Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Integrative Transcriptomic Analysis Identifies CXCL12 as a Candidate Hub Gene Associated with CD4$^+$ T-Cell Immune Network Remodeling in Secondary Lymphedema

Xiaoxiao Huang[1,2], Chuan Lu[1,2], Zhibiao Tan[2,3], Jianwen Xu[1,2]

[1]Department of Rehabilitation Medicine, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, People's Republic of China; [2]The First Clinical Medical College, Guangxi Medical University, Nanning, Guangxi, People's Republic of China; [3]Department of Rehabilitation Medicine, Guigang City People's Hospital, Guigang, Guangxi, People's Republic of China

Correspondence: Jianwen Xu, Email xujianwen@gxmu.edu.cn

Introduction: Secondary lymphedema is accompanied by a persistent chronic inflammatory response and remodeling of the immune microenvironment. Among them, CD4$^+$ T cells and their functional subgroups play a crucial role. This study identified potential candidate molecules in the CD4$^+$ T cell-related immune microenvironment, providing new research clues for subsequent mechanism-oriented intervention studies.

Methods: In this study, we used the hdWGCNA method to establish a co-expression network of CD4$^+$ T cells in secondary lymphedema, enabling the identification of key immune regulatory modules. We performed differential expression analysis using bulk transcriptomics and systematically analyzed the immune microenvironment of secondary lymphedema using CIBERSORT, Hallmark ssGSEA, and CellChat. This study primarily relies on computational analysis, employing numerous bioinformatics tools and algorithms for data integration and interpretation, aiming to explore the potential regulatory role of CD4$^+$ T cells in the immune microenvironment from a holistic perspective. To further verify the results of the computational analysis, the selected key genes were experimentally validated using qPCR in a mouse model. Furthermore, computational virtual cell knockout analysis was applied for the first time in single-cell analysis of secondary lymphedema to assess the potential regulatory roles of key candidate genes in the immune regulatory network. This method provides more computational predictions and hypotheses for the study, and further experimental validation is needed to establish the actual functions and mechanisms of action of these genes.

Results: This study identified 11 cell types using single-cell sequencing and found that CD4$^+$ T cells transitioned from a resting to an activated state in SL tissue (resting CD4$^+$ T cells decreased from 43.9% to 22.8%, while activated CD4$^+$ T cells increased from 17.7% to 41.9%). Cell communication analysis showed that CD4$^+$ T cells are key signal receivers in the microenvironment, with the CXCL12–CXCR4 axis significantly enriched in communication between fibroblasts and CD4$^+$ T cells. Six co-expressed modules were identified using hdWGCNA, with CXCL12 associated with both the immune regulation-related M1 module and the T cell signaling-related M4 module. CXCL12 was significantly upregulated in SL tissue, exhibiting good diagnostic efficacy (AUC = 0.844), and was positively correlated with follicular helper T cells and negatively correlated with resting CD4$^+$ memory T cells. In addition, this study validated the upregulation of Cxcl12 expression in a mouse model; virtual knockout analysis showed that only about 2% of gene expression was significantly altered after CXCL12 deletion, involving immune regulation and matrix remodeling-related genes such as AXL, IFITM3, and IGFBP7.

Conclusion: This study reveals CXCL12 as a potential candidate molecule for the immune regulation of secondary lymphedema, which may mediate the interaction between fibroblasts and CD4$^+$ T cells through the CXCL12–CXCR4 axis. However, the causal relationship still needs functional verification.

Keywords: CXCL12, CD4$^+$ T cells, immune remodeling, immune network, secondary lymphedema

https://doi.org/10.2147/JIR.S597037
Received: 29 January 2026
Accepted: 25 April 2026
Published: 15 May 2026

Introduction

Secondary lymphedema (SL) is a common chronic complication following tumor surgery, radiotherapy, infection, or trauma.[1] Its core pathological basis is impaired lymphatic drainage and the resulting retention of tissue fluid, chronic inflammatory response, and progressive tissue remodeling.[2] As the disease progresses, patients may experience persistent swelling, fibrosis of the skin and subcutaneous tissue, abnormal fat deposition, and recurrent infections, which severely affect limb function and quality of life.[3,4] Currently, clinical treatment mainly focuses on physical methods to reduce swelling and symptomatic management, while effective targeted interventions targeting its molecular mechanisms remain limited.

In recent years, advances in research have reshaped our understanding of SL, which is no longer viewed simply as a fluid retention disorder resulting from mechanical damage to lymphatic vessels, but is now considered a complex pathological process characterized by abnormalities in the immune microenvironment. Research evidence suggests that the immune microenvironment plays a crucial role in the occurrence and progression of SL, with various immune cell subsets and their secreted cytokines jointly participating in disease-related inflammatory responses and tissue remodeling.[5] In this process, $CD4^+$ T cells are considered important immunoregulatory cells mediating immune abnormalities and tissue responses. Their significant infiltration in tissues affected by SL and the imbalance of their functional subsets are closely related to chronic inflammation, tissue remodeling, and lymphatic dysfunction.[6] A case-control study of human secondary lymphedema tissue further confirmed that $CD4^+$ T cells showed significant infiltration in the affected skin, providing direct human histological evidence for thismechanism.[7] Previous studies have shown that $CD4^+$ T cells can cooperate with macrophages to promote lymphangiogenesis and contribute to the development of lymphedema.[8] In mouse models, depletion of $CD4^+$ T cells significantly reduces edema formation and lymphatic remodeling, highlighting their critical role in disease progression.[8] In addition, antigen-driven expansion of $CD4^+$ T cells has been observed in lymphedematous tissues.[9]

Following lymphatic system injury, $CD4^+$ T cells can be activated by antigen-presenting cells in regional draining lymph nodes and subsequently migrate to the affected skin tissue. Their differentiation state often exhibits immune characteristics combining T helper 1 (Th1) and T helper 2 (Th2) phenotypes and is considered closely related to the progression of fibrosis and the formation of lymphatic dysfunction.[10] Animal model studies have further shown that even with a relatively low overall number of $CD4^+$ T cells, a lymphedema phenotype characterized by swelling, inflammation, and fibrosis can still occur.[11] In other words, $CD4^+$ T cells do not simply rely on numerical superiority, but may amplify inflammation and tissue remodeling responses in local tissues through specific immune regulatory networks. In addition, the migration and functional status of $CD4^+$ T cells are regulated by lymphatic endothelial cell (LEC) signaling pathways. Abnormal sphingosine-1-phosphate (S1P) signaling has been associated with altered localization of $CD4^+$ T cells in local tissues and the Th1/Th2 differentiation balance, thereby exacerbating the inflammatory response and lymphatic dysfunction.[5]

Although $CD4^+$ T cells are thought to be involved in the development of SL, whether they function through a specific transcriptional regulatory network in disease states, and the synergistic regulatory mechanisms between them and other immune cells and tissue structural cells, still need to be systematically elucidated. Current molecular research on SL mainly relies on bulk transcriptome analysis or candidate factor studies.[12] Previous transcriptomic studies have also identified $CD4^+$ T cell–associated hub genes involved in immune dysregulation during lymphedema progression.[13] These methods have some value in revealing overall expression trends, but due to the lack of cell resolution, it is difficult to distinguish the specific transcriptional features of different immune cell populations, and even more difficult to analyze the immune regulatory network centered on $CD4^+$ T cells. Furthermore, identifying the core molecules that truly reside at key regulatory nodes in the remodeling of the immune microenvironment among a large number of differentially expressed genes (DEGs) remains a scientific problem that urgently needs to be solved.

In this study, we focused on the chemokine C-X-C motif chemokine ligand 12 (CXCL12), identifying it as a potential molecule in the $CD4^+$ T cell-associated immune regulatory network. Previous studies have shown that CXCL12 can directly induce the migration and tubule formation of LECs, promoting lymphangiogenesis in vivo.[14] Simultaneously, CXCL12 also participates in regulating the interaction between immune cells and stromal cells through its specific

receptor C-X-C chemokine receptor type 4 (CXCR4).[15] Furthermore, in vitro experiments have confirmed that CXCL12 has the ability to stimulate $CD4^+$ T cells to migrate across the endothelial cell layer.[16] However, in the specific pathological context of SL, the precise role of CXCL12 in the immune network related to $CD4^+$ T cells has not been fully elucidated. Therefore, this study aims to, based on bioinformatics analysis and experimental verification, preliminarily explore the potential role by which CXCL12 regulates the immune response mediated by $CD4^+$ T cells in the SL, and thereby affects the process of chronic inflammation and tissue fibrosis.

Single-cell RNA sequencing (scRNA-seq) technology offers new possibilities for analyzing the complex immune microenvironment at the cellular resolution level.[17] However, single-cell data are characterized by high dimensionality, high noise, and strong heterogeneity,[18] and relying solely on differential expression analysis often makes it difficult to identify core molecules and key transcriptional procedures that have real regulatory significance in the disease process from a massive amount of genes. Therefore, this study introduced high-dimensional weighted gene co-expression network analysis (hdWGCNA)[19] focusing on the $CD4^+$ T cell population, and combined it with cell communication analysis, feature screening models, and network perturbation assessment to gradually converge from complex transcriptional heterogeneity and locate candidate molecules at key nodes in the immune regulatory network. Furthermore, by introducing a computational simulation-based virtual cell knockout strategy, we further evaluated the potential regulatory roles of key candidate molecules in network structure stability and cell communication patterns, and explored their functional significance in secondary lymphedema-related immune remodeling at the systemic level.

This study provides a new perspective for understanding the immune regulation mechanism in secondary lymphedema at the network level, and lays the foundation for subsequent functional studies and the exploration of potential intervention strategies.

Methods

Transcriptome Data Sources

We searched the Gene Expression Omnibus (GEO) database using "lymphedema" as the keyword. Then we obtained the raw expression data of the GSE132936, GSE255848, and GSE302563 datasets for subsequent analysis. The GSE132936 dataset analyzed adipose-derived mesenchymal stem cells (ASCs) from 10 patients with malignant tumor-associated SL of the limbs. The study employed a paired design, with ASCs derived from adipose tissue in the lymphedema thigh of the same patient serving as the disease group, and ASCs derived from normal upper abdominal adipose tissue serving as the control group. The GSE255848 dataset was constructed based on adipose tissue samples and included lymphedema patients (n = 5) and normal controls (n = 5). GSE302563 is a bulk RNA sequencing dataset containing subcutaneous tissue samples collected from patients with stage III SL (n = 5) and normal controls (n = 3) from the International Lymphatic Society.

scRNA-Seq Data Processing and Analysis

The scRNA-seq data used in this study were derived from a previously published study by Liu et al[20] The dataset included adipose tissue samples from five patients with stage III SL and four healthy controls. scRNA-seq analysis was performed after isolating the stromal vascular fraction (SVF). The raw sequencing data were stored in FASTQ format and uploaded to the Human Genome Sequence Archive (GSA-Human) of the China National Center for Bioinformation (CNCB) (https://ngdc.cncb.ac.cn/), with accession number HRA000901. This study downloaded and reanalyzed the dataset for subsequent research.

The obtained scRNA-seq data were analyzed using Seurat (v5.1) in the R environment (R version 4.4.1). The filtered expression matrix was read using the Read10X function. During quality control, cells with fewer than 200 genes and genes expressed in fewer than 3 cells were first screened out. Simultaneously, the proportions of mitochondrial genes (percent.mt), ribosome genes, and erythrocyte-related genes (percent.HB) were calculated for each cell. Cells with a mitochondrial gene proportion exceeding 10% or a red blood cell gene proportion exceeding 5% were excluded. Additionally, cells expressing fewer than 200 or more than 5000 genes were also removed. After quality control, 70,114 cells were retained for subsequent analysis. To remove doublets that may form during sequencing library preparation,

`DoubletFinder` (v2.0.6) was used. Based on the principal component analysis (PCA) results, the first 20 principal components (PCs = 1:20) were selected for paramSweep, and the optimal pK value was determined by maximizing the MeanBC index. The `modelHomotypic` function was calculated for each cell estimate the proportion of homologous double cells, and the expected number of double cells (nExp_poi) was calculated based on a preset double cell formation rate of 5% (doublet rate = 0.05), which was then further adjusted to the actual nExp_poi.adj used for detection. In the `DoubletFinder` function, pN = 0.25, pK was set to its optimal value, and nExp = nExp_poi.adj, to identify and remove double cells from the data. During data preprocessing, standard quality control procedures were used to reduce the impact of technical noise, but no additional correction was performed using a dedicated ambient RNA decontamination algorithm. After quality control, the data were standardized using the `LogNormalize` method (scale factor = 10,000), and 2000 hypervariable genes were selected using the `vst` method for subsequent analysis. The data were then standardized using the `ScaleData` function, followed by PCA. To eliminate batch effects among different samples, the Harmony algorithm was used to integrate the data based on the PCA results (using default parameters). The integrated low-dimensional space was used for subsequent analysis. A shared nearest neighbor (SNN) graph was constructed based on the first six principal components, and the Louvain algorithm was used for cluster analysis. The clustering stability at different resolutions was evaluated using the clustree function, and a resolution of 0.1 was finally selected for cell clustering. Finally, the `RunUMAP` function was used for Uniform Manifold Approximation and Projection (UMAP) to perform nonlinear dimensionality reduction, and the results were visualized using the `DimPlot` function.

Cell Type Annotation

The `FindAllMarkers` function was used to screen for marker genes in each cell cluster (log2FC > 0.25, adjusted p-value < 0.05). Cell types were annotated using classic marker genes and the CellMarker database, and a total of 11 cell types were identified. To further elucidate T cell heterogeneity, T cell subsets were extracted from the annotated cell population. PCA was re-performed on these subsets, and the top 12 principal components were determined using elbow plots and quantitative criteria for subsequent analysis. Based on the selected principal components, a SNN graph was constructed, multi-resolution clustering (0.1–1.2, step size 0.1) was employed. Cluster stability and hierarchical relationships were evaluated using `clustree`, and considering cluster number and biological interpretability, the optimal resolution was ultimately determined to be 0.5. Marker genes for each subset were screened using the `FindAllMarkers` function (min.pct = 0.25, logfc.threshold = 0.25). Manual annotation was performed using literature reports and typical marker genes from the `CellMarker` database, and `UMAP` plots were generated to illustrate the distribution characteristics of T cell subsets under different conditions.

CellChat Cell Communication Analysis Methods

Cell–cell communication analysis was performed using the `CellChat` R package (v1.6.0). The expression matrix was first normalized using the `LogNormalize` method (scale factor = 10,000). A `CellChat` object incorporating cell type annotations was constructed using the `createCellChat` function. The ligand–receptor interaction database was set to the "Secreted Signaling" category for the human species. Subsequently, the expression data were subsetted using the `subsetData` function, and overexpressed ligands and receptors in each cell type were identified using the `identifyOverExpressedGenes` and `identifyOverExpressedInteractions` functions, respectively. Communication probabilities were then calculated using the `computeCommunProb` function with default parameters (population.size = TRUE, raw.use = TRUE, trim = 0.1). Communication probability was computed using the `triMean` method, and significant interactions were defined as those with a p-value < 0.05 based on 100 permutations. Finally, cell–cell communication at the signaling pathway level was inferred using the `computeCommunProbPathway` function.

High-Dimensional Weighted Gene Co-Expression Network Analysis (hdWGCNA)

hdWGCNA analysis was performed using the hdWGCNA R package (v0.3.20). Metacells were constructed using the `MetacellsByGroups` function with the following parameters: `k = 25` (number of nearest neighbors) and `max_shared = 10` (maximum number of shared cells between metacells), with dimensionality reduction performed using the Harmony

algorithm. Prior to network construction, expression data were normalized using the `NormalizeMetacells` function. The soft-thresholding power was determined based on the scale-free topology criterion ($R^2 > 0.8$) and set to 6. The minimum module size was set to 50, and modules with a correlation greater than 0.75 were merged using a cut height of 0.15. Module eigengenes were calculated, and module–trait relationships were assessed using Pearson correlation analysis. P-values were adjusted using the Benjamini-Hochberg method. Module activity was visualized in UMAP space using module scores, which also revealed module correlations, cell type-module associations, and hub gene network structure.

Integration of Bulk Transcriptomic Datasets and Batch Effect Correction

Subsequently, the expression matrices of the three datasets GSE132936, GSE255848, and GSE302563 were standardized, duplicate genes were removed, and the gene intersection between the datasets was taken as the common gene set. Based on this, a corresponding dataset prefix was added to each sample to construct the merged expression matrix. PCA was performed before and after correction to assess potential batch effects among the datasets. Batch effects were corrected using the `ComBat` method from the `sva` package. Finally, the batch-corrected expression matrix was obtained for subsequent analysis.

Biomarker Screening and Machine Learning Analysis

After batch correction, duplicate genes were removed, the expression values of candidate genes were extracted and grouped according to the sample annotation information. We used the Wilcoxon rank-sum test ($P < 0.05$) to screen for DEGs, and visualized their expression patterns through heatmaps and box plots. Subsequently, we conducted Receiver Operating Characteristic (ROC) curve analysis using the `pROC` software package, calculating the area under the curve (AUC) value and its bootstrap confidence interval, which were used to evaluate its classification performance. Genes with AUC > 0.7 were included in subsequent analyses. Machine learning models were constructed using the `caret` framework. The dataset was split into training (70%) and testing (30%) sets using stratified random sampling. 5-fold repeated cross-validation (repeated 5 times) was performed within the training set, with all preprocessing steps performed within each training fold to prevent data leakage. Multiple classifiers were evaluated, including PLS, Random Forest (RF), Decision Tree (DTS), Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbors (KNN), XGBoost, Gradient Boosting Machine (GBM), Neural Network, and Generalized Linear Model Boosting (glmBoost). Variable importance was calculated using the `DALEX` package to identify key predictive genes.

Chromosomal Localization and Visualization Analysis of Key Genes

Chromosomal localization analysis was performed on the key genes identified through screening. Gene location information was obtained based on the Human Genome Reference Consortium Build 38 (hg38), and Circos diagrams were drawn using the `circlize` package in R to show the distribution characteristics of key genes on different chromosomes.

Immune Cell Infiltration Analysis

The Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) algorithm (v1.03) was used for immune cell deconvolution analysis on the transcriptome data, and the relative proportions of immune cells were estimated based on the linear kernel ν-support vector regression model. Statistical significance was evaluated through permutation tests.

Assessment of Single-Sample Gene Set Enrichment Analysis (ssGSEA) Pathway Activity and Its Correlation with Characteristic Genes and Immune Cell Infiltration

The GSVA package in R was used to perform ssGSEA on the normalized expression matrix to assess pathway activity in each sample. The expression matrix was averaged to merge duplicate genes using the `avereps` function from the `limma` package. The gene set file was in GMT format and read through `GSEABase`. The gene set size thresholds were set to `minSize = 5` and `maxSize = 500`, and the ssGSEA scores were normalized to 0 to 1. Based on this, the

expression levels of the characteristic genes screened in Biomarker Screening and Machine Learning Analysis were extracted. Subsequently, these levels were combined with the ssGSEA pathway score and the proportion of immune cell infiltration calculated by CIBERSORT for subsequent correlation analysis. Only the common samples in the two groups of data were included for subsequent analysis. We conducted a Spearman rank correlation analysis to examine the correlation between the expression of characteristic genes and the ssGSEA pathway score as well as the infiltration of immune cells. The correlation coefficient and P value were calculated, and the significance level was marked with "*, **, ***" ($P<0.05$, $P<0.01$, $P<0.001$). Heatmaps and the `linkET` package were used to visualize the results, illustrating the overall association patterns among pathway activity, immune cell infiltration, and characteristic genes.

Animals and Establishment of the SL Mouse Model

Male C57BL/6J mice (6–8 weeks old, 20 ± 2.5 g; supplier: Huachuang Xinno) were used in this study. All animals were maintained under specific pathogen-free (SPF) conditions with controlled environmental parameters (temperature 22–24 °C, 12-h light/dark cycle) and had free access to food and water. Mice were acclimatized for at least 7 days prior to experimentation. Animals were randomly divided into a sham-operated control group and a SL model group using a random number table (n = 6 per group). The sample sizes for this study were determined in accordance with the principles of the the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines 2.0 guidelines 2.0, with an emphasis on the 3Rs principle of Reduction to minimize the number of animals used while ensuring reliable data acquisition.[21] As this was a model characterization study comparing a sham-operated group with a lymphedema model group, the sample size was selected based on previous experience and published literature in the field. Specifically, a group size of n=5-6 for physiological measurements (hind limb volume) has been widely reported as sufficient to detect the substantial and consistent volumetric changes induced by lymphatic disruption.[22] For exploratory histological and molecular analyses, a minimum of n=3 mice per group is standard practice for detecting robust morphological and gene expression changes associated with surgical modeling.[23] All reported animal numbers represent the final sample sizes after accounting for experimental attrition.

Male mice were chosen for this experiment to reduce the potential interference of hormonal fluctuations in female mice with immune and inflammatory responses. Hormonal fluctuations in female mice during their estrous cycle may cause variability in immune responses, leading to inconsistent experimental results.[24] By using male mice, we are able to reduce these confounding factors, ensuring the consistency and reproducibility of experimental results. This practice is widely used in immunology-related research to improve experimental stability and comparability.

Establishment of the SL model was performed as follows. Mice were anesthetized using inhalational isoflurane delivered via a precision vaporizer (induction at 4–5% and maintenance at 2–3% in oxygen). Adequate depth of anesthesia was confirmed by the absence of pedal withdrawal reflexes. After routine skin preparation and disinfection of the right groin and hind limb regions, Evans blue solution (0.1%, 0.1 mL) was injected subcutaneously into the right paw pad to facilitate intraoperative identification of lymphatic vessels and lymph nodes. After 5–10 minutes, a circular skin incision was made in the right inguinal region, and the ipsilateral superficial inguinal, popliteal, and deep inguinal lymph nodes, as well as femoral lymphatic vessels, were carefully exposed and surgically removed to induce lymphatic drainage impairment. Electrocoagulation was used to minimize intraoperative bleeding. In the sham group, only a circular skin incision was made without disruption of lymphatic vessels or lymph nodes, followed by closure with 6–0 sutures. At postoperative day 7, mice were deeply anesthetized with 5% isoflurane in oxygen and euthanized by cervical dislocation, in accordance with the American Veterinary Medical Association (AVMA) Guidelines for the Euthanasia of Animals. Skin and subcutaneous tissues from standardized anatomical sites of the affected hind limb were collected, rapidly frozen in liquid nitrogen, and stored at −80°C for subsequent analyses. During outcome assessment and data analysis, investigators were blinded to group allocation.

Measurement of Edema Volume

On postoperative days 0, 3, and 6, the length, width, and height of the hind limb were recorded using calipers, and the cross-sectional area and volume of hind limb edema were estimated accordingly. The cross-sectional area (S) and volume

(V) were calculated according to a pre-defined formula, where S represents the cross-sectional area and V represents the hind limb volume.

$$S = \pi \times \text{length} \times \text{width} \times 1/4$$

$$V = S \times \text{height} \times 1/3$$

HE Staining

Skin tissue samples were collected from the right hind limb of lymphedema mice, routinely embedded in paraffin, and prepared into 5μm thick tissue sections. After dewaxing with xylene and hydration with gradient alcohols, the sections were stained according to the instructions of the hematoxylin and eosin (H&E) staining kit (Servicebio). After staining, the sections were dehydrated with gradient alcohols, cleared with xylene, and mounted with neutral resin. At least 3 mice were randomly selected from each group, and 3–5 non-overlapping fields of view were selected from each mouse for observation. Finally, bright-field microscopy was used to observe and acquire images of the tissue morphological changes.

RNA Extraction and Quantitative PCR Analysis

Total ribonucleic acid (RNA) was extracted using the SteadyPure RNZ kit (AG, China), and RNA concentration and purity were assessed using a NanoDrop spectrophotometer. Reverse transcription was performed according to the manufacturer's instructions to synthesize complementary DNA (cDNA) (Servicebio). Quantitative polymerase chain reaction (qPCR) was carried out using SYBR Green chemistry (TransGen Biotech, China), with β - actin used as the internal control. Relative gene expression levels were calculated using the $2^{-}\Delta\Delta Ct$ method. Primer sequences are listed in Supplementary Table 1. All pre-defined molecular assays were fully reported. During the experiment, one mouse in the SL group died and was not included in subsequent molecular analysis; all other animals were included in the final data statistics.

Statistical Analysis

Statistical analysis was performed using GraphPad Prism software (version 9.5.1). Experimental data are expressed as mean ± standard deviation (mean ± SD). Normality was assessed using the Shapiro–Wilk test. For comparisons between two groups with data conforming to a normal distribution, an unpaired Student's *t*-test was used. A value of $P < 0.05$ was considered statistically significant. Hind limb volume data were analyzed using two-way analysis of variance (two-way ANOVA) with group (Sham, Model) as the between-subjects factor and time point (Days 0, 3, and 6) as the within-subjects factor. The main effects of group and time, as well as the group × time interaction effect, were reported. When a significant interaction was detected, Sidak's multiple comparisons test was further performed for post-hoc analysis to compare differences between the two groups at each time point. Results are presented as mean difference, 95% confidence interval (CI), and adjusted P value. Statistical significance was set at $P < 0.05$.

All six mice in the Sham group survived and were included in all analyses. The Model group initially consisted of six mice; however, one mouse died accidentally on postoperative day 6, an event unrelated to the lymphedema model itself. This mouse was therefore excluded from subsequent molecular analyses, including qPCR, but its available hind limb volume data (Days 0, 3, and 6) were included in the longitudinal analysis. Accordingly, the final sample sizes were: for volume analysis (including available data from the deceased mouse) Sham n = 6, Model n = 6; for molecular analysis Sham n=6, Model n = 5. All available animal samples were included in the final statistical analysis for each respective outcome measure.

Multiple testing correction was performed where applicable using adjusted P values (false discovery rate, FDR). Potential confounding factors such as batch effects were considered and controlled during the analysis.

scTenifoldKnk-Based in silico Virtual Knockout Analysis

To investigate the regulatory role of CXCL12, we performed in silico virtual knockout analysis using `scTenifoldKnk`. scRNA-seq samples were integrated using Seurat (v5.1) and Harmony for batch correction. The expression matrix was filtered to retain CXCL12 and the top 1000 highly variable genes. Virtual knockout was performed with the following parameters: `nc_nNet = 10`, `nc_nCells = 500`, and `nc_nComp = 3`. Genes with adjusted p-value < 0.05 were considered significantly differentially regulated. The top 20 genes ranked by fold change (FC) were visualized in a bar plot, and a volcano plot displayed Z-scores versus -log10(adjusted p-values), with significantly dysregulated genes (|Z-score| > 2, adjusted $p < 0.05$) highlighted.

Ethics Statement

This study complied with the Declaration of Helsinki. The human transcriptomic data analyzed in this study were obtained from publicly available databases. All original studies had received ethical approval from their respective institutional review boards (IRB) and informed consent from participants.

The present study only performed secondary analysis of publicly available and de-identified datasets and did not involve identifiable personal information or direct interaction with human participants. Therefore, additional ethical approval was not required. According to Article 32 of the Measures for Ethical Review of Life Science and Medical Research Involving Human Subjects (2023, China), studies based on publicly available and de-identified data are exempt from further ethical review.

All animal experiments were approved by the Animal Ethics Committee of Guangxi Medical University (Approval No. 202401033) and were conducted in accordance with the relevant guidelines for the care and use of laboratory animals.

Results

Quality Control and Dimensionality Reduction Analysis of scRNA-Seq Data

To ensure the reliability and accuracy of scRNA-seq data, this study first conducted a systematic quality control assessment of the raw data. Before filtering, the results showed that most cells had a moderate number of genes detected (nFeature_RNA) and transcript abundance (nCount_RNA), and the overall proportion of mitochondrial genes was low, indicating acceptable initial data quality (Supplementary Figure 1A). After applying the preset screening criteria (200–5000 genes, mitochondrial gene ratio <10%, erythrocyte-related gene ratio <5%), a total of 70,114 high-quality cells were obtained for subsequent analysis, with post-filtering quality metrics demonstrating improved data consistency (Supplementary Figure 1B).Further analysis revealed a significant positive correlation between nCount_RNA and nFeature_RNA (correlation coefficient r = 0.96), indicating a good consistency between sequencing depth and the number of genes detected, while the correlation with the proportion of mitochondrial genes was weak, further supporting the stability of data quality (Supplementary Figure 1C). Based on this, the data were standardized and highly variable genes (HVGs) were screened. Approximately 2000 HVGs were identified, exhibiting significant expression differences across different cell types, providing crucial information for subsequent dimensionality reduction analysis. PCA based on HVGs showed that cells were continuously distributed in the low-dimensional space, with no obvious separation trend observed between different samples, suggesting that batch effects had a relatively small impact on the overall data structure. Elbow plot analysis was then used to identify the first six principal components for subsequent analysis (Supplementary Figure 1D). To further improve data accuracy, potential doublets were identified and removed. Results showed that doublets were mainly distributed in some peripheral regions (Supplementary Figure 1E). After screening, 67,703 high-confidence single cells were obtained for subsequent analysis. A SNN graph was constructed based on the selected principal components, and Louvain's algorithm was used for cluster analysis. Ultimately, 13 cell clusters were identified at a resolution of 0.1 (Supplementary Figure 1F). UMAP dimensionality reduction results showed that each cell cluster was clearly distributed and structurally stable in the low-dimensional space.

Cell Type Annotation Reveals CD4$^+$ T Cell Heterogeneity

Cells were annotated based on canonical marker genes, and major cell populations, including T cells, Natural Killer (NK) cells, monocytes, fibroblasts, macrophages, adipocytes, neutrophils, dendritic cells, B cells, basophils, and epithelial cells were identified (Supplementary Figure 2A). Each cluster displayed distinct marker gene expression patterns, indicating robust and biologically meaningful cell type classification. UMAP-based visualization further confirmed the distribution of CD4$^+$ T cells expressing CD4 and IL7R, while CD40LG expression was enriched in some cells, suggesting the presence of an activated CD4$^+$ T cell population (Supplementary Figure 2B). These results highlight the heterogeneity of CD4$^+$ T cells and support their potential role in downstream functional remodeling analyses.

scRNA-Seq Reveals Expansion and Heterogeneity of CD4$^+$ T Cells in SL

To systematically analyze the cellular composition characteristics of SL tissues, this study performed single-cell transcriptome analysis on the control and SL groups. UMAP dimensionality reduction results showed that multiple cell types could be clearly distinguished in both groups of samples, including T cells, NK cells, monocytes, fibroblasts, macrophages, adipocytes, neutrophils, dendritic cells, B cells, basophils, and epithelial cells (Figure 1A). Overall, the distribution patterns of each cell population were basically consistent in the two groups, suggesting that cell type annotation has good stability and comparability. Further analysis of the proportions of each cell type (Figure 1B) revealed significant changes in the cell composition of the SL group. Among them, the proportion of T cells increased slightly

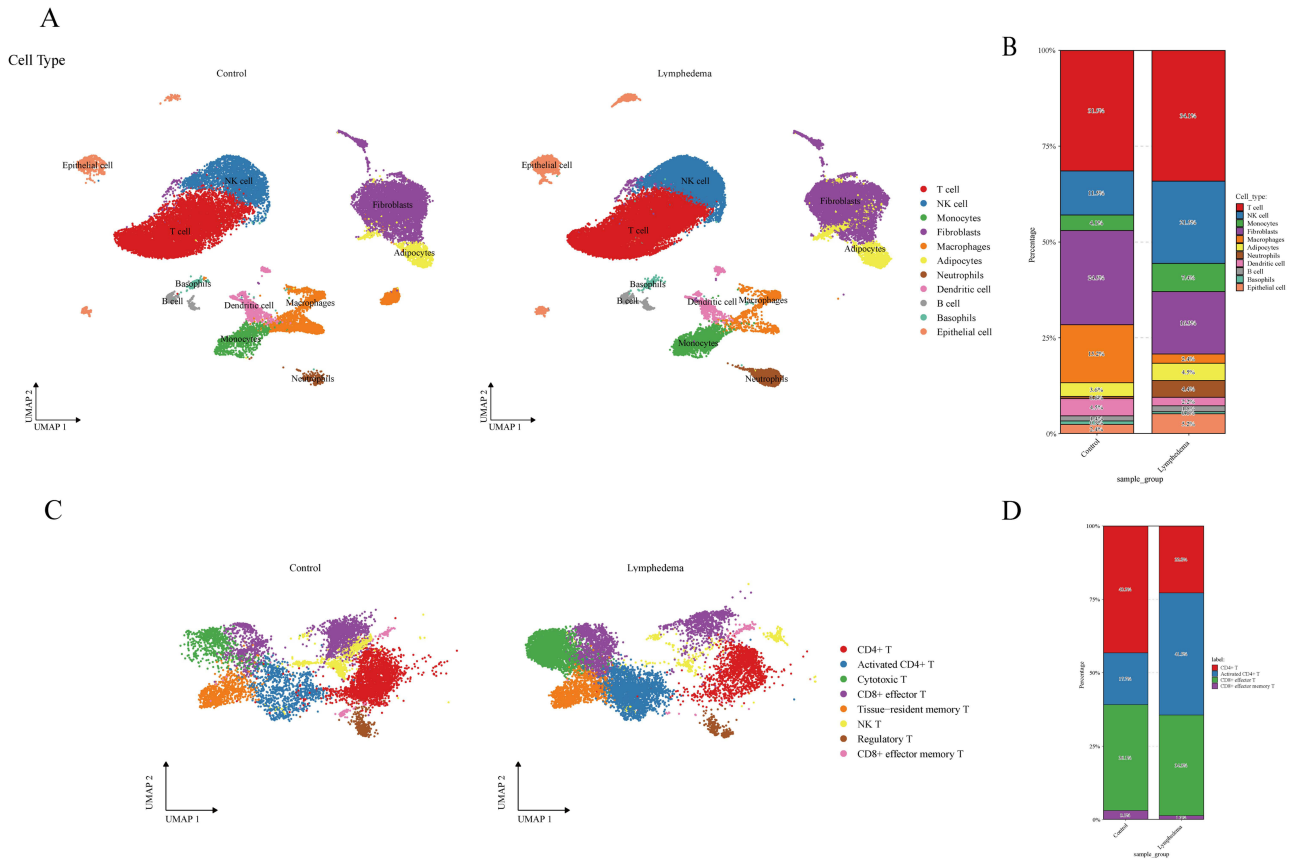

Figure 1 scRNA-seq reveals immune cell heterogeneity and the expansion of CD4$^+$ T cells in SL tissue. (**A**) UMAP visualization of all cell populations in control and SL groups. Cells are colored by annotated cell types, including T cells, NK cells, monocytes, fibroblasts, macrophages, adipocytes, neutrophils, dendritic cells, B cells, basophils, and epithelial cells. (**B**) Stacked bar plot showing the proportional distribution of major cell types in control and SL groups. Each bar represents the relative abundance of each cell type within the total cell population. (**C**) UMAP visualization of T cell subsets in control and SL groups. Cells are colored by annotated T cell subsets, including CD4$^+$ T cells, activated CD4$^+$ T cells, cytotoxic T cells, CD8$^+$ effector T cells, tissue-resident memory T cells, NK T cells, regulatory T cells, and CD8$^+$ effector memory T cells. (**D**) Stacked bar plot showing the proportional distribution of T cell subsets in control and SL groups. Each bar represents the relative abundance of each T cell subset within the total T cell population. All analyses were based on single-cell RNA sequencing data as described in the Methods section.

https://doi.org/10.2147/JIR.S597037

(from about 31.5% to 34.1%), NK cells increased significantly (from about 11.5% to 21.5%), and the proportion of monocytes also increased (from about 4.1% to 7.4%), suggesting enhanced infiltration of immune cells in SL tissue. In contrast, the proportions of fibroblasts and macrophages decreased in the SL group (fibroblasts decreased from approximately 24.3% to 16.3%, and macrophages decreased from approximately 15.2% to 2.4%), suggesting that the tissue matrix and some resident immune cell components underwent remodeling. In addition, the proportions of other cell types, such as adipocytes, neutrophils, dendritic cells, and B cells, also changed to varying degrees. SL tissue exhibited a microenvironmental remodeling characterized by an increased proportion of immune cells, suggesting that inflammatory cellular components may play an important role in the development and progression of the disease.

Based on the overall cellular composition analysis, this study conducted a re-clustering analysis of the T cell population to further elucidate the heterogeneity of T cells. UMAP results showed that T cells could be further divided into several subsets, including $CD4^+$ T cells, activated $CD4^+$ T cells, cytotoxic T cells, $CD8^+$ effector T cells, tissue-resident memory T cells, NK T cells, regulatory T cells, and $CD8^+$ effector memory T cells (Figure 1C). Each subset could be clearly distinguished in both the Control and Lymphedema groups, suggesting that the T cell subset classification has good stability. Further analysis of T cell subset proportions (Figure 1D) revealed significant changes in the T cell subset composition in the lymphedema group. Specifically, the proportion of activated $CD4^+$ T cells increased significantly, from 17.7% in the control group to 41.9%, while the proportion of resting $CD4^+$ T cells decreased from 43.9% to 22.8%. Consequently, the overall proportion of $CD4^+$ T cells (resting and activated combined) increased from 61.6% in controls to 64.7% in lymphedema samples. Simultaneously, the proportion of cytotoxic T cells slightly decreased (from 36.1% to 34.3%), while the proportion of $CD8^+$ effector T cells decreased slightly (from 3.1% to 1.3%). It is noteworthy that although the proportion of resting $CD4^+$ T cells decreased, the proportion of activated $CD4^+$ T cells increased significantly, indicating that $CD4^+$ T cells as a whole showed a trend of transitioning from a resting to an activated state, suggesting that they underwent functional reprogramming in the lymphedema-associated inflammatory microenvironment.

$CD4^+$ T Cells Function as Key Signal-Receiving Cells in the Lymphedema Immune Microenvironment

To further evaluate the potential role of multicellular synergistic regulation in the tissue microenvironment of SL, we systematically analyzed the intercellular communication relationships among different cell types based on scRNA-seq data. CellChat results showed that intercellular communication signals are mainly composed of multiple types, including secretory signals, extracellular matrix (ECM)-receptor interactions, and cell-cell contacts (Figure 2A). Among them, secretory signals accounted for the highest proportion (61.8%), followed by ECM–Receptor interaction (21.7%), while Cell–Cell Contact signals accounted for 16.5%. Analysis of the molecular composition of ligand-receptor interactions showed that ligand-receptor pairs in monomeric or non-classical complex form accounted for 52.1%, while ligand-receptor complexes in heterodimer form mediated signal transduction accounted for 47.9%. Furthermore, a statistical analysis of the signal pathway annotation sources used by CellChat revealed that most communication signals originated from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (approximately 73%); in addition, approximately 27% of the signals were sourced from literature-curated data, supplementing novel or specific regulatory axes not yet fully covered by the database. These results suggest that secreted signaling molecules represent the dominant mode of intercellular communication in the SL microenvironment. Such signaling patterns may facilitate extensive cross-talk between stromal and immune cells and contribute to the recruitment and activation of immune cell populations, including $CD4^+$ T cells, thereby promoting immune remodeling during lymphedema progression.

The results show that adipocytes, fibroblasts, and T cells (particularly $CD4^+$ T cells) exhibit the strongest interaction intensity among all cell populations. This is reflected in two indicators: the total number of interactions (Figure 2B) and interaction weights (Figure 2C). This finding suggests a dense interaction between $CD4^+$ T cells and stromal cells (particularly adipocytes and fibroblasts), highlighting the importance of immune-stromal interactions in shaping the lymphedema microenvironment. The high interaction centrality of $CD4^+$ T cells positions them as key receivers and integrators of signals from the stromal microenvironment, potentially influencing their recruitment, retention, and

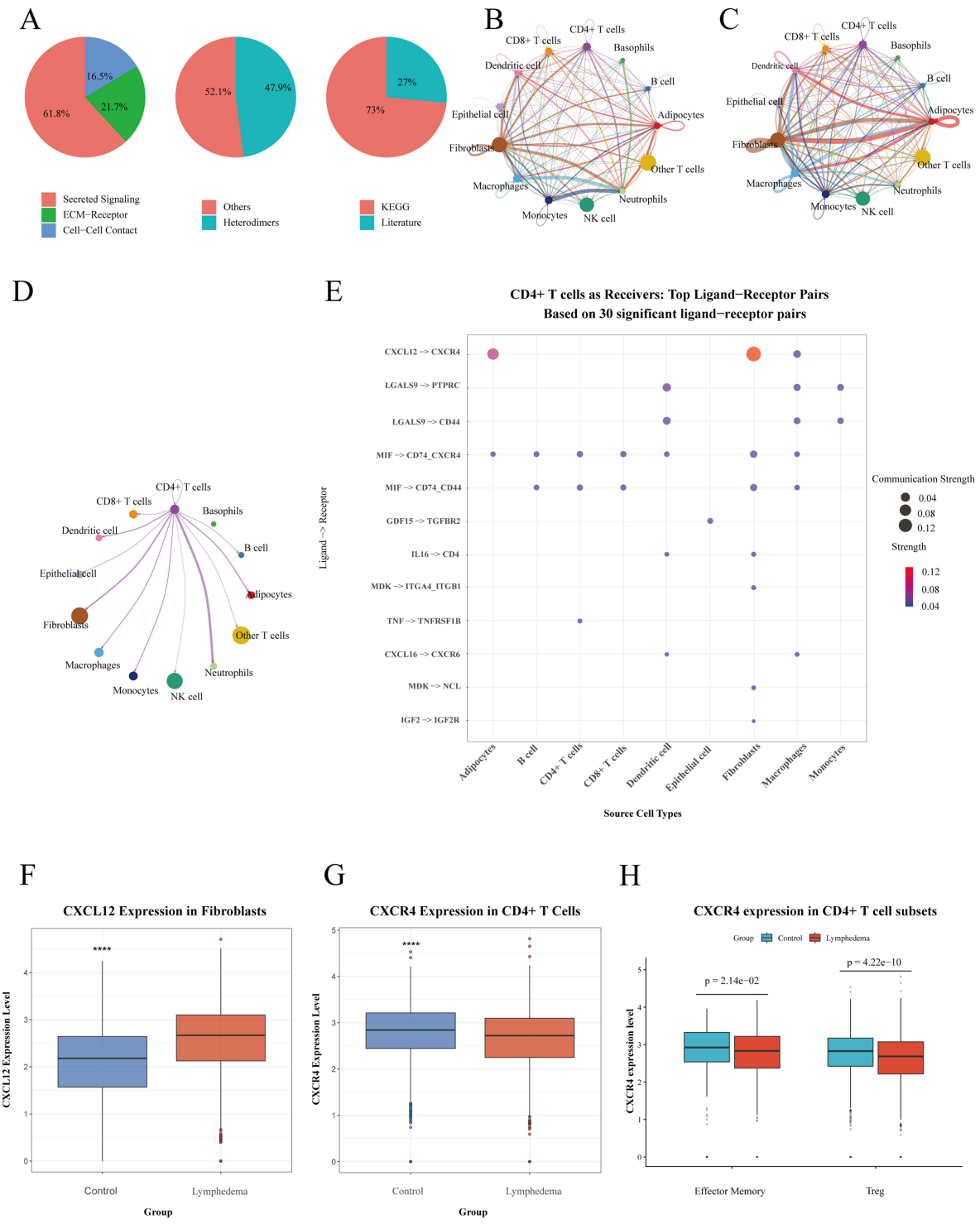

Figure 2 Intercellular communication network and CXCL12-CXCR4 axis features of CD4$^+$ T cells in SL. (**A**) Pie chart showing the proportion of intercellular signaling modes, including secreted signaling, ECM-receptor interaction, and cell-cell contact, based on CellChat analysis. (**B** and **C**) Global intercellular communication networks. (**B**) Interaction counts and (**C**) Interaction strength. Node size represents cell population size; edge thickness indicates interaction frequency (**B**) or interaction strength (**C**). (**D**) CD4$^+$ T cell-centered communication network, illustrating interactions between CD4$^+$ T cells and other immune (macrophages, dendritic cells, NK cells) and stromal (fibroblasts) cell populations. (**E**) Dot plot showing the top ligand-receptor pairs with CD4$^+$ T cells as receiver cells, based on 30 significant interactions. Dot size and color indicate communication strength. Notably, the CXCL12-CXCR4 axis exhibits strong signaling across multiple source cell types. (**F**) CXCL12 expression levels in fibroblasts from control and lymphedema samples. (**G**) CXCR4 expression in CD4$^+$ T cells was slightly decreased in the Lymphedema group compared with the Control group. (**H**) CXCR4 expression in CD4$^+$ T cell subsets (effector memory T cells and Treg cells). CXCR4 expression was slightly decreased in effector memory T cells and markedly decreased in Treg cells in the Lymphedema group compared with the Control group. Data are presented as ± SD. Statistical significance was determined using the Wilcoxon rank-sum test. ****P < 0.0001.

https://doi.org/10.2147/JIR.S597037

functional polarization in lymphedema tissues. To further identify key responding cells in the cell communication network, we analyzed the main signal receiving cells (Figure 2D). The results showed that $CD4^+$ T cells can receive communication signals from a variety of immune and non-immune cells, including dendritic cells, macrophages, monocytes, neutrophils, as well as B cells, adipocytes, and fibroblasts. These results highlight the central role of $CD4^+$ T cells in coordinating multicellular communication within the lymphedema microenvironment. To further elucidate the key ligand-receptor interactions targeting $CD4^+$ T cells, this study systematically analyzed signal inputs from different cell sources (Figure 2E). The results showed that various stromal cells and immune cells could transmit signals to $CD4^+$ T cells. Among them, the interaction between fibroblasts and $CD4^+$ T cells were the most significant and extensive, suggesting that they play an important role in regulating the functional state of $CD4^+$ T cells. In the specific analysis of ligand-receptor pairs, multiple extracellular matrix-related and inflammation-related signaling pathways were significantly enriched, reflecting the active communication state between stromal cells and immune cells. It is noteworthy that the CXCL12–CXCR4 signaling pathway was significantly enriched in the communication between fibroblasts and $CD4^+$ T cells. To further validate this finding, we subsequently analyzed the expression patterns of CXCL12 and its receptor, CXCR4, in the relevant cell populations. In fibroblasts, the expression of CXCL12 was significantly elevated in the lymphedema group compared to the control group (Figure 2F); this difference was statistically significant (****$P < 0.0001$). In $CD4^+$ T cells, CXCR4 expression showed a downward trend in the lymphedema group compared to the control group (Figure 2G), and this difference was statistically significant (****$P < 0.0001$). Further analysis of $CD4^+$ T-cell subsets (Figure 2H) revealed that, within both effector memory T cells and regulatory T cells (Tregs), CXCR4 expression was lower in the lymphedema group compared to the control group, with the differences reaching statistical significance (Effector Memory: $P = 2.14 \times 10^{-2}$; Treg: $P = 4.22 \times 10^{-10}$).

$CD4^+$ T Cell Co-Expression Network Reveals Functional Heterogeneity and Remodeling-Associated Transcriptional Programs

To investigate the transcriptional program underlying $CD4^+$ T cell remodeling in lymphedema, we constructed a co-expression network using hdWGCNA. A soft-threshold of 6 was selected based on scale-free topological fitting ($R^2 > 0.8$) to ensure robust network structure (Figure 3A). Using this parameter, we identified several co-expression modules with high internal connectivity (Figure 3B), revealing a coordinated transcriptional program within $CD4^+$ T cells of the SL tissue. Different modules exhibited distinct distribution patterns in the low-dimensional embedding space (see Figure 3C), demonstrating significant transcriptional heterogeneity among $CD4^+$ T cell populations. This heterogeneity likely reflects the functional diversity of $CD4^+$ T cell responses to the inflammatory microenvironment. The expression of module-specific genes showed a continuous and stable gradient across individual cells (Figure 3D), further supporting the presence of structured transcriptional states within $CD4^+$ T cells. Most modules were highly enriched in $CD4^+$ T cells (Figure 3E), confirming their $CD4^+$ T cell specificity. Notably, several modules also showed partial expression in other immune (NK cells, monocytes) and stromal (fibroblasts, adipocytes) populations, suggesting potential transcriptional crosstalk between $CD4^+$ T cells and the broader tissue microenvironment. Further analysis revealed that the CD4T-M1 module not only exhibited high expression in T cells but also showed prominent expression in fibroblasts (Figure 3F). The M1 module may reflect not only $CD4^+$ T cell transcriptional features but also functional interactions between T cells and stromal cells, potentially contributing to immune-stromal crosstalk in lymphedema. Inter-module correlation analysis revealed distinct correlation patterns among $CD4^+$ T cell co-expression modules (Figure 3G). To further characterize the transcriptional architecture of $CD4^+$ T cell modules, we constructed gene co-expression networks for CD4T-M1 to CD4T-M6 (Figure 3H–M). Specifically, the CD4T-M1 module was enriched in immune regulation and chemotaxis-related genes, including CXCL12, CXCL14, and CD63, and exhibited a densely connected network structure, suggesting coordinated regulation of cell migration and microenvironmental interactions. The CD4T-M2 module was dominated by interferon-stimulated genes (eg, ISG15, IFI6, and RSAD2), reflecting a classical inflammatory or antiviral response program. In contrast, both the CD4T-M3 and CD4T-M5 modules were enriched in ribosomal protein genes (RPS/RPL families). The CD4T-M4 module was characterized by genes involved in T cell signaling and transcriptional regulation, such as ETS1, ITK, and FYN, supporting its association with T cell activation and intracellular signaling. Meanwhile, the

CD4T-M6 module contained genes related to inflammatory signaling and T cell differentiation, including STAT3, IL7R, and CCR7. Notably, CXCL12 showed consistent co-expression within the M1 module and exhibited moderate connectivity (kME = 0.25), while also demonstrating comparable association with the M4 module (kME = 0.24). Given that the M4 module is enriched in genes associated with T cell activation and cytoskeletal dynamics (eg, ITK, SKAP1, ETS1, and IQGAP1), these findings suggest that CXCL12 may function as a cross-module regulatory node linking microenvironmental signaling with T cell functional remodeling.

Cross-Dataset Integration Analysis and Batch Effect Correction

To integrate multiple independent transcriptome datasets for subsequent analysis, this study first performed PCA on samples from GSE132936, GSE255848, and GSE302563 to assess potential batch effects. The results showed that before

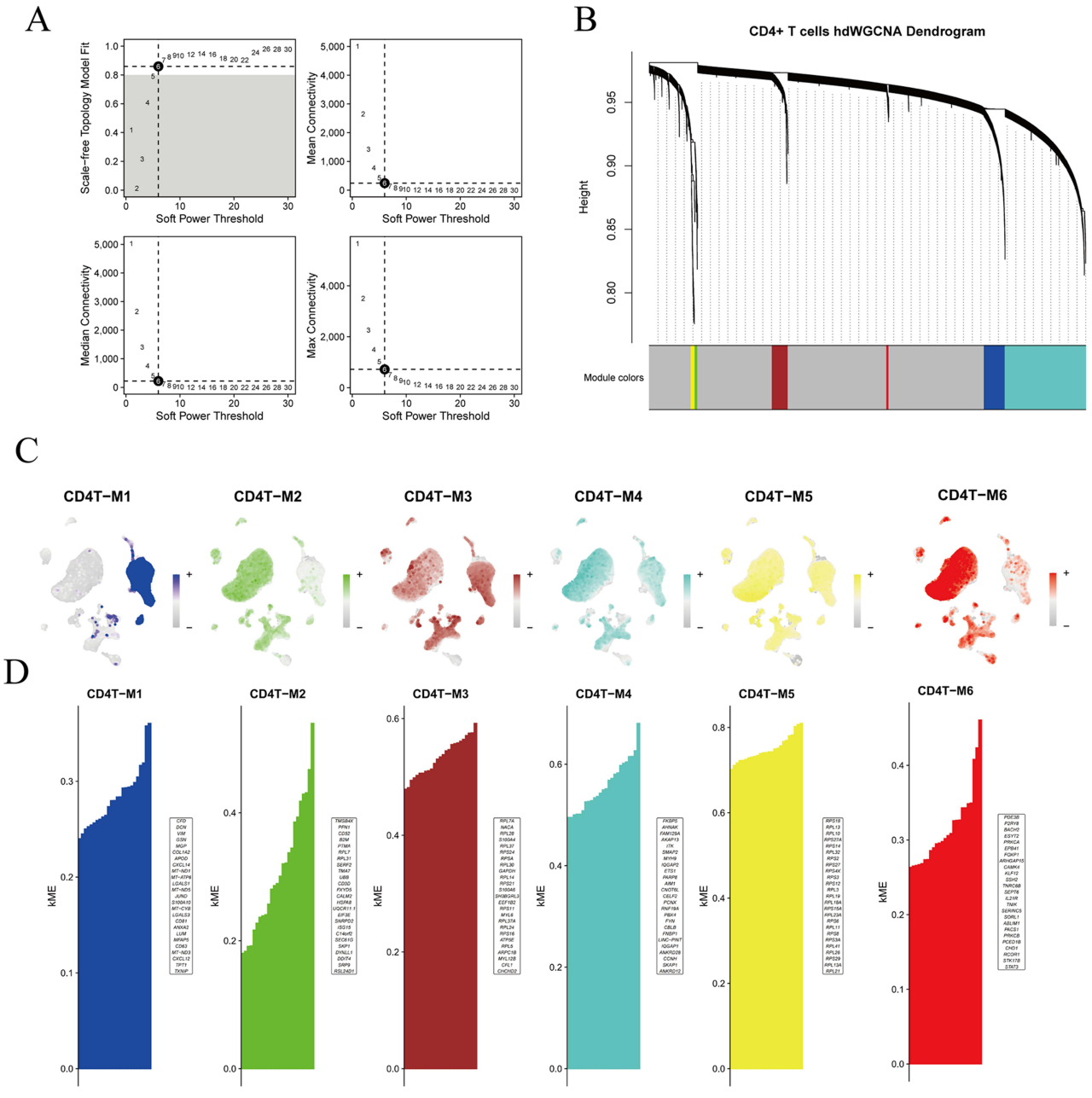

Figure 3 continued.

https://doi.org/10.2147/JIR.S597037

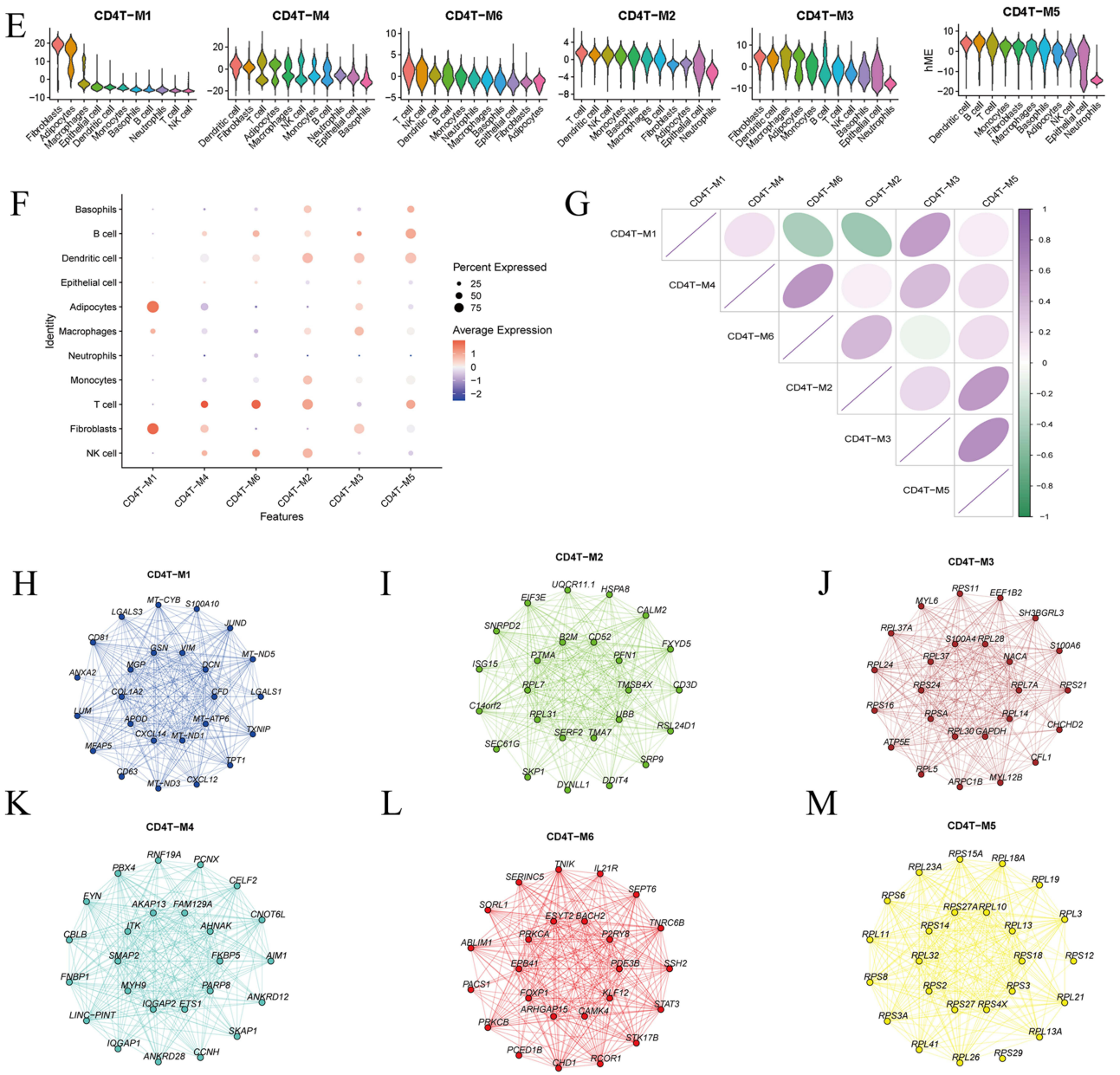

Figure 3 hdWGCNA reveals distinct co-expression modules reflecting CD4$^+$ T cell functional heterogeneity in SL. (**A**) Selection of the soft-thresholding power for network construction. Based on the evaluation of scale-free topology fit index and network connectivity across different powers, a soft-thresholding power of 6 was selected to ensure scale-free network properties. (**B**) Hierarchical clustering of genes based on the topological overlap matrix (TOM) derived from CD4$^+$ T cell transcriptomic data. Different colors represent distinct gene co-expression modules within CD4$^+$ T cells. (**C**) UMAP visualization of module eigengene expression patterns (CD4T-M1 to CD4T-M6). (**D**) Module membership (kME) ranking of genes within each CD4$^+$ T cell co-expression module (CD4T-M1 to CD4T-M6). Genes are ordered based on their correlation with the corresponding module eigengene, and the top hub genes for each module are highlighted. (**E**) Distribution of module eigengene (ME) expression across different cell types, derived from CD4$^+$ T cell co-expression modules. (**F**) Dot plot showing the expression patterns of CD4$^+$ T cell-derived co-expression modules (CD4T-M1 to CD4T-M6) across different cell types. Dot size represents the percentage of expressing cells, and color intensity indicates the average expression level. (**G**) Correlation analysis among CD4$^+$ T cell co-expression modules. Module analysis was performed based on single-cell transcriptomic data as described in the Methods section. (**H**–**M**) Gene co-expression network visualization of each CD4$^+$ T cell module (CD4T-M1 to CD4T-M6). Nodes represent genes, and edges indicate co-expression relationships derived from hdWGCNA.

batch correction, samples from different datasets were clearly separated in the low-dimensional space. The samples were mainly clustered according to the source of the dataset rather than biological characteristics, indicating that there was a significant batch effect in the original data (Supplementary Figure 3A).

Subsequently, the expression matrix was corrected using a batch effect correction method. After correction, PCA analysis was performed again, revealing significant mixing of samples from different datasets in the low-dimensional space. The original separation phenomenon driven by the dataset source was significantly reduced (Supplementary Figure 3B), indicating that the batch effect was effectively corrected.

CXCL12-Centered Identification and Validation of Key Candidate Genes

Differential expression analysis showed that CXCL12 was significantly upregulated in the SL group compared with the Control group (Figure 4A, $P < 0.001$). Meanwhile, genes such as FNBP1 and DCN were also upregulated, whereas CSN was relatively downregulated in the SL group, indicating substantial transcriptional alterations associated with SL. Hierarchical clustering analysis further demonstrated that these candidate genes exhibited stable expression patterns with strong discriminative ability across samples (Figure 4B), effectively distinguishing the SL group from the Control group. ROC curve analysis revealed that CXCL12 showed strong discriminative performance (AUC = 0.844, 95% CI: 0.706–0.950), while FNBP1 also exhibited good predictive value (AUC = 0.831, 95% CI: 0.683–0.956) (Figure 4C and D). Logistic regression was selected for subsequent analysis due to its robust performance and lower risk of overfitting, along with better interpretability compared with more complex models. (Supplementary Figure 4) Furthermore, feature importance analysis based on the `LogitBoost` model identified CXCL12 as the most important contributor among all candidate genes (Figure 4E), further supporting its central role in distinguishing SL from Control samples. To further characterize the genomic distribution of these key genes, a Circos plot was constructed (Figure 4F). The selected genes were distributed across multiple chromosomes without apparent genomic clustering, suggesting that they may function through coordinated regulatory networks rather than physical co-localization.

Immune Microenvironment Remodeling Characterized by $CD4^+$ T Cells

To systematically evaluate the immune microenvironment characteristics of tissues with SL, we analyzed the composition and interrelationships of immune cells based on integrated transcriptome data. CIBERSORT results showed that there were significant differences in the overall composition of immune cells between the control group and the SL group (Figure 4G). Correlation analysis further revealed that there was a significant negative correlation between multiple immune cell pairs, mainly involving cell subpopulations in different activation states or differentiation stages within the same immune lineage. As shown in Figure 4H, resting NK cells and activated NK cells showed a significant negative correlation ($r = -0.80$); similarly, memory B cells and naive B cells ($r = -0.72$), follicular helper T cells and resting $CD4^+$ memory T cells ($r = -0.76$), and macrophages M0 and M1 ($r = -0.67$) all showed significant negative correlations. Furthermore, there was a significant negative correlation between resting mast cells and activated mast cells ($r = -0.72$). In addition, significant positive correlations were observed among multiple immune cell subsets. Among them, naive B cells were positively correlated with monocytes ($r = 0.60$), macrophage M1 was significantly positively correlated with plasma cells ($r = 0.67$) and resting mast cells ($r = 0.75$), and monocytes also showed a strong positive correlation with $CD8^+$ T cells ($r = 0.79$). Differential immune cell analysis showed a significant difference in the number of resting NK cells between the control group and the SL group ($P < 0.05$) (Figure 4I).

CXCL12-Centered Correlation Analysis of Candidate Genes and Immune Cell Infiltration

We next analyzed the correlations between key candidate genes (CXCL12, CBLB, FKBP5, PCED1B, and GSN) and immune cell infiltration characteristics. As shown in Figure 5A, these genes exhibited varying degrees of association with multiple immune cell subsets, suggesting their involvement in immune microenvironment regulation. Among them, CXCL12 showed relatively high connectivity within the network and was correlated with multiple immune cell populations. Notably, its associations with $CD4^+$ T cell–related subsets were more prominent, indicating a potential role in $CD4^+$ T cell–mediated immune regulation. In addition, genes such as CBLB, FKBP5, PCED1B, and GSN also showed associations with specific immune cell subsets, supporting their potential cooperative roles in immune regulation. Further correlation analyses (Figure 5B–D) revealed distinct associations between CXCL12 expression and different

https://doi.org/10.2147/JIR.S597037

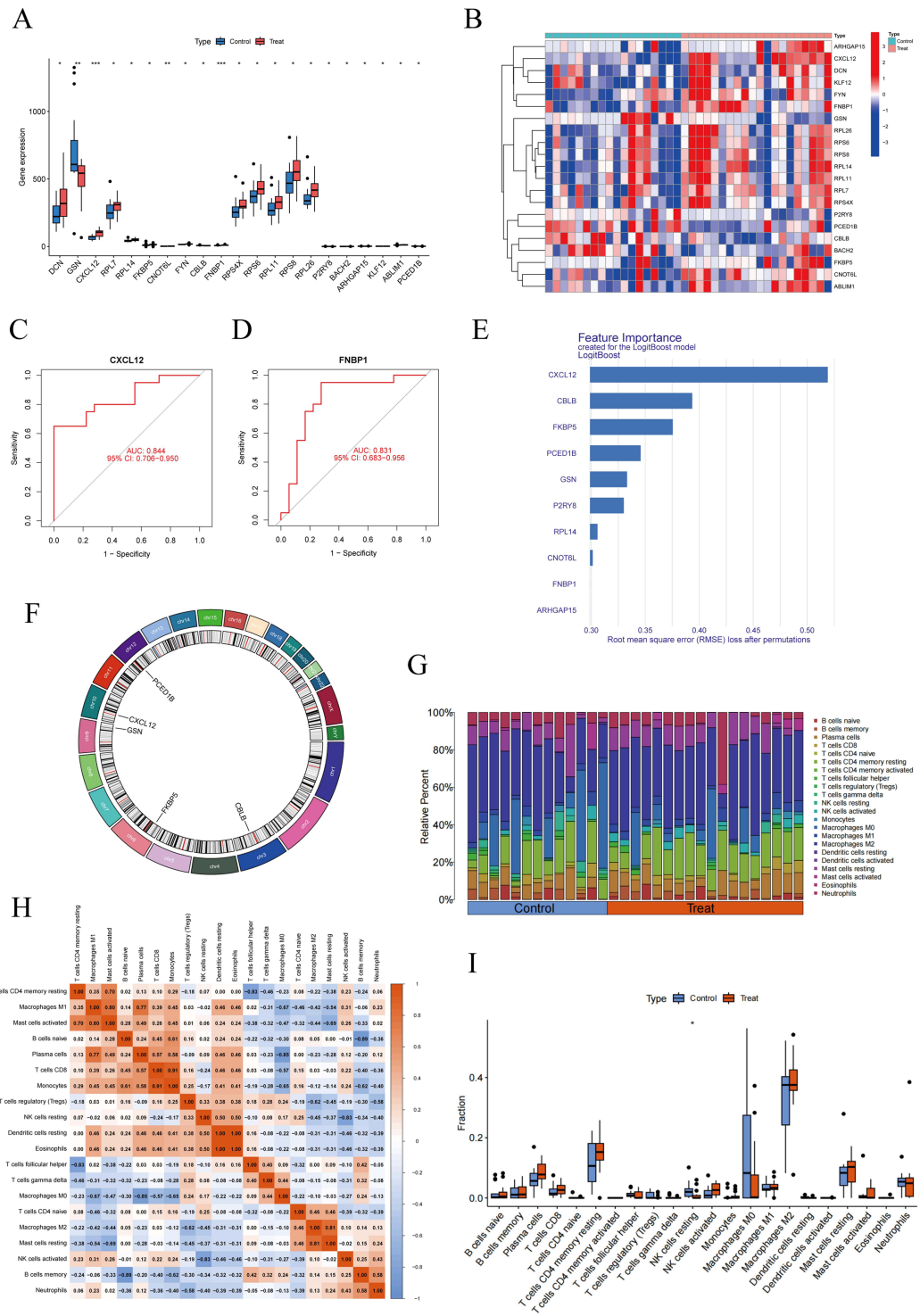

Figure 4 CXCL12-centered immune network and pathway enrichment analysis in lymphedema. (**A**) Boxplot showing the expression levels of selected hub genes between control and lymphedema groups. CXCL12 and its associated genes exhibited significant differential expression between the two groups. (**B**) Heatmap of gene expression patterns across samples, with hierarchical clustering revealing distinct expression profiles between control and lymphedema groups. CXCL12 is highlighted, demonstrating its differential expression and potential role as a candidate hub gene. (**C**) ROC curve analysis of CXCL12 (AUC = 0.844, 95% CI: 0.706–0.950). (**D**) ROC curve analysis of FNBP1 (AUC = 0.831, 95% CI: 0.683–0.956). AUC values were calculated using the Wilcoxon rank-sum test with 95% confidence intervals estimated by bootstrap resampling. (**E**) Feature importance analysis based on the LogitBoost model. CXCL12 ranked as the most important feature among candidate genes, indicating its dominant contribution to classification performance. (**F**) Circos plot showing the chromosomal distribution of CXCL12 and its associated hub genes. (**G**) Relative abundance of immune cell populations in control and lymphedema groups. (**H**) Correlation heatmap of immune cell populations. (**I**) Comparison of immune cell fractions between control and lymphedema groups. Data are presented as mean ± SD. Statistical significance was determined using the Wilcoxon rank-sum test (**G** and **I**). *P < 0.05, **P < 0.01, ***P < 0.001. Statistical analyses were performed as described in the Methods section.

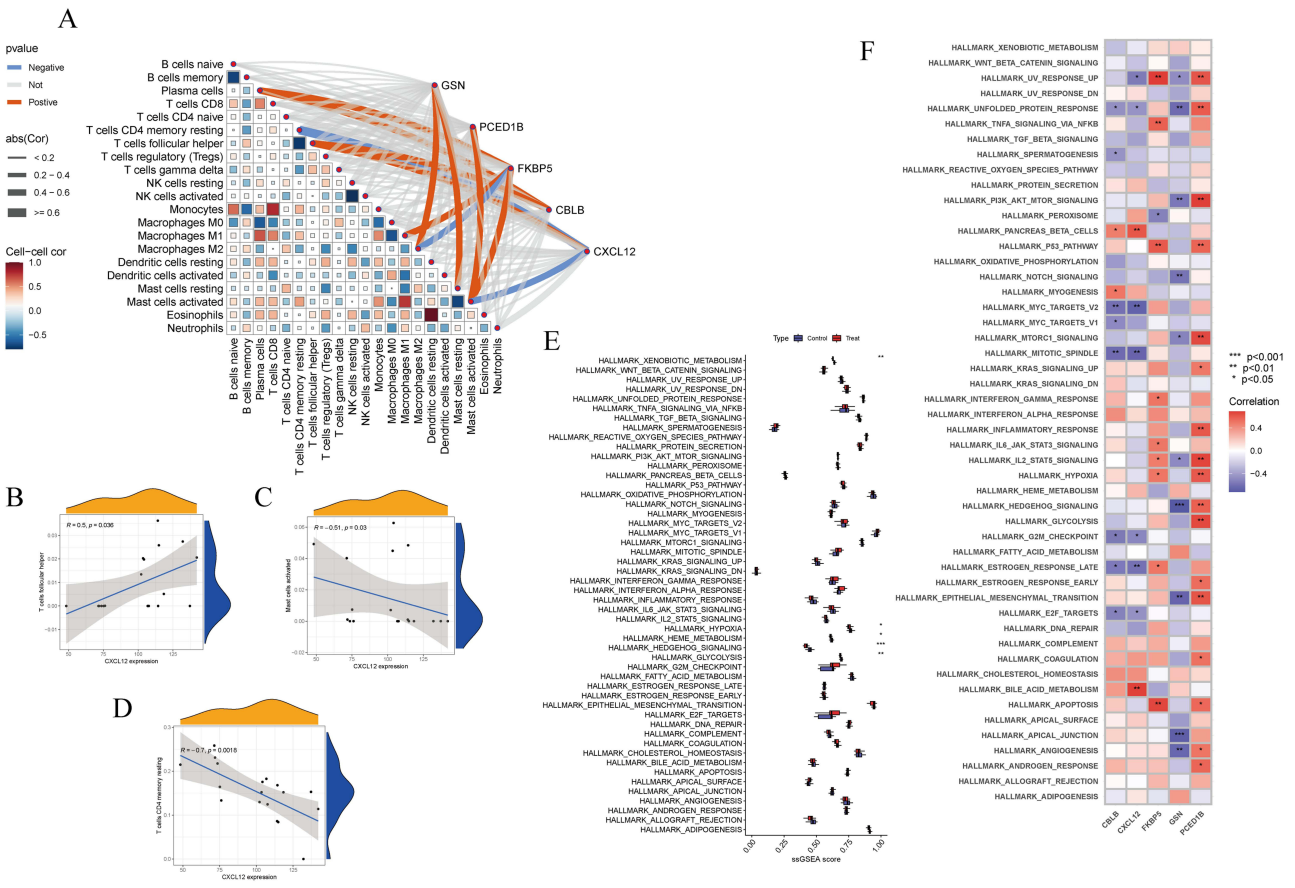

Figure 5 CXCL12-centered immune interaction network and its association with CD4$^+$ T cell subsets. (**A**) Correlation network of immune cell populations and CXCL12-associated genes. The heatmap shows pairwise correlations among immune cell types, while the network highlights CXCL12-related interactions with key genes. Edge thickness represents correlation strength; edge color indicates correlation direction (red: positive, blue: negative). (**B**–**D**) Correlation analysis between CXCL12 expression and immune cell subsets. CXCL12 expression was positively correlated with Tfh (**B**; R = 0.5, P = 0.036), but negatively correlated with mast cells (**C**; R = −0.51, P = 0.03) and resting memory CD4$^+$ T cells (**D**; R = −0.7, P = 0.0018). Correlation analysis was performed using Spearman's rank correlation test. (**E**) Gene set variation analysis (GSVA) of Hallmark pathways between control and lymphedema groups. Differences in pathway activity scores are shown, with significant alterations observed in multiple immune- and inflammation-related pathways. (**F**) Correlation heatmap showing the associations between CXCL12 and key signaling pathways. Color intensity represents correlation strength; statistical significance is indicated. Statistical significance was assessed using the Wilcoxon rank-sum test (**E**) and Spearman's rank correlation test (**F**). *P < 0.05, **P < 0.01, ***P < 0.001. Statistical analyses were performed as described in the Methods section.

immune cell subsets. Specifically, CXCL12 expression was positively correlated with follicular helper T (Tfh) cells (R = 0.50, P = 0.036), whereas it showed negative correlations with activated mast cells (R = −0.51, P = 0.03) and resting memory CD4$^+$ T cells (R = −0.70, P = 0.0018).

Hallmark Pathway Activity Changes and Pathway Associations of Key Genes

Hallmark pathway ssGSEA analysis based on integrated transcriptomic data revealed significant differences in pathway activities between the control and secondary lymphedema groups (Figure 5E). The altered pathways mainly included metabolism-related pathways (such as xenobiotic metabolism and heme metabolism), signaling and developmental pathways (such as Hedgehog signaling and epithelial–mesenchymal transition, EMT), as well as immune- and inflammation-related pathways (such as IL2/STAT5 signaling and hypoxia), indicating coordinated alterations across multiple biological processes in secondary lymphedema.

To further elucidate the potential biological functions of key candidate genes at the pathway level in SL, we performed Hallmark gene set–based pathway correlation analyses for CXCL12, CBLB, FKBP5, PCED1B, and GSN (Figure 5F). Pathway correlation analysis revealed that CXCL12 exhibited distinct association patterns with multiple signaling pathways. CXCL12 was positively correlated with metabolic pathways, including pancreas beta cells and bile acid metabolism, suggesting a potential link between CXCL12-mediated immune regulation and metabolic processes. In

https://doi.org/10.2147/JIR.S597037

contrast, CXCL12 was negatively correlated with cell cycle– and proliferation-related pathways, including E2F targets, G2M checkpoint, mitotic spindle, and MYC targets V2, as well as stress-related pathways, such as estrogen response (late), unfolded protein response, and UV response (up).

Establishment and Phenotypic Validation of a Mouse Model of SL

Histological Changes in Lymphedema Tissue

To further evaluate the pathological characteristics of the lymphedema tissue, HE staining analysis was performed on the skin tissue of the right lower extremity. The results showed that the skin tissue structure in the model group was significantly altered. Under low magnification, the dermal layer structure is loose, the collagen fibers are arranged in a disordered manner, and the interstitial spaces are significantly widened, indicating the formation of tissue edema; at the same time, inflammatory cells are seen to accumulate in the dermal and subcutaneous junction areas. High-magnification analysis further revealed inflammatory cell infiltration, primarily composed of mononuclear cells, in the dermal stroma, with localized patches of these cells accompanied by destruction of collagen fiber structure. In contrast, the Sham group had intact skin tissue structure, dense and orderly arrangement of collagen fibers in the dermis, small interstitial spaces, and no obvious inflammatory cell infiltration (Figure 6A; scale bars = 100μm).

Successful Establishment of a Mouse Model of SL and Changes in Edema

To assess the establishment of the SL model, dynamic monitoring was conducted on the changes in hind limb volume in mice at different time points. The results showed that at baseline (day 0), there was no statistically significant difference

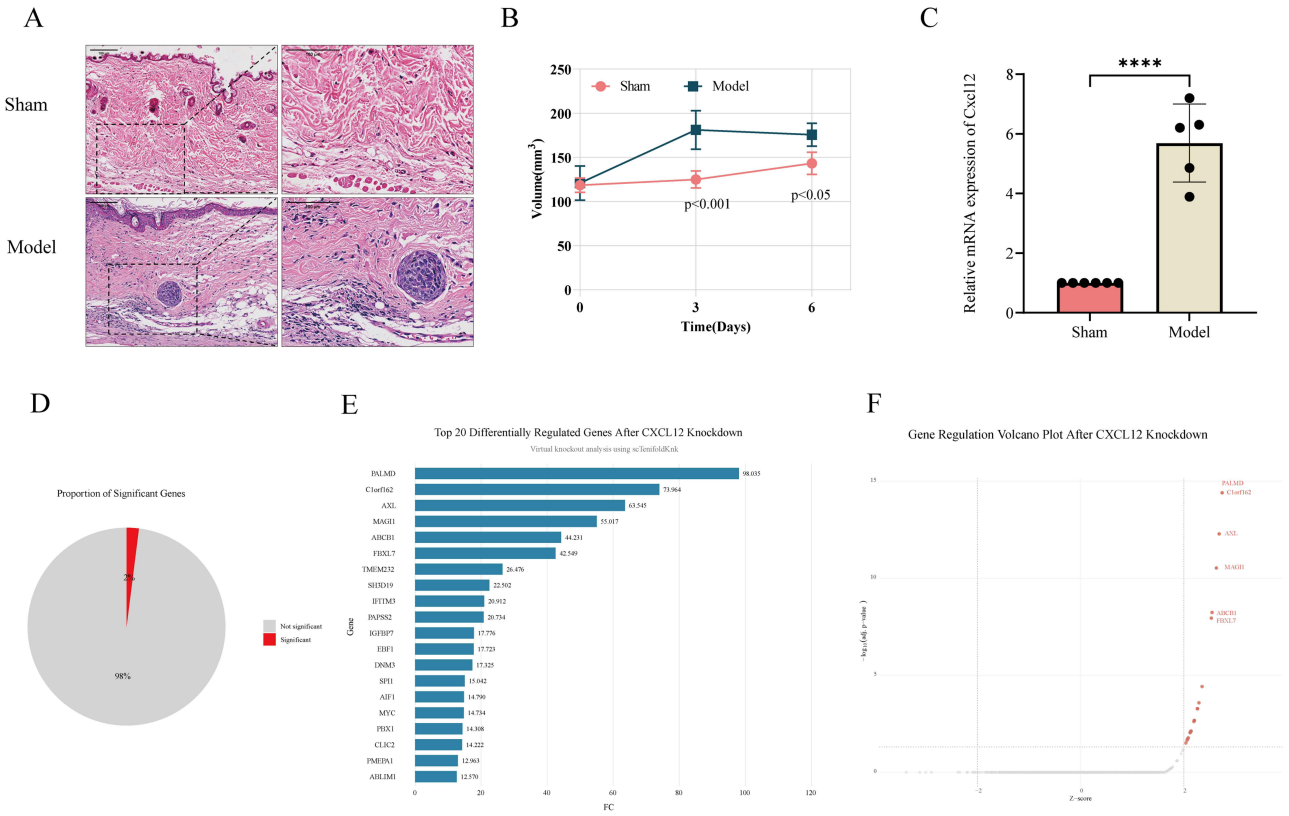

Figure 6 In vivo validation of Cxcl12 expression in a mouse model and transcriptomic consequences of in silico CXCL12 knockout. (**A**) H&E staining of right hindlimb skin tissues from mice in the sham and model groups. Scale bars = 100μm. (**B**) Changes in limb volume over time in Sham and Model groups. Data are presented as mean ± SD (n = 6). Statistical analysis was performed using two-way repeated-measures ANOVA followed by Sidak's multiple comparisons test. (**C**) qPCR validation of Cxcl12 expression in mouse skin tissues; Relative mRNA expression of Cxcl12 was measured by quantitative real-time PCR and normalized to the reference gene; Data are presented as mean ± SD (Sham, n = 6; Model, n = 5). Statistical significance was determined using an unpaired Student's *t*-test. ****P < 0.0001 vs. Sham group. (**D**) The proportion of significantly differentially expressed genes in the whole transcriptome after CXCL12 virtual knockout. (**E**) The top 20 differentially expressed genes with the most significant changes after CXCL12 virtual knockout (sorted by the magnitude of expression change); (**F**) Scatter plot based on Z-score and multiple correction significance analysis.

https://doi.org/10.2147/JIR.S597037

in limb volume between the Sham group and the Model group ($P > 0.05$).As time progressed, the limb volume of mice in the Model group increased significantly, and was significantly higher than that in the Sham group on days 3 and 6 (day 3: $P < 0.001$; day 6: $P < 0.05$), indicating that the model was successfully constructed and showed progressive edema changes.(Figure 6B) Two-way repeated-measures ANOVA was further used to analyze the time and grouping effects. The results showed that the time factor [$F(2,30) = 24.68$, $P < 0.0001$], the grouping factor [$F(1,30) = 37.38$, $P < 0.0001$], and the interaction between the two [$F(2,30) = 9.859$, $P = 0.0005$] were all statistically significant, indicating that there were significant differences in the trend of limb volume change between the two groups at different time points. Sidak's multiple comparison analysis showed that, compared with the Sham group, the Model group had significantly higher volumes on day 3 (mean difference = −56.24, 95% CI: −77.99 ~ −34.49, $P < 0.0001$) and day 6 (mean difference = −32.49, 95% CI: −54.24 ~ −10.74, $P = 0.0021$), while there was no statistically significant difference between the two groups on day 0 ($P = 0.9903$).

qPCR Validation Confirms Cxcl12 Upregulation in Lymphedema Tissues

To validate the transcriptome analysis results, we further used qPCR to detect changes in the expression of key genes in lymphedema tissue. As shown in Figure 6C, compared with the Sham group (n = 6), the mRNA expression level of Cxcl12 in the model group (n = 5) was significantly upregulated. One mouse in the model group died accidentally on postoperative day 6 and was excluded from molecular analysis.

Transcriptomic Alterations Following Virtual CXCL12 Knockout

To assess the overall impact of CXCL12 on the transcriptional regulatory network, we performed an in silico knockout using `scTenifoldKnk` and compared genome-wide gene expression profiles before and after perturbation (Figure 6D). The results revealed that only approximately 2% of genes showed significant expression changes following CXCL12 depletion, while the remaining 98% did not reach statistical significance. Despite the limited number of differentially expressed genes, their functional distribution was broad, suggesting a specific regulatory role for CXCL12.

To identify potential downstream molecules affected by CXCL12 deficiency, we conducted a ranking analysis of significantly altered genes. As shown in Figure 6E, CXCL12 depletion led to widespread transcriptional changes, with the top differentially regulated genes including PALMD, C1orf162, AXL, MAGI1, and ABCB1. Notably, several genes associated with cell signaling, transcriptional regulation, and immune responses, such as AXL, IFITM3, IGFBP7, and MYC were significantly altered. A volcano plot was generated to further characterize the transcriptional landscape following CXCL12 knockout (Figure 6F). The results showed that CXCL12 depletion resulted in the significant upregulation of multiple genes, with the most prominently affected being PALMD, C1orf162, AXL, MAGI1, ABCB1, and FBXL7. These genes exhibited both high fold changes and strong statistical significance, indicating robust transcriptional perturbations upon CXCL12 knockout. Notably, many of these genes are involved in cell signaling, proliferation, and immune-related processes, suggesting that CXCL12 may function as an upstream regulator influencing multiple biological pathways.

Discussion

This study systematically analyzed the immune microenvironment characteristics of SL tissues by integrating multiple batches of transcriptome data and single-cell sequencing data. The results showed that the composition of immune cells in SL tissue underwent significant remodeling, especially $CD4^+$ T cells, with activated $CD4^+$ T cells rising from 17.7% to 41.9% and resting $CD4^+$ T cells declining from 43.9% to 22.8%. Consequently, the overall proportion of $CD4^+$ T cells (resting and activated combined) increased from 61.6% in controls to 64.7% in lymphedema samples, indicating a transition from a resting to an activated state. This quantitative expansion, together with the functional shift, suggests that $CD4^+$ T cells may play a key regulatory role in the disease-related inflammatory microenvironment. However, a key question arises: what are the core regulatory factors driving $CD4^+$ T cell functional reprogramming in SL? To answer this question, this study screened CXCL12 as a potential key molecule through multi-omics integrated analysis. The comprehensive analysis results indicate that CXCL12 is closely related to changes in $CD4^+$ T cell functional status,

 https://doi.org/10.2147/JIR.S597037

intercellular communication, and the interaction between immune cells and the tissue microenvironment, suggesting that it may play a crucial pivotal role in the remodeling of the SL immune microenvironment.

CXCL12 (stromal cell-derived factor-1, SDF-1) is a typical homeostatic chemokine that plays an important role in maintaining tissue homeostasis and the spatial localization of immune cells.[25]

Previous studies have shown that stromal cells including LECs and fibroblasts participate in immune cell recruitment by secreting CXCL12, thereby regulating the local immune microenvironment.[26,27] In addition, a specific type of CXCL12 fibroblast reticular cells in lymph nodes has been shown to regulate $CD4^+$ T cell differentiation and maintain immune homeostasis through the CXCL12-CXCR4 axis.[28]

Based on the above background, we further propose: Does CXCL12 act as a key signaling molecule between stromal cells and $CD4^+$ T cells and participate in the regulation of the SL microenvironment?

CXCL12 exerts its biological functions primarily by binding to its receptor CXCR4. Previous studies have confirmed that the CXCL12-CXCR4 axis plays a crucial role in immune cell migration, tissue retention, and spatial distribution.[29] This is consistent with the results of our intercellular communication analysis, which showed that in the SL microenvironment, $CD4^+$ T cells mainly act as signal receiving cells, while CXCL12-related signals mainly originate from non-lymphocyte populations such as fibroblasts, participating in the construction of communication networks between immune cells and stromal cells. These findings suggest that intercellular communication in the lymphedema microenvironment relies predominantly on soluble factor-mediated paracrine signaling rather than direct cell-cell contact. This mode of communication is closely associated with chronic inflammation and tissue remodeling processes.[30] In this context, the CXCL12–CXCR4 axis, as a canonical secreted signaling pathway, may play a critical role in mediating stromal–immune cell interactions, thereby contributing to $CD4^+$ T cell functional remodeling and fibrotic progression.[31] Therefore, CXCL12 is more likely to serve as a communication signal transmitted from stromal cells such as fibroblasts to $CD4^+$ T cells. It may thus participate in the immune remodeling process of SL by regulating the functional state of T cells.

Further analysis revealed that CXCL12 expression was significantly upregulated in SL tissues, while its classical receptor CXCR4 showed a decreasing trend in $CD4^+$ T cells. This seemingly contradictory expression pattern suggests that the regulation of the CXCL12-CXCR4 axis may not be a simple activation or inhibition process, but rather a dynamic regulatory process. Previous studies have shown that CXCR4, as a G protein-coupled receptor, can undergo GRK-mediated phosphorylation and β-arrestin recruitment under continuous CXCL12 stimulation, thereby inducing receptor endocytosis and transport to the lysosomal degradation pathway, resulting in receptor desensitization.[32,33] Therefore, the pattern of ligand upregulation and receptor downregulation does not necessarily indicate signaling pathway inactivation. Rather, it is more likely to reflect an adaptive feedback regulatory mechanism induced by continuous stimulation in the chronic inflammatory microenvironment.[31] This mechanism helps prevent T cells from becoming functionally exhausted due to excessive chemotactic signaling, thereby maintaining immune system homeostasis.[31] It is worth noting that this overall downregulation of CXCR4 is not entirely consistent across all $CD4^+$ T cell subsets. Given the heterogeneous expression of CXCR4 across $CD4^+$ T cell subsets, future studies should further characterize $CXCR4^+$ $CD4^+$ T cell clusters to determine their functional properties, migratory behavior, and contribution to lymphedema pathogenesis. Such investigations could reveal whether $CXCR4^+$ subsets are preferentially recruited to CXCL12-rich niches and whether they play distinct roles in promoting chronic inflammation and fibrosis. Differences in expression levels among different subsets may endow them with different responsiveness to CXCL12 signaling, thereby creating a fine functional stratification at the site of inflammation.[34,35] This phenomenon further suggests that the CXCL12-CXCR4 axis not only participates in the regulation of cell migration, but may also participate in immune function remodeling by regulating the sensitivity of different T cell subsets. These findings are consistent with and further deepen previous research results. Beyond the dynamic regulation of CXCR4 expression, the functional roles of $CD4^+$ T cells in lymphedema pathogenesis have been well documented.

Previous studies have shown that $CD4^+$ T cells can promote lymphangiogenesis and drive the development of lymphedema through synergistic effects with macrophages, while the depletion of $CD4^+$ T cells can significantly reduce the degree of edema. Our findings are consistent with and further develop previous research confirming that $CD4^+$ T cells play a crucial role in the pathogenesis of SL. For example, Ogata et al demonstrated that $CD4^+$ T cells cooperate with

macrophages to drive the pathogenesis of lymphedema,[8] while Campbell et al reported antigen-driven amplification of $CD4^+$ T cells in the skin.[9] Consistent with this, our intercellular communication analysis revealed extensive interactions between $CD4^+$ T cells and myeloid populations, including macrophages and dendritic cells. Furthermore, this study, through multi-omics integrated analysis, identified fibroblast-derived CXCL12 as a candidate hub molecule in the $CD4^+$ T cell-related network, suggesting that it may regulate the functional state of $CD4^+$ T cells through the CXCL12-CXCR4 pathway. This finding provides a new molecular entry point for understanding the regulatory mechanisms of $CD4^+$ T cells in SL.

From the perspective of functional pathways, CXCL12 expression is positively correlated with various metabolic homeostasis and tissue function maintenance pathways, while it is negatively correlated with cell proliferation and stress response pathways,[36] suggesting that SL tissues are more likely to exhibit a chronic regulatory state characterized by immune microenvironment remodeling rather than an acute inflammatory response.[30] Collectively, module analysis revealed that the M1 module is primarily enriched in genes associated with extracellular matrix remodeling and microenvironmental signaling, suggesting its relevance to stromal cell-related transcriptional programs. In contrast, the M4 module is enriched in genes involved in T cell receptor signaling and cell migration, including ITK, SKAP1, ETS1, and IQGAP1, indicating its association with $CD4^+$ T cell functional regulation. Notably, CXCL12 exhibited moderate associations with both M1 and M4 modules, suggesting its potential role in linking distinct transcriptional programs. Combined with its role as a secreted signaling molecule identified by intercellular communication analysis, CXCL12 may act as a stromal-derived mediator that connects microenvironmental cues with T cell-intrinsic functional programs, thereby contributing to immune–stromal interactions and $CD4^+$ T cell remodeling in SL.

CIBERSORT-based immune cell infiltration analysis showed that CXCL12 expression levels were significantly positively correlated with the proportion of follicular Tfh and significantly negatively correlated with the proportion of resting $CD4^+$ memory T cells. An increase in the proportion of Tfh cells may reflect the persistence of chronic immune activation or local tissue immune responses.[37] In contrast, a decrease in the proportion of resting $CD4^+$ memory T cells suggests a disruption of immune homeostasis.[38] This indicates that the correlation between CXCL12 and immune cell infiltration primarily reflects remodeling of $CD4^+$ T cell functional subsets, rather than changes in the overall number of $CD4^+$ T cells. This characteristic is consistent with previous findings that immune abnormalities in chronic inflammation often manifest as changes in T cell functional subset distribution rather than simple expansion.[39,40] Together with the established role of CXCL12 in regulating T cell migration and tissue colonization,[41,42] these results further support that CXCL12 may participate in SL immune microenvironment remodeling by regulating $CD4^+$ T cell functional lineages.[43]

Notably, qPCR validation results in the mouse model showed that Cxcl12 was significantly upregulated in the model tissues. This in vivo result is highly consistent with the features of Cxcl12 associated with immune microenvironment remodeling revealed in multi-transcriptome and single-cell analyses.

This study further evaluated the overall impact of CXCL12 in the transcriptional regulatory network by virtually knocking it out. The results showed that only about 2% of genes showed significant expression changes after CXCL12 knockout, while the vast majority of genes did not show statistically significant changes. This result suggests that CXCL12 does not act as a global transcriptional driver, but rather its function is more likely limited to a specific set of genes and depends on specific cellular states and microenvironmental contexts. Notably, the genes that showed significant changes after virtual knockout were mainly concentrated in cellular stress response, extracellular matrix regulation, and immune-related processes, including AXL, IFITM3, IGFBP7, PALMD, MAGI1 and ABCB1. The observed changes in these genes collectively suggest that the depletion of CXCL12 leads to alterations in genes associated with immune regulation, matrix remodeling, and cellular stress responses, potentially reflecting a role for CXCL12 in maintaining immune–stromal crosstalk homeostasis.[44,45]

Our findings suggest potential translational applications for CXCL12 in secondary lymphedema. From a diagnostic standpoint, CXCL12 and its associated immune network features may serve as biomarker for assessing immune dysregulation, with potential utility in patient stratification and monitoring disease progression.[46,47] Therapeutically, targeting the CXCL12–CXCR4 axis with CXCR4 inhibitors may complement existing surgical approaches, potentially improving outcomes and reducing relapse.[35,48] Additionally, CXCL12-functionalized biomaterials could offer localized delivery strategies for lymphatic tissue engineering, promoting cell recruitment and immune regulation to support

lymphangiogenesis and tissue integration. Future studies are needed to systematically evaluate these strategies in preclinical models.

Limitations

This study has several limitations. First, all analyses were based on retrospective, publicly available datasets, which may introduce heterogeneity in sampling sites, disease stages, and patient characteristics, as well as potential batch effects and biological confounding. Although we applied integrated multi-omics approaches, these findings remain hypothesis-generating and require validation in independent cohorts using orthogonal methods. Second, the animal model was limited to early-stage changes (up to post-operative day 7), which may not fully recapitulate chronic fibrotic remodeling observed in human lymphedema. Future studies incorporating extended time points and histological analyses are needed to clarify the role of the CXCL12/CXCR4 axis in fibrosis, particularly in fibroblast–$CD4^+$ T cell interactions. In addition, only male mice were used, and potential sex-dependent effects require further investigation. Third, although CXCL12 was identified as a candidate gene associated with immune regulation through multi-omics analysis, functional validation was limited to qPCR, and no gain- or loss-of-function experiments were performed. Moreover, immune infiltration and cell–cell communication were inferred computationally. Future studies should include targeted functional experiments, such as CXCL12 overexpression or knockdown, as well as in vivo validation using genetic or pharmacological approaches. Integration of spatial transcriptomics and longitudinal analyses will further clarify the CXCL12-mediated immune–stromal interaction network in SL.

Conclusion

In summary, this study systematically elucidated the immune regulatory network centered on $CD4^+$ T cells in secondary lymphedema by integrating multi-omics data. Among various analytical strategies, CXCL12 was consistently identified as a leading candidate molecule, and its receptor CXCR4 is stably expressed on $CD4^+$ T cells. Further intercellular communication analysis suggested that the CXCL12–CXCR4 axis may mediate signaling from fibroblasts to $CD4^+$ T cells, highlighting important immune-matrix interactions in the lymphedema microenvironment. Therefore, this study suggests that the CXCL12–CXCR4 axis may play an important role in the immune-matrix interaction and immune microenvironment remodeling in secondary lymphedema. This finding provides a new perspective for understanding the immune regulatory mechanisms of secondary lymphedema and lays a theoretical foundation for further exploration of its feasibility as a potential intervention target. However, direct functional validation is required to establish the causal role of this axis in lymphedema pathogenesis.

Abbreviations

3Rs, Replacement, Reduction, Refinement; ARRIVE, Animal Research: Reporting of In Vivo Experiments; ASCs, Adipose-derived Mesenchymal Stem Cells; AUC, Area Under the Curve; AVMA, American Veterinary Medical Association; cDNA, complementary DNA; CI, confidence interval; CIBERSORT, Cell-type Identification by Estimating Relative Subsets of RNA Transcripts; CNCB, China National Center for Bioinformation; CXCL12, chemokine C-X-C motif chemokine ligand 12; CXCR4, C-X-C chemokine receptor type 4; DEGs, Differentially Expressed Genes; DTS, Decision Tree; ECM, Extracellular Matrix; FC, fold change; FDR, False Discovery Rate; GBM, Gradient Boosting Machine; GEO, Gene Expression Omnibus; glmBoost, Generalized Linear Model Boosting; GMT, Gene Matrix Transposed; GSA-Human, Human Genome Sequence Archive; GSEABase, Gene Set Enrichment Analysis Base; GSVA, Gene Set Variation Analysis; H&E, Hematoxylin and Eosin; hdWGCNA, High-dimensional Weighted Gene Co-expression Network Analysis; hg38, Human Genome Reference Consortium Build 38; HVGs, Highly Variable Genes; IRB, Institutional Review Board; KEGG, Kyoto Encyclopedia of Genes and Genomes; kME, Module Membership; KNN, K-Nearest Neighbors; LEC, Lymphatic Endothelial Cell; NK, Natural Killer; PCA, principal component analysis; PLS, Partial Least Squares; qPCR, Quantitative polymerase chain reaction; RF, Random Forest; RNA, ribonucleic acid; ROC, Receiver Operating Characteristic; S1P, sphingosine-1-phosphate; scRNA-seq, Single-cell RNA sequencing; SDF-1, stromal cell-derived factor-1; SNN, shared nearest neighbor; SPF, specific pathogen-free; SVF, stromal vascular

fraction; SVM, Support Vector Machine; Tfh, follicular helper T; Th1, T helper 1; Th2, T helper 2; Tregs, regulatory T cells; two-way ANOVA, two-way analysis of variance; UMAP, Uniform Manifold Approximation and Projection.

Data Sharing Statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments

We are grateful to Liu et al (2022) for generating and publicly releasing the scRNA-seq dataset (HRA000901) analyzed in this study. Their foundational work provided an essential resource that enabled our investigation of immune network remodeling in secondary lymphedema. We also thank Zi Li for helpful comments and suggestions during the revision of the article.

Author Contributions

Jianwen Xu: Conceptualization; Funding Acquisition; Project Administration; Supervision; Writing - Original Draft Preparation; Writing - Review & Editing.

Xiaoxiao Huang: Data Curation; Methodology, Software; Validation; Visualisation; Writing - Original Draft Preparation.

Chuan Lu: Software; Validation; Writing - Review & Editing.

Zhibiao Tan: Writing - Review and Editing; Software; Data curation.

All authors gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This research received no external funding.

Disclosure

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Brenner E, Hägerling R, Schacht V, et al. Definition, epidemiology and pathophysiology of lymphoedema. *Cells*. 2025;14(24):1955. doi:10.3390/cells14241955
2. Sung CJ, Gupta K, Wang J, et al. Lymphatic tissue bioengineering for the treatment of postsurgical lymphedema. *Bioengineering*. 2022;9(4). doi:10.3390/bioengineering9040162
3. Li CY, Kataru RP, Mehrara BJ, Li CY, Kataru RP, Mehrara BJ. Histopathologic features of lymphedema: a molecular review. *Int J Mol Sci*. 2020;21(7). doi:10.3390/ijms21072546
4. Duhon BH, Phan TT, Taylor SL, et al. Current mechanistic understandings of lymphedema and lipedema: tales of fluid, fat, and fibrosis. *Int J Mol Sci*. 2022;23(12). doi:10.3390/ijms23126621
5. Kim D, Tian W, Wu TT, et al. Abnormal lymphatic sphingosine-1-phosphate signaling aggravates lymphatic dysfunction and tissue inflammation. *Circulation*. 2023;148(16):1231–1249. doi:10.1161/CIRCULATIONAHA.123.064181
6. Fu A, Liu C. The function of T cell immunity in lymphedema: a comprehensive review. *Lymphat Res Biol*. 2023;21(6):556–564. doi:10.1089/lrb.2023.0002
7. Spörlein A, Hirche C, Berner JE, Kneser U, Will PA. Characterization of immune cell infiltration and collagen type III disorganization in human secondary lymphedema: a case-control study. *Plast Reconstr Surg Glob Open*. 2024;12(6):e5906. doi:10.1097/GOX.0000000000005906
8. Ogata F, Fujiu K, Matsumoto S, et al. Excess lymphangiogenesis cooperatively induced by macrophages and CD4+ T cells drives the pathogenesis of lymphedema. *J Invest Dermatol*. 2016;136(3):706–714. doi:10.1016/j.jid.2015.12.001
9. Campbell AC, Stull-Lane AR, Baik JE, et al. Lymphedema pathogenesis involves antigen-driven expansion of CD4+ T cells in skin. *Front Immunol*. 2025;16. 1620571. doi:10.3389/fimmu.2025.1620571
10. García Nores GD, Ly CL, Cuzzone DA, et al. CD4+ T cells are activated in regional lymph nodes and migrate to skin to initiate lymphedema. *Nat Commun*. 2018;9:1970. doi:10.1038/s41467-018-04418-y
11. Ly CL, Cuzzone DA, Kataru RP, Mehrara BJ. Small numbers of CD4+ T cells can induce development of lymphedema. *Plast Reconstr Surg*. 2019;143(3):518e–526e. doi:10.1097/PRS.0000000000005322

12. Yang JCS, Huang LH, Wu SC, et al. Recovery of dysregulated genes in cancer-related lower limb lymphedema after supermicrosurgical lymphaticovenous anastomosis - a prospective longitudinal cohort study. *J Inflamm Res*. 2022;15:761–773. doi:10.2147/JIR.S350421
13. Fu A, Liu C. Analysis of CD4+ T-helper-associated hub gene signature and immune dysregulation via RNA-sequencing data in a mouse tail model of lymphedema. *Gland Surg*. 2023;12(9):1141–1157. doi:10.21037/gs-23-48
14. Zhuo W, Jia L, Song N, et al. The CXCL12–CXCR4 chemokine pathway: a novel axis regulates lymphangiogenesis. *Clin Cancer Res*. 2012;18 (19):5387–5398. doi:10.1158/1078-0432.CCR-12-0708
15. Do LN, Delgado E, Lim CG, et al. A neuro-lymphatic communication guides lymphatic development by CXCL12 and CXCR4 signaling. *Development*. 2024;151(22):dev202901. doi:10.1242/dev.202901
16. Liu KKY, Dorovini-Zis K. Regulation of CXCL12 and CXCR4 expression by human brain endothelial cells and their role in CD4+ and CD8+ T cell adhesion and transendothelial migration. *J Neuroimmunol*. 2009;215(1):49–64. doi:10.1016/j.jneuroim.2009.08.003
17. Cao Y, Qiu Y, Tu G, Yang C. Single-cell RNA sequencing in immunology. *Curr Genomics*. 2020;21(8):564–575. doi:10.2174/ 1389202921999201020203249
18. Ge S, Sun S, Xu H, Cheng Q, Ren Z. Deep learning in single-cell and spatial transcriptomics data analysis: advances and challenges from a data science perspective. *Brief Bioinform*. 2025;26(2):bbaf136. doi:10.1093/bib/bbaf136
19. Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Meth*. 2023;3(6). doi:10.1016/j.crmeth.2023.100498
20. Liu X, Yuan M, Xiang Q, et al. Single-cell RNA sequencing of subcutaneous adipose tissues identifies therapeutic targets for cancer-associated lymphedema. *Cell Discov*. 2022;8:58. doi:10.1038/s41421-022-00402-5
21. du SNP, Hurst V, Ahluwalia A, et al. The ARRIVE guidelines 2.0: updated guidelines for reporting animal research. *PLOS Biol*. 2020;18(7): e3000410. doi:10.1371/journal.pbio.3000410
22. Morita Y, Sakata N, Kawakami R, et al. Establishment of a simple, reproducible, and long-lasting hind limb animal model of lymphedema. *Plast Reconstr Surg Glob Open*. 2023;11(9):e5243. doi:10.1097/GOX.0000000000005243
23. Bustin SA, Benes V, Garson JA, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*. 2009;55(4):611–622. doi:10.1373/clinchem.2008.112797
24. Shepherd R, Cheung AS, Pang K, Saffery R, Novakovic B. Sexual dimorphism in innate immunity: the role of sex hormones and epigenetics. *Front Immunol*. 2021;11. doi:10.3389/fimmu.2020.604000
25. Hang H, Bailey JL, Elks CM. Oncostatin M mediates adipocyte expression and secretion of stromal-derived factor 1. *Biology*. 2019;8(1):19. doi:10.3390/biology8010019
26. Meadows V, Gao N. New kids on the block: immature myeloid cells in intestinal regeneration. *Cell Mol Gastroenterol Hepatol*. 2024;17 (3):499–500. doi:10.1016/j.jcmgh.2023.11.011
27. Nagasawa T. CXCL12/SDF-1 and CXCR4. *Front Immunol*. 2015;6:301. 10.3389/fimmu.2015.00301
28. Yamamura Y, Sabiu G, Zhao J, et al. CXCL12+ fibroblastic reticular cells in lymph nodes facilitate immune tolerance by regulating T cell-mediated alloimmunity. *J Clin Invest*. 2025;135(9):e182709. doi:10.1172/JCI182709
29. De Filippo K, Rankin SM. CXCR4, the master regulator of neutrophil trafficking in homeostasis and disease. *Eur J Clin Invest*. 2018;48(Suppl Suppl 2):e12949. doi:10.1111/eci.12949
30. Hossain L, Gomes KP, Safarpour S, Gibson SB. The microenvironment of secondary lymphedema. The key to finding effective treatments? *Biochimica Et Biophysica Acta*. 2025;1871(3):167677. doi:10.1016/j.bbadis.2025.167677
31. Lu L, Li J, Jiang X, Bai R. CXCR4/CXCL12 axis: “old” pathway as “novel” target for anti-inflammatory drug discovery. *Med Res Rev*. 2024;44 (3):1189–1220. doi:10.1002/med.22011
32. Busillo JM, Benovic JL. Regulation of CXCR4 signaling. *Biochimica et Biophysica Acta*. 2007;1768(4):952–963. doi:10.1016/j. bbamem.2006.11.002
33. Marchese A. Endocytic trafficking of chemokine receptors. *Curr Opin Cell Biol*. 2014;27:72–77. doi:10.1016/j.ceb.2013.11.011
34. Yang Y, Li J, Lei W, et al. CXCL12-CXCR4/CXCR7 axis in cancer: from mechanisms to clinical applications. *Int J Biol Sci*. 2023;19 (11):3341–3359. doi:10.7150/ijbs.82317
35. Mezzapelle R, Leo M, Caprioglio F, et al. CXCR4/CXCL12 activities in the tumor microenvironment and implications for tumor immunotherapy. *Cancers*. 2022;14(9):2314. doi:10.3390/cancers14092314
36. Yagita-Sakamaki M, Ito T, Sakaguchi T, et al. Intestinal Foxl1+ cell-derived CXCL12 maintains epithelial homeostasis by modulating cellular metabolism. *Int Immunol*. 2025;37(4):235–250. doi:10.1093/intimm/dxae068
37. Gao X, Luo K, Wang D, et al. T follicular helper 17 (Tfh17) cells are superior for immunological memory maintenance. *Elife*. 2023;12:e82217. doi:10.7554/eLife.82217
38. Okoye AA, Picker LJ. CD4+ T cell depletion in HIV infection: mechanisms of immunological failure. *Immunol Rev*. 2013;254(1):54–64. doi:10.1111/imr.12066
39. Kulesh V, Peskov K, Helmlinger G, Bocharov G. Systematic review and quantitative meta-analysis of age-dependent human T-lymphocyte homeostasis. *Front Immunol*. 2025;16:1475871. doi:10.3389/fimmu.2025.1475871
40. Mao L, Feng Q, Luo OJ, Chen G, Leng XS. Human T cell development and aging: remodeling throughout the lifespan. *Aging Res*. 2024;2 (1):9340021. doi:10.26599/AGR.2024.9340021
41. Miao R, Lim VY, Kothapalli N, et al. Hematopoietic stem cell niches and signals controlling immune cell development and maintenance of immunological memory. *Front Immunol*. 2020:11. 10.3389/fimmu.2020.600127.
42. Shi J, Wei Y, Xia J, et al. CXCL12–CXCR4 contributes to the implication of bone marrow in cancer metastasis. *Fut Oncol*. 2014;10(5):749–759. doi:10.2217/fon.13.193
43. Biajoux V, Bignon A, Bouchet-Delbos L, Émilie D, Balabanian K. Anomalies de l’axe de signalisation CXCL12 (SDF-1)/CXCR4 dans le syndrome WHIM et la lymphopénie T $CD4^+$ idiopathique. *Biologie Aujourd’hui*. 2010;204(4):273–284. doi:10.1051/jbio/2010022
44. Lemke G. Biology of the TAM Receptors. *Cold Spring Harb Perspect Biol*. 2013;5(11):a009076. doi:10.1101/cshperspect.a009076
45. Zhang B, Wang R, Tao S, et al. Stromal cell–derived chemokines modulate immune cells in inflammation: new findings and future perspectives. *J Immunol*. 2025;214(11):2822–2835. doi:10.1093/jimmun/vkaf136

46. Ma M, Wang S, Wang K, Jiang B, Li J, Hou S. CXCL12 links bladder cancer and diabetes as a potential biomarker. *Sci Rep*. 2025;15:19017. doi:10.1038/s41598-025-01357-9
47. Kim SK, Choe JY, Park KY. CXCL12 and CXCR4 as novel biomarkers in uric acid-induced inflammation and patients with gouty arthritis. *Biomedicines*. 2023;11(3):649. doi:10.3390/biomedicines11030649
48. Karin N, Wildbaum G. The role of chemokines in shaping the balance between CD4+ T cell subsets and its therapeutic implications in autoimmune and cancer diseases. *Front Immunol*. 2015;6:609. doi:10.3389/fimmu.2015.00609