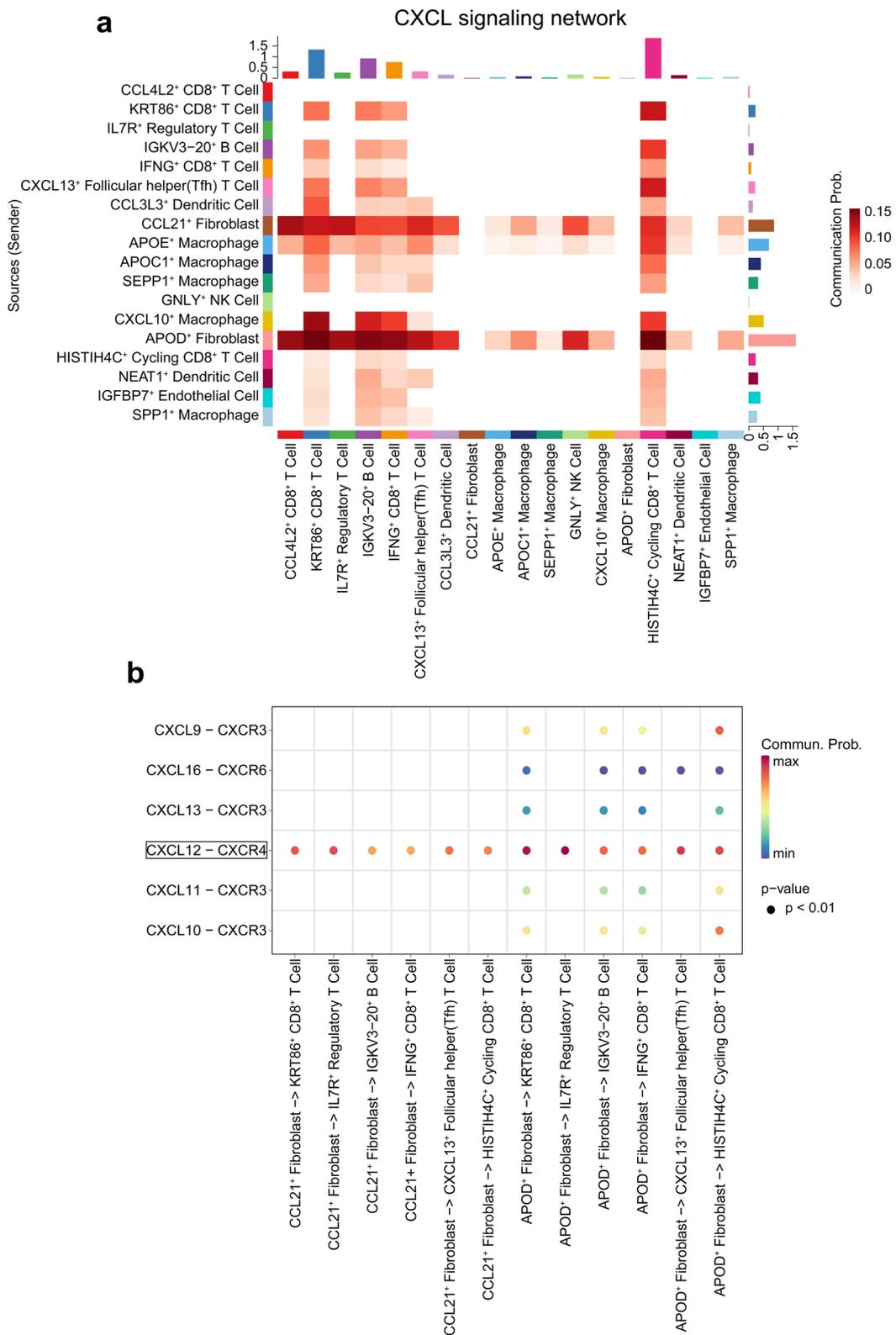
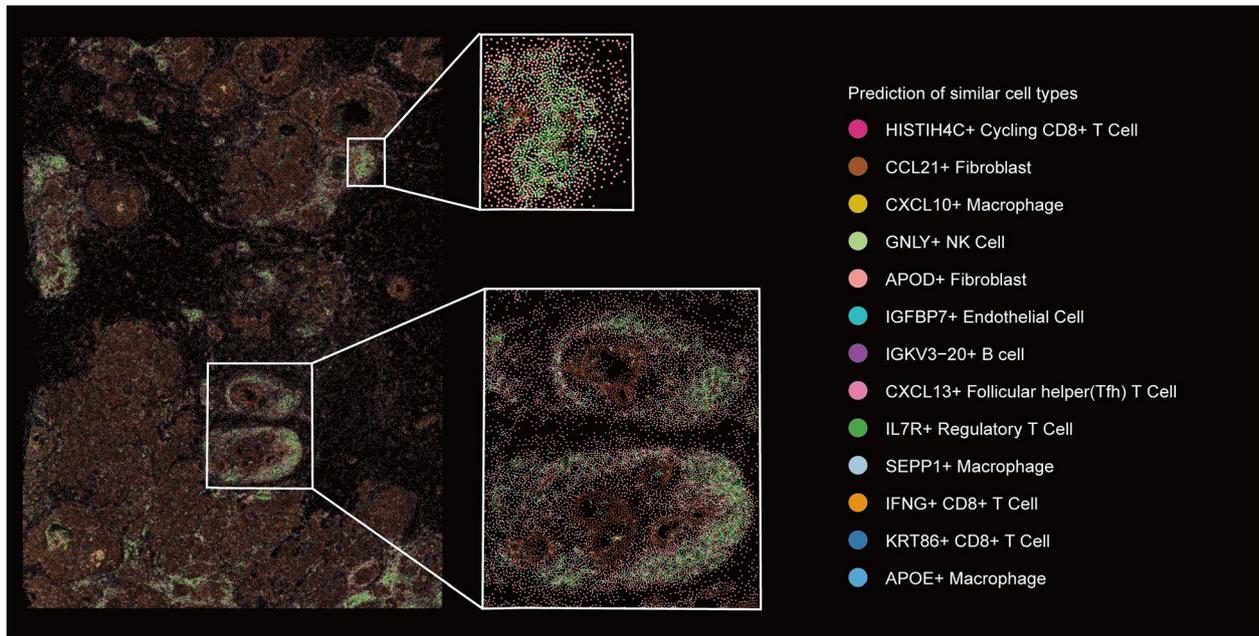
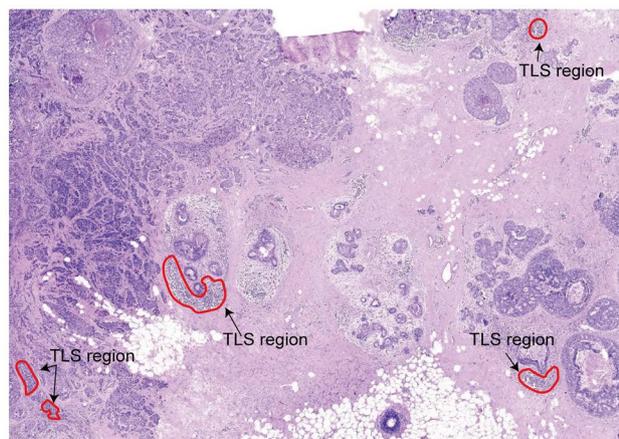


Fig. 4 Cell-cell communications in the TLS associated regions. **a** The strength of the CXCL signaling pathway among each cell type in the TLS-associated regions. **b** Ligand-receptor analysis of the CXCL signaling pathway between fibroblasts and T/B cells is presented

a



b



c CXCL12--CXCR4 ligand receptor pair

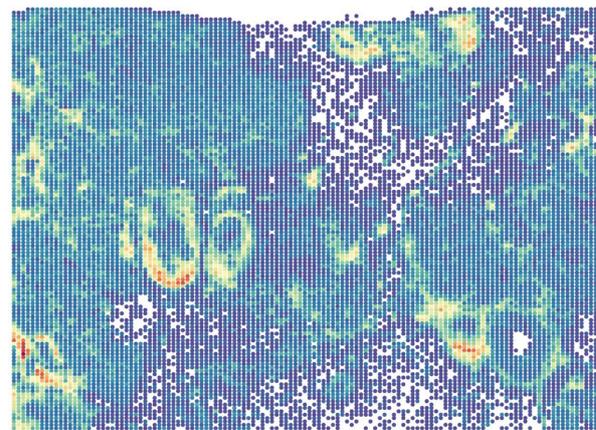
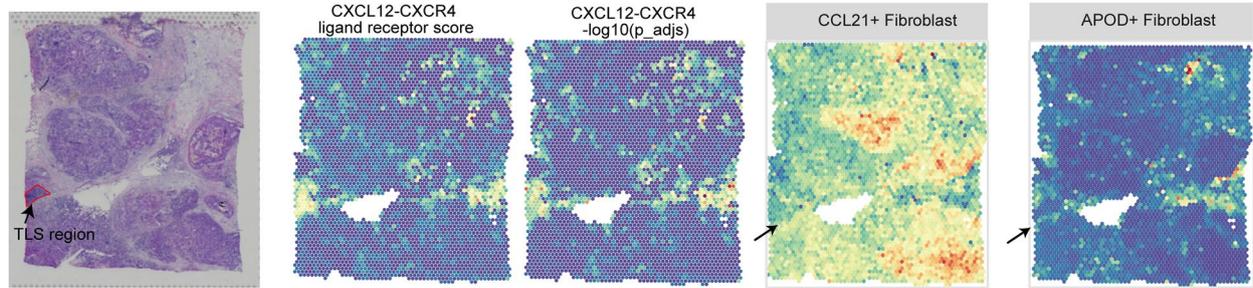


Fig. 5 Predicted cell types and CXCL12-CXCR4 ligand-receptor pair communication strength in Xenium breast cancer data. **a** The cell types in Xenium breast cancer dataset were predicted using Geneformer, which was trained on our annotated 18 cell types; the white box highlights enlarged TLS regions. **b** The H&E image slice of the Xenium breast cancer dataset, with TLS regions annotated by a pathologist indicated by the red circle. **c** The communication strength of the CXCL12-CXCR4 ligand-receptor pair

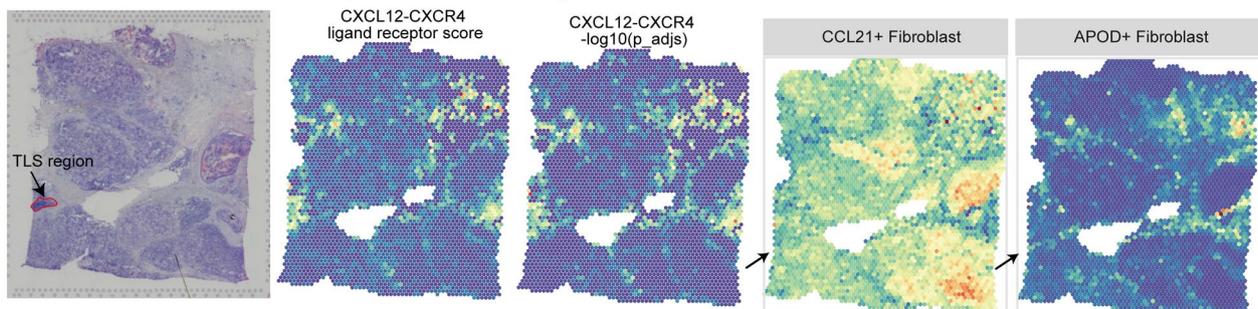
focused on this in our study. We utilized breast cancer scRNA-seq data, spatial transcriptome data and a TLS-specific gene signature to investigate cells in the TLS associated regions, identify differential genes between these regions and others, analyze the macrophage lineage trajectory, and explore cell communication within the TLS. Our findings indicate that there

are approximately 18 distinct cell types in the TLS-associated regions, and that macrophages in these areas undergo significant lineage and functional transformations. The recruitment of T cells and B cells is critical to the formation of TLS-associated regions [19], our multiple lines of evidence shows that two types of fibroblasts (CCL21⁺ fibroblast, APOD⁺ fibroblast) play a key

a 10x Human Breast Cancer (Block A Section 1)



b 10x Human Breast Cancer (Block A Section 2)



c GSM6177599

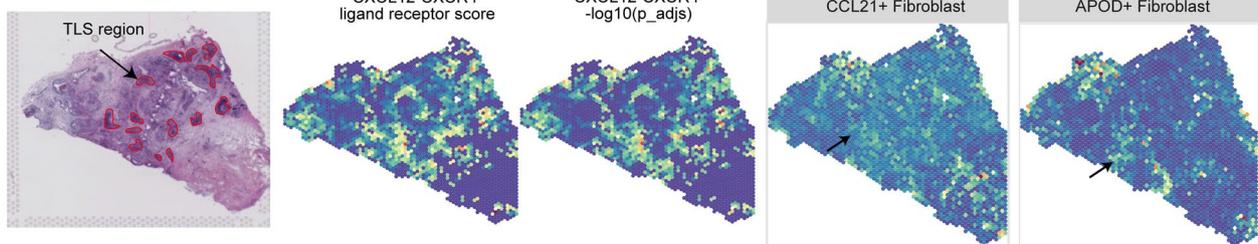


Fig. 6 Inferred CCL21⁺ fibroblast and CXCL12-CXCR4 ligand-receptor pair communication strength in Visium breast cancer data. **a** The H&E image of 10× human breast cancer (Block A, Sect. 1), alongside the CXCL12-CXCR4 ligand-receptor communication strength; the TLS region is annotated by a pathologist and indicated by the red circle. **b** The H&E image of 10× human breast cancer (Block A, Sect. 2), with CXCL12-CXCR4 ligand-receptor communication strength, the TLS region is highlighted in red as annotated by a pathologist. **c** The H&E image of GSM6177599 and the CXCL12-CXCR4 ligand-receptor communication strength, the TLS region is indicated by a red circle annotated by a pathologist

role in this process, primarily through the CXCL12-CXCR4 ligand-receptor pair to recruit T cells and B cells into TLS-associated regions. In immunotherapy, the benefit to breast cancer patients are closely related to the composition and activity of immune cells in their microenvironment. TLS serves as a potential structure that can enhance the immune activity of the tumor microenvironment [28]. Therefore, it is significant to utilize CCL21⁺/APOD⁺ fibroblasts, which highly express CXCL12 or CXCR4, to induce the infiltration of

T and B cells into the low-immunogenic breast cancer microenvironment. This approach aims to increase the presence of new T/B cells or replace exhausted T cells, thereby improving the immune activity of the microenvironment in breast cancer patients. However, further experimental validation is needed to demonstrate the practicality of this strategy. In summary, these findings contribute to our understanding of the TLS associated regions in breast cancer and have implications for future applications in immunotherapy.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12935-025-03635-y>.

Additional file 1: Fig.1 The performance of the 12-chemokine signature in predicting TLS associated regions (a) Using the 12-chemokine signature calculates the TLS score, which is distributed across the patients' tissue slides. (b) The proportions of major cell types within the TLS of each patient and across different subtypes of patients are presented. (c) A comparison of TLS scores among the various subtypes of breast cancer.

Additional file 2: Fig.2 Survival analysis of cells in the TLS associated regions (a-r) Using specific marker genes of each cell type to calculate infiltration score of each cell type in the TCGA BRCA cohort. Survival analysis was then conducted to assess the impact of each cell type on patients' survival time, with p-values calculated using the log-rank test.

Additional file 3: Fig.3 Cell-cell communications in the TLS associated regions The weights of cell-cell communications in the TLS associated regions were analyzed using the CellChat package, and the heatmap was showing the communication strength.

Additional file 4: Fig.4 Inferred cell proportions of 18 cell types in the 10x human breast cancer Block A, Section1 The proportion of 18 cell types in the 10x human breast cancer Block A section1 spatial transcriptome dataset.

Additional file 5: Fig.5 Inferred proportions of 18 cell types in 10x human breast cancer Block A section2 The proportion of 18 cell types in 10x human breast cancer Block A section1 spatial transcriptome dataset.

Additional file 6: Fig.6 Inferred proportions of 18 cell types in GSM6177599 The proportion of 18 cell types in GSM6177599 spatial transcriptome dataset.

Acknowledgements

We sincerely thank Alexander Swarbrick (Garvan Institute of Medical Research) for sharing their single-cell RNA sequencing data in the GEO database and spatial transcriptome data in the Zenodo database.

Author contributions

KK.F., L.L., Y.J. Y conceived and designed the study and drafted the manuscript. KK.F., DQ.F, DG.W, AQL., FY.W., SF.D., M.L.S., CZ.Q. performed the analysis of the data. The author(s) read and approved the final manuscript.

Funding

The study was supported by the First-class discipline innovation-driven talent program of Guangxi Medical University, the National Natural Science Foundation of China (No. 82260554), the Natural Science Foundation of Guangxi Province (No. 2023GXNSFBA026092 and No. 2024GXNSFAA010100), the Key Research and Development Program of Guangxi Province (No. 2023AB22116), and the Guangxi medical and health appropriate technology development and application project (No. S2018013).

Availability of data and materials

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 24 September 2024 Accepted: 2 January 2025

Published online: 13 January 2025

References

- Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol*. 2010;28(10):1684–91.
- Barzaman K, Karami J, Zarei Z, Hosseinzadeh A, Kazemi MH, Moradi-Kalbolandi S, Safari E, Farahmand L. Breast cancer: biology, biomarkers, and treatments. *Int Immunopharmacol*. 2020;84: 106535.
- Chahat NN, Murti Y, Yadav S, Rawat P, Dhiman S, Kumar B. Advancements in targeting tumor suppressor genes (p53 and BRCA 1/2) in breast cancer therapy. *Mol Divers*. 2024. <https://doi.org/10.1007/s11030-024-10964-z>.
- Umar A, Dunn BK, Greenwald P. Future directions in cancer prevention. *Nat Rev Cancer*. 2012;12(12):835–48.
- Egen JG, Ouyang W, Wu LC. Human anti-tumor immunity: insights from immunotherapy clinical trials. *Immunity*. 2020;52(1):36–54.
- Chandrasekaran J, Elumalai S, Murugesan V, Kunjiappan S, Pavada P, Theivendren P. Computational design of PD-L1 small molecule inhibitors for cancer therapy. *Mol Divers*. 2023;27(4):1633–44.
- Gooden MJ, de Bock GH, Leffers N, Daemen T, Nijman HW. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis. *Br J Cancer*. 2011;105(1):93–103.
- Loi S, Sirtaine N, Piette F, Salgado R, Viale G, Van Eenoo F, Rouas G, Francis P, Crown JP, Hitre E, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02–98. *J Clin Oncol*. 2013;31(7):860–7.
- Emens LA. Breast cancer immunotherapy: facts and hopes. *Clin Cancer Res*. 2018;24(3):511–20.
- Engelhard VH, Rodriguez AB, Mauldin IS, Woods AN, Peske JD, Slingluff CL Jr. Immune cell infiltration and tertiary lymphoid structures as determinants of antitumor immunity. *J Immunol*. 2018;200(2):432–42.
- Collbeck EJ, Ager A, Gallimore A, Jones GW. Tertiary lymphoid structures in cancer: drivers of antitumor immunity, immunosuppression, or bystander sentinels in disease? *Front Immunol*. 1830;2017:8.
- Aoyama S, Nakagawa R, Mule JJ, Mailloux AW. Inducible tertiary lymphoid structures: promise and challenges for translating a new class of immunotherapy. *Front Immunol*. 2021;12: 675538.
- Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, Thennavan A, Wang C, Torpy JR, Bartonicek N, et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet*. 2021;53(9):1334–47.
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)*. 2021;2(3):100141.
- Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
- Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979–82.
- Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using Cell Chat. *Nat Commun*. 2021;12(1):1088.
- Schumacher TN, Thommen DS. Tertiary lymphoid structures in cancer. *Science*. 2022;375(6576):9419.
- Sautes-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat Rev Cancer*. 2019;19(6):307–25.
- Baxevasis CN, Fortis SP, Perez SA. The balance between breast cancer and the immune system: challenges for prognosis and clinical benefit from immunotherapies. *Semin Cancer Biol*. 2021;72:76–89.
- Munoz-Erazo L, Rhodes JL, Marion VC, Kemp RA. Tertiary lymphoid structures in cancer—considerations for patient prognosis. *Cell Mol Immunol*. 2020;17(6):570–5.
- Prabhakaran S, Rizk VT, Ma Z, Cheng CH, Berglund AE, Coppola D, Khalil F, Mule JJ, Soliman HH. Evaluation of invasive breast cancer samples using a

- 12-chemokine gene expression score: correlation with clinical outcomes. *Breast Cancer Res.* 2017;19(1):71.
23. Tokunaga R, Zhang W, Naseem M, Puccini A, Berger MD, Soni S, McSkane M, Baba H, Lenz HJ. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation—a target for novel cancer therapy. *Cancer Treat Rev.* 2018;63:40–7.
 24. Korbecki J, Kojder K, Siminska D, Bohatyrewicz R, Gutowska I, Chlubek D, Baranowska-Bosiacka I. CC Chemokines in a Tumor: a review of pro-cancer and anti-cancer properties of the ligands of receptors CCR1, CCR2, CCR3, and CCR4. *Int J Mol Sci.* 2020. <https://doi.org/10.3390/ijms21218412>.
 25. Theodoris CV, Xiao L, Chopra A, Chaffin MD, Al Sayed ZR, Hill MC, Mantineo H, Brydon EM, Zeng Z, Liu XS, et al. Transfer learning enables predictions in network biology. *Nature.* 2023;618(7965):616–24.
 26. Li H, Wang J, Liu H, Lan T, Xu L, Wang G, Yuan K, Wu H. Existence of intratumoral tertiary lymphoid structures is associated with immune cells infiltration and predicts better prognosis in early-stage hepatocellular carcinoma. *Aging (Albany NY).* 2020;12(4):3451–72.
 27. Germain C, Gnjjatic S, Tamzalit F, Knockaert S, Remark R, Goc J, Lepelley A, Becht E, Katsahian S, Bizouard G, et al. Presence of B cells in tertiary lymphoid structures is associated with a protective immunity in patients with lung cancer. *Am J Respir Crit Care Med.* 2014;189(7):832–44.
 28. Dieu-Nosjean MC, Giraldo NA, Kaplon H, Germain C, Fridman WH, Sautes-Fridman C. Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. *Immunol Rev.* 2016;271(1):260–75.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.