



# Gene expression signatures from single-cell transcriptomics predict Sjögren's syndrome

Shumin Cheng<sup>1\*</sup>, Ling Yan<sup>1\*</sup>, Yilin Wu<sup>2\*</sup>, Yan Li<sup>1\*</sup>, Ziheng Ye<sup>1\*</sup>, Zhen Zhang<sup>3\*</sup>,  
Bihun Huang<sup>3,4\*</sup>, Yingqian Mo<sup>3,5\*</sup>, Guan Yang<sup>6\*</sup>, HaiYan Yan<sup>3\*</sup>, Jun Liang<sup>1\*</sup>

<sup>1</sup>The Seventh Affiliated Hospital, Sun Yat-sen University;

<sup>2</sup>Shenzhen University of Advanced Technology General Hospital

<sup>3</sup>Shenshan Medical Center, Sun Yat-sen Memorial Hospital, Sun Yat-Sen University;

<sup>4</sup>The People's Hospital of Guangxi Zhuang Autonomous Region;

<sup>5</sup>Sun Yat-sen Memorial Hospital, Sun Yat-sen University;

<sup>6</sup>Department of Infectious Diseases and Public Health, City University of Hong Kong



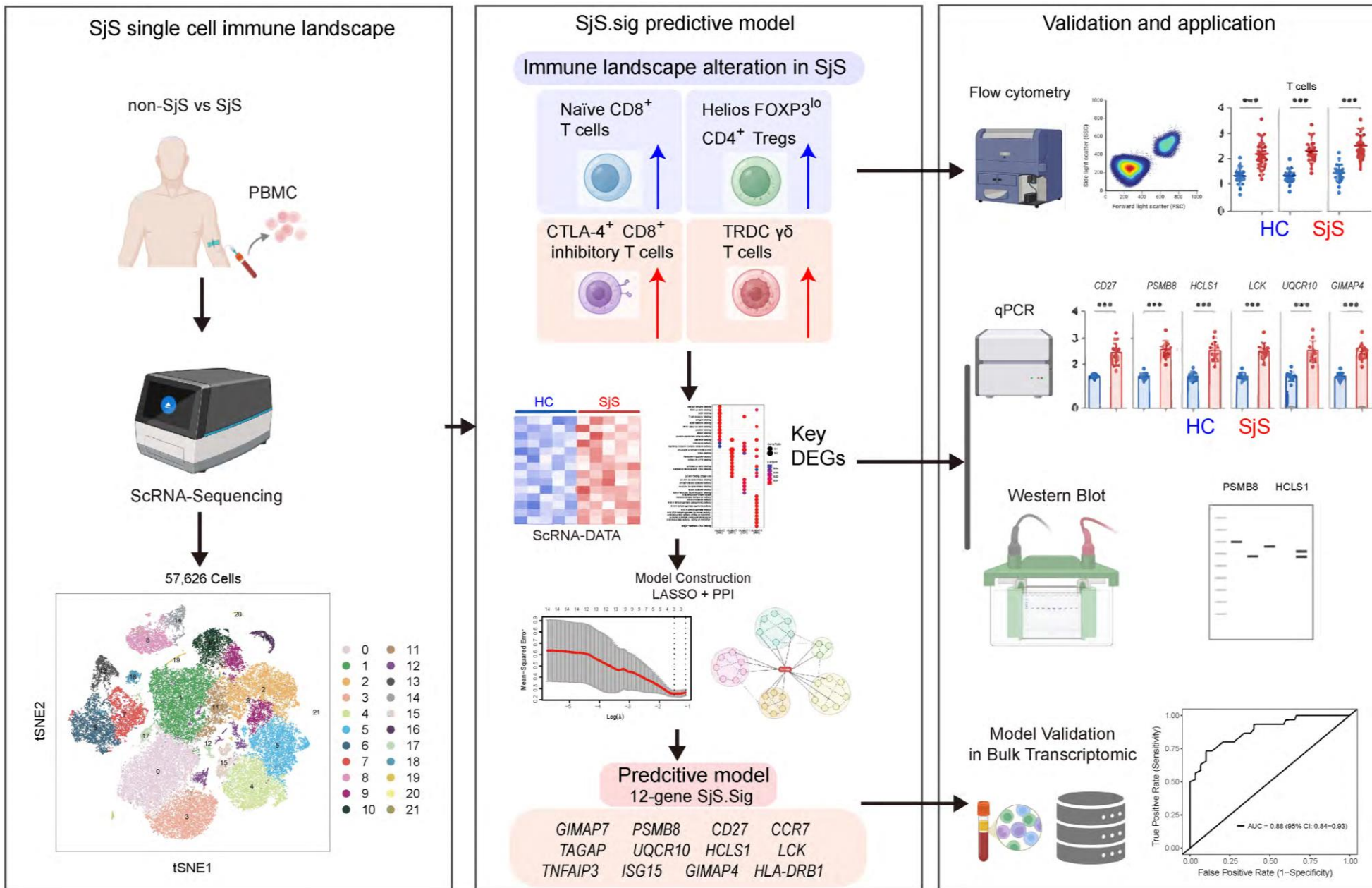
Shumin Cheng, Ling Yan, Yilin Wu, Yan Li, Ziheng Ye, Zhen Zhang, Yi Zhang, et al. 2026.

Gene expression signatures from single-cell transcriptomics predict Sjögren's syndrome.

*iMetaOmics* 3: e70106. <https://doi.org/10.1002/imo2.70106>



# Introduction

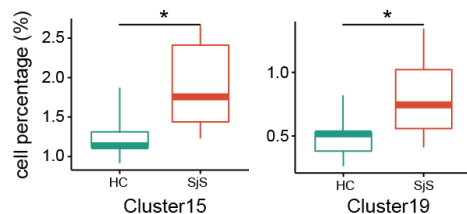




# Highlights

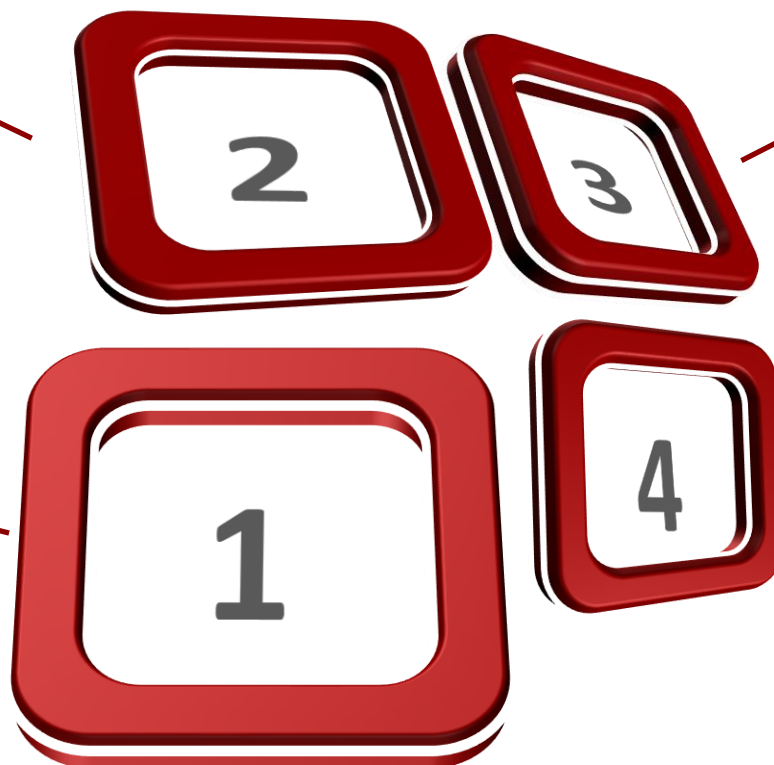
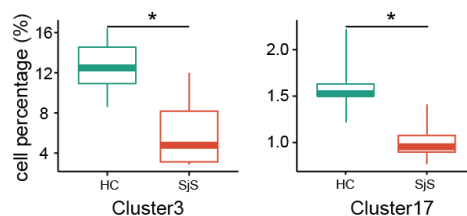
## Increased cluster

The levels of **CTLA-4<sup>+</sup> CD8<sup>+</sup> Inhibitory T cells** and **TRDC  $\gamma\delta$  T cells** were significantly elevated in SjS patients.



## Decreased cluster

Patients with Sjögren's syndrome (SjS) have significantly reduced levels of **Naïve CD8<sup>+</sup> T cells** and **Helios FOXP3<sup>lo</sup> CD4<sup>+</sup> Treg cells** in their peripheral blood.



## Upregulation of core genes

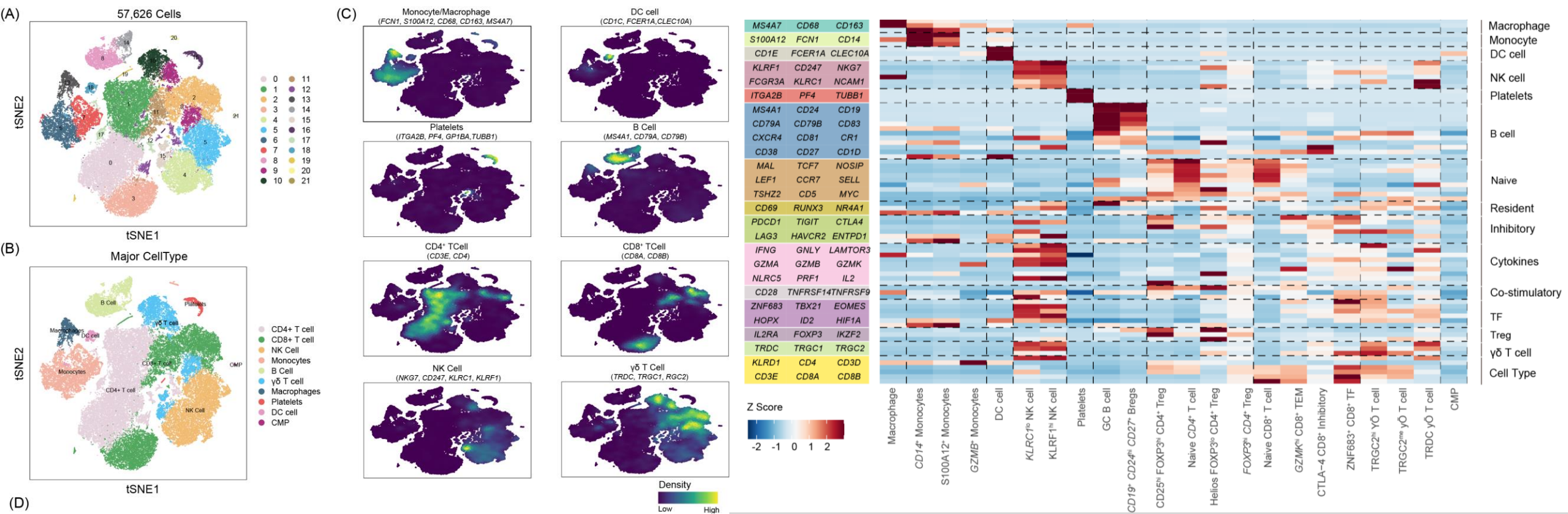
The expression levels of key genes such as *CD27*, *PSMB8*, *HCLS1*, *LCK*, *UQCR10*, and *GIMAP4* were significantly upregulated in SjS patients.

## Model Development and Validation

Based on single-cell transcriptome sequencing (scRNA-seq) data, a predictive model for the clinical diagnosis of SjS was successfully constructed and validated.



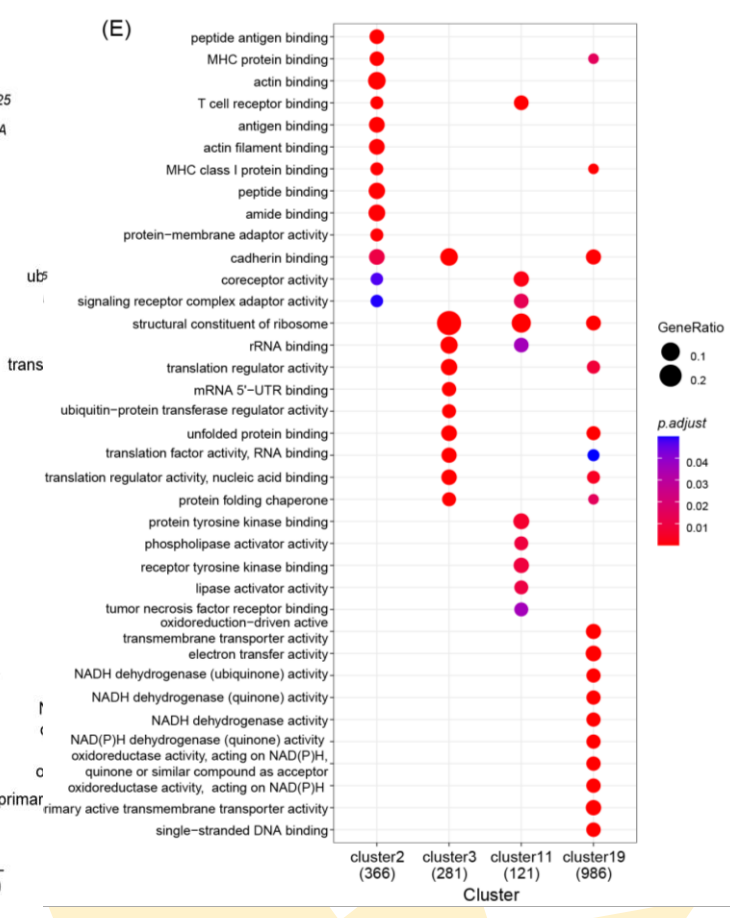
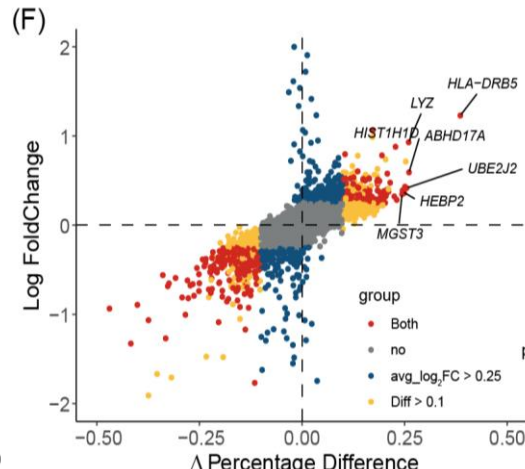
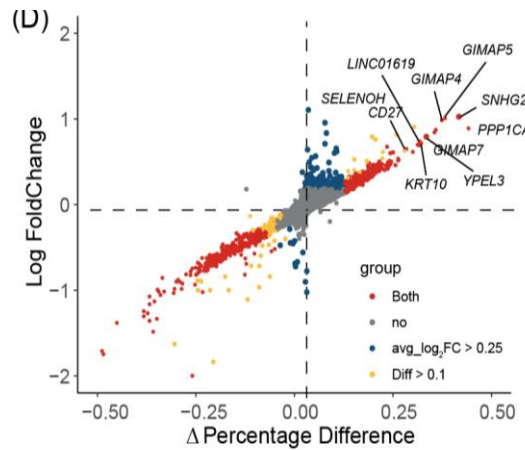
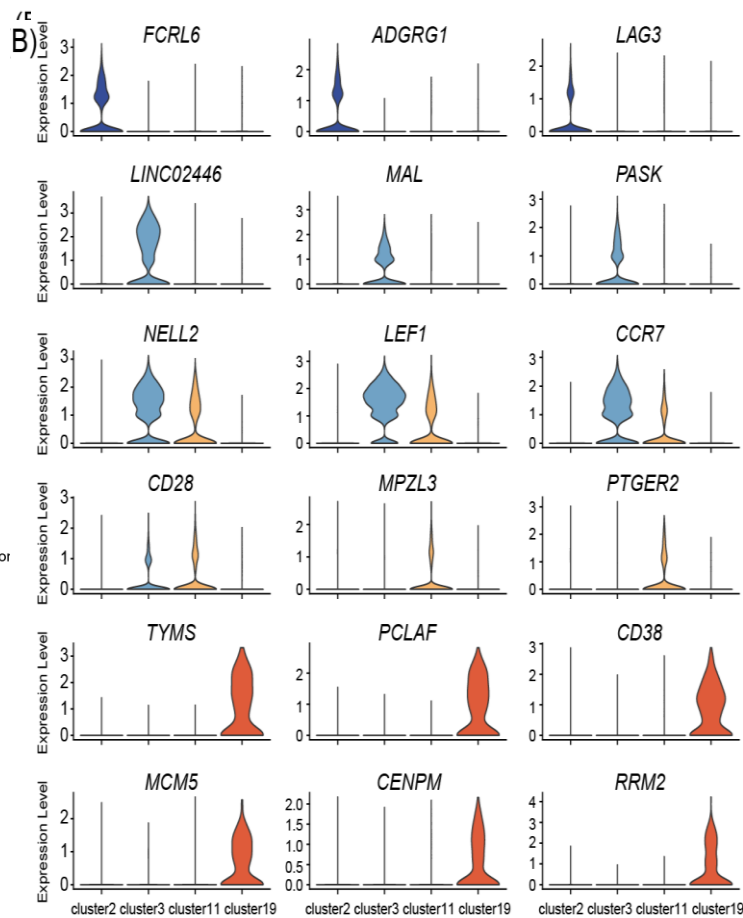
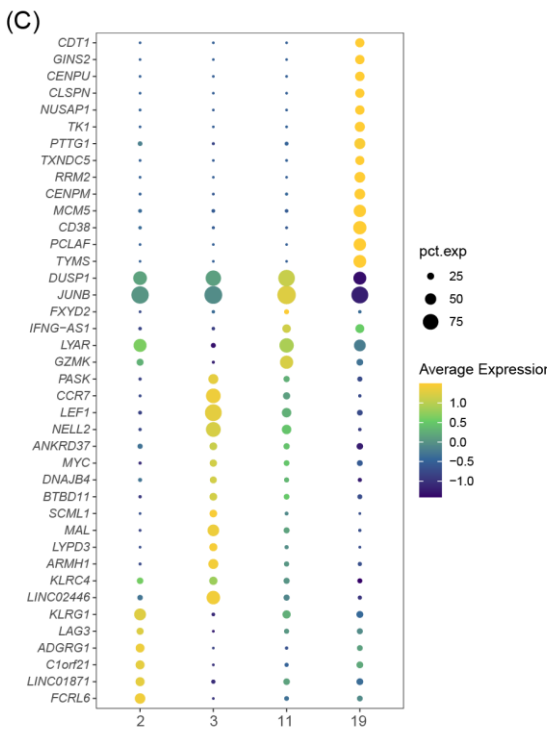
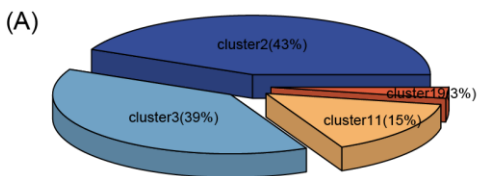
# Result 1: Single-cell transcriptome atlas and cell cluster annotation of PBMCs from SJS patients



- **Data Scale:** 57,626 high-quality single-cell datasets were successfully captured and analyzed.
- **Clustering Results:** 22 initial clusters were identified and further integrated into 11 major cell populations.
- **Fine Annotation:** Subpopulation subdivision of CD4<sup>+</sup> and CD8<sup>+</sup> T cells was emphasized, covering functional subtypes such as Naïve, repressive, Treg, and  $\gamma\delta$  T.
- **Validation Methods:** The biological accuracy of cell population definitions was ensured by combining Nebulosa gene density plotting and characteristic gene heatmaps.



# Result 2: Identification and Characterization of Core Cell Subpopulations



## Core Cell Clusters:

- 1) The levels of Naïve CD8<sup>+</sup> T cells (Cluster 3) and Helios FOXP3<sup>lo</sup> CD4<sup>+</sup> Treg cells (Cluster 17) were significantly decreased.
- 2) The levels of TRDC γδ T cells (Cluster 15) and CTLA-4<sup>+</sup> CD8<sup>+</sup> T cells (Cluster 19) were significantly increased.

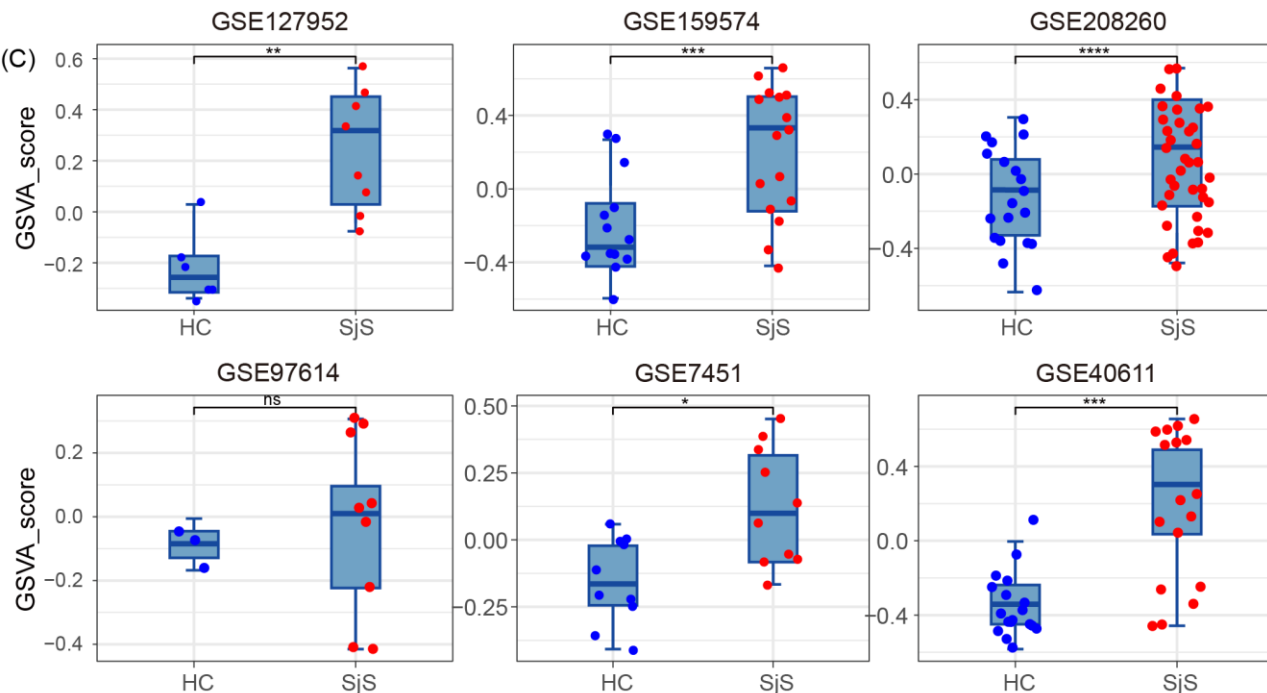
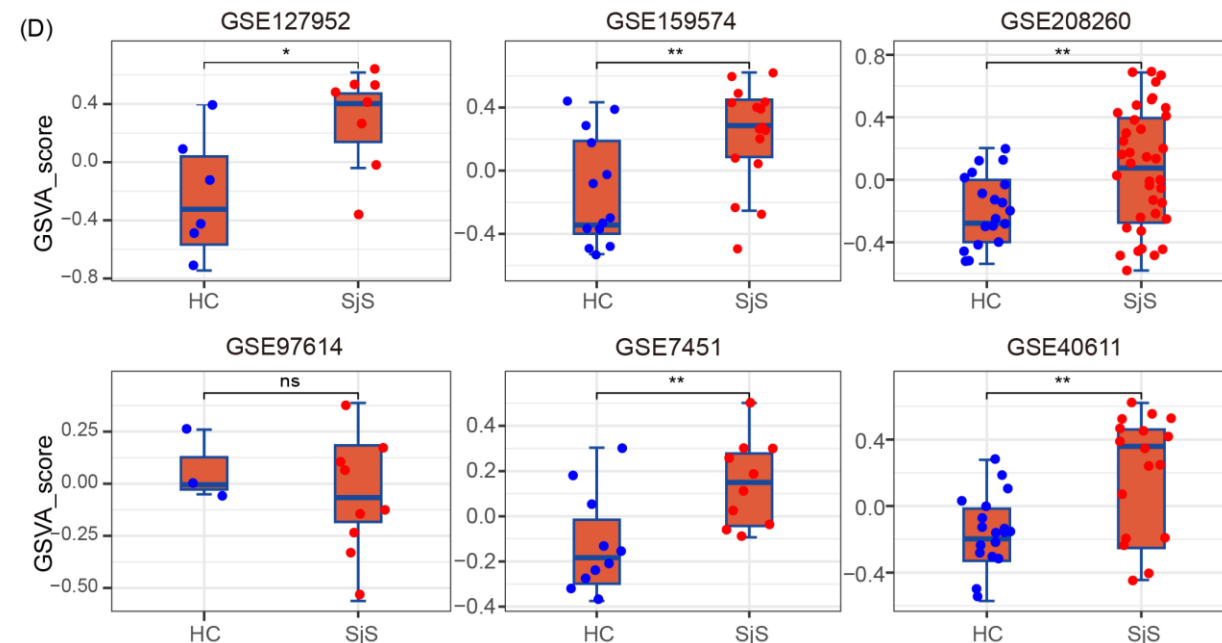
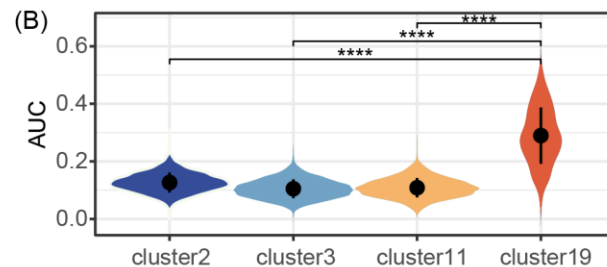
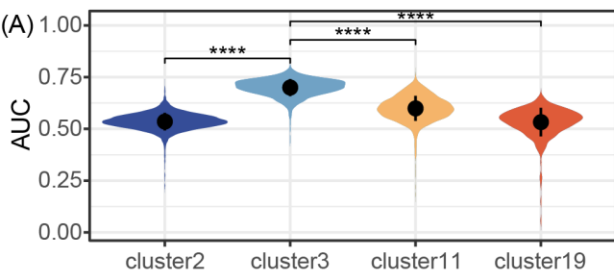
Comprehensive Analysis: Taking Cluster 3 as an example, the standardized gene set selection has covered the three key subpopulations: Cluster 17, Cluster 15, and Cluster 19.

## Selection of Core Gene Set:

- 1) Cluster-specific genes,
- 2) Pathway-specific genes with functional enrichment,
- 3) Pathway-specific genes with functional enrichment.



# Result 3: Computational Validation and External Dataset Evaluation of Core Cell Subpopulations



## Model Training and Screening (LASSO)

The most representative gene set was refined from the candidate genes highly expressed in SjS using the LASSO regression algorithm.

## Single-Cell Level Scoring (AUCell)

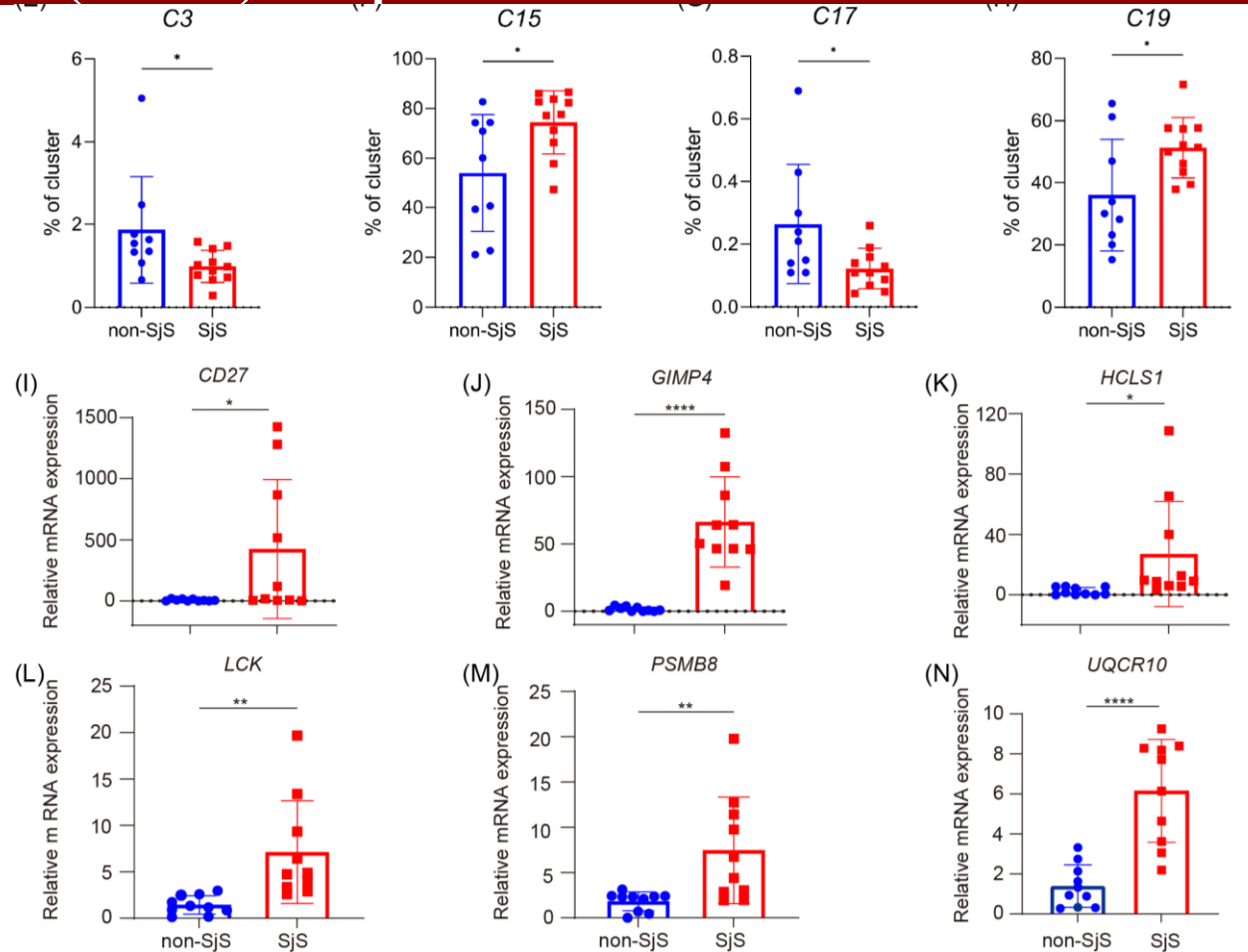
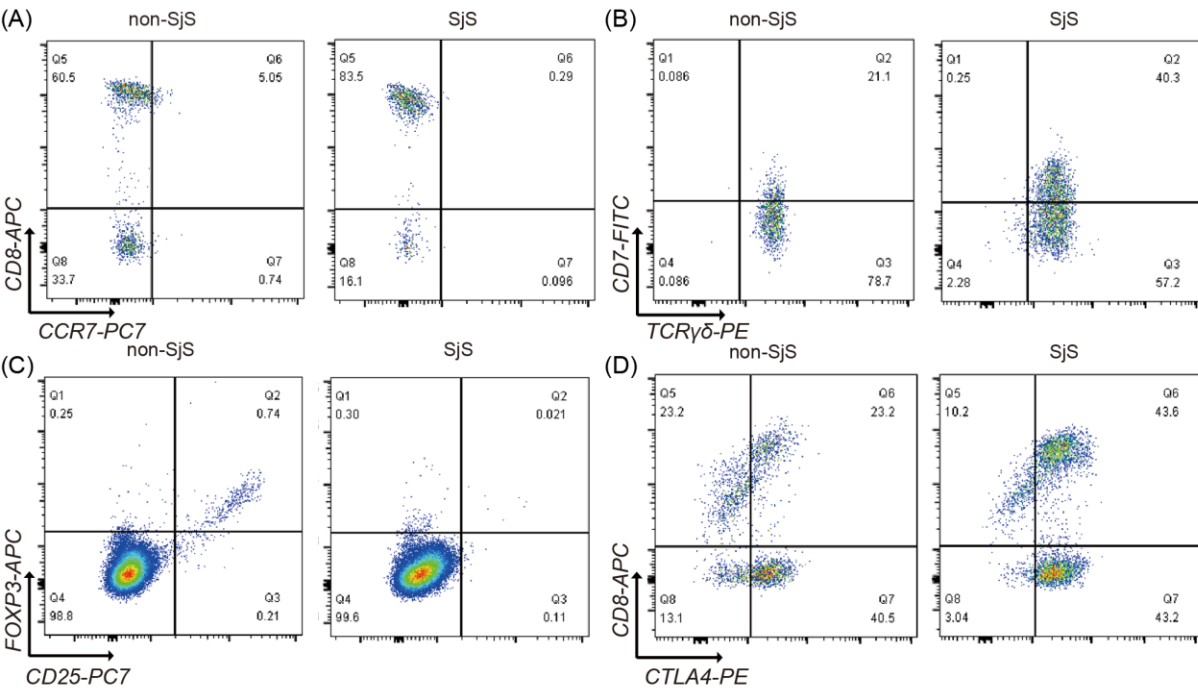
High activity of characteristic genes in corresponding subpopulations (e.g., Naïve CD8<sup>+</sup> and CTLA-4<sup>+</sup> CD8<sup>+</sup> Inhibitory T cells) was determined by gene set scoring.

## Gene Set Variation Analysis (GSVA)

Validation was performed on 6 publicly available gene expression datasets. The GSVA scores of SjS patients were consistently higher than those of the healthy group.



# Result 4: Experimental Validation of Human Peripheral Blood Mononuclear Cell (PBMC) Samples



## Flow Cytometry (FACS) Analysis

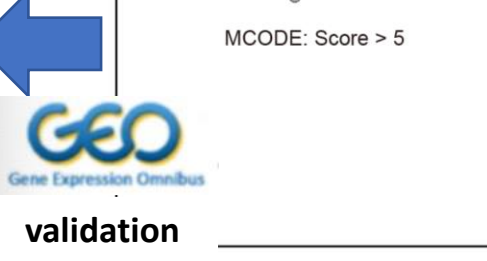
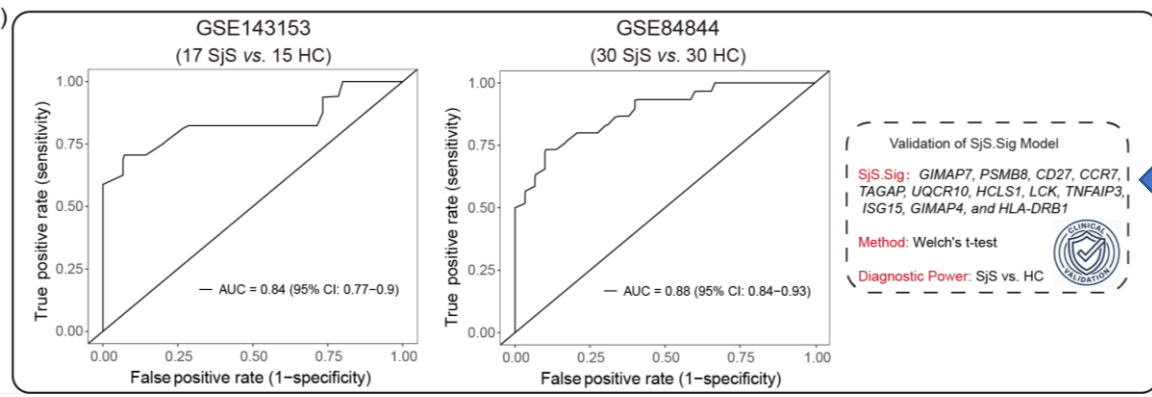
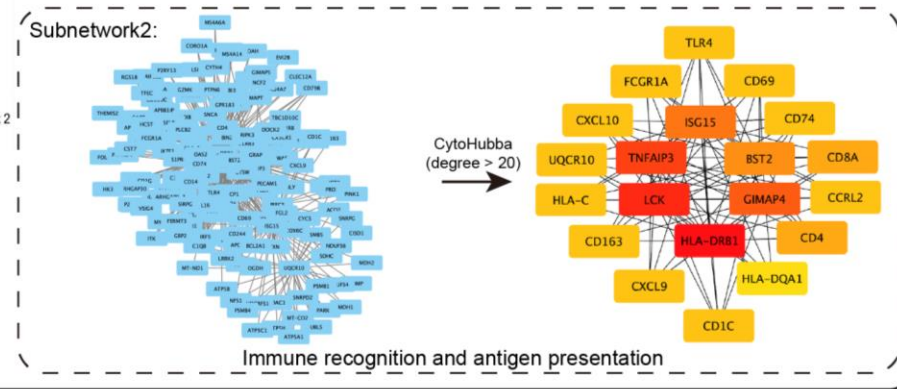
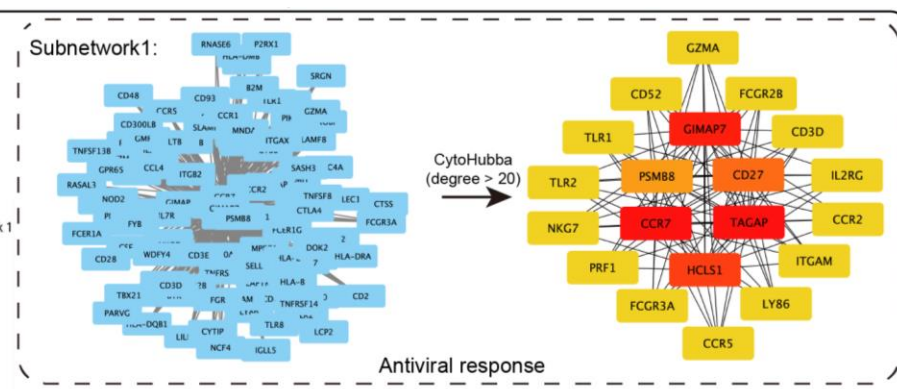
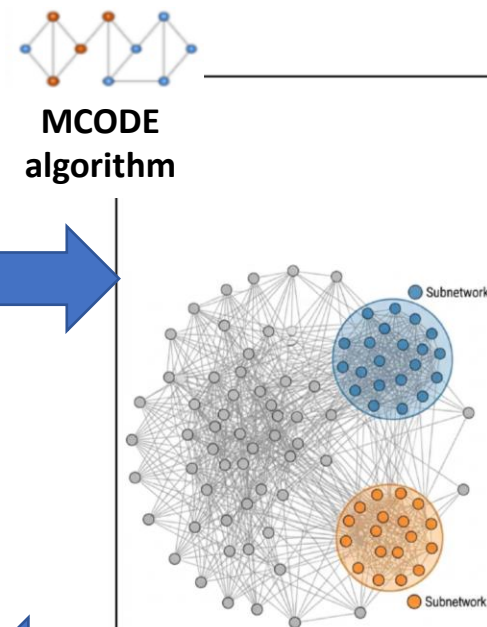
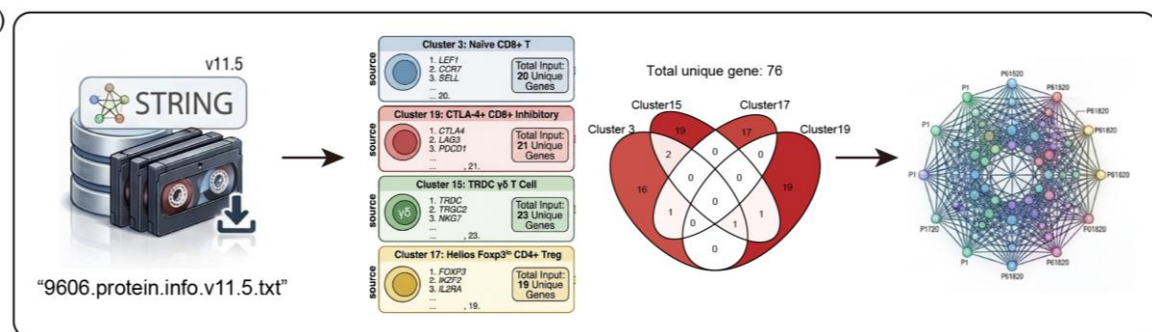
- Sample Size:** PBMC samples from 11 SjS patients and 9 controls (non-SjS) were analyzed.
- Cluster Proportions:**
  - 1) Significantly decreased : C3 (Naïve CD8<sup>+</sup> T) and C17 (Helios FOXP3<sup>lo</sup> CD4<sup>+</sup> Treg).
  - 2) Significantly increased: C15 (TRDC γδ T) and C19 (CTLA-4<sup>+</sup> CD8<sup>+</sup> Inhibitory T)

## Core Feature Gene qPCR Analysis

- Sample Size:** PBMC samples from 10 SjS patients and 10 controls (non-SjS) were analyzed.
- Gene expression results:** *CD27*, *GIMAP4*, *HCLS1*, *LCK*, *PSMB8* and *UQCR10* were all significantly upregulated in the SjS group ( $p < 0.05$ ).



# Result 5: Gene Prediction Model Construction and Validation



## 01 SjS.Sig Establishment

- MCODE (score > 5) + CytoHubba (Degree > 20)
- **Conclusion:** A prediction model (SjS.Sig) consisting of 12 core genes was finally refined.)

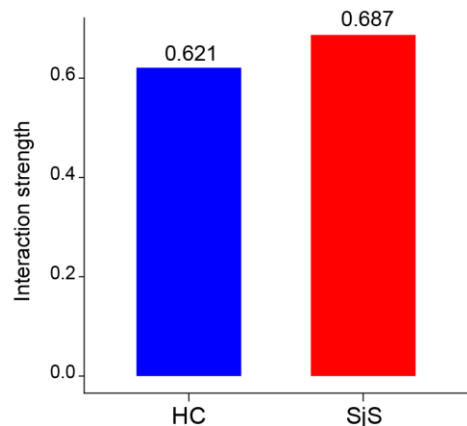
## 02 External Independent Dataset Validation

- Excellent discrimination was demonstrated in both GSE84844 (AUC=0.88) and GSE143153 (AUC=0.84).
- **Conclusion:** SjS.Sig is a biomarker with high clinical translational potential.

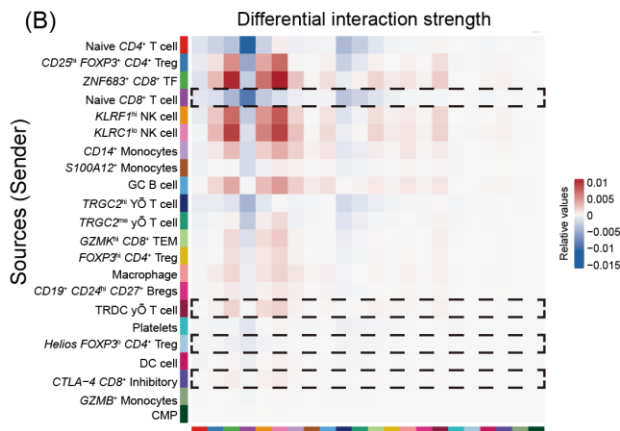


# Result 6: Remodeling of the Sjs Intercellular Communication Map

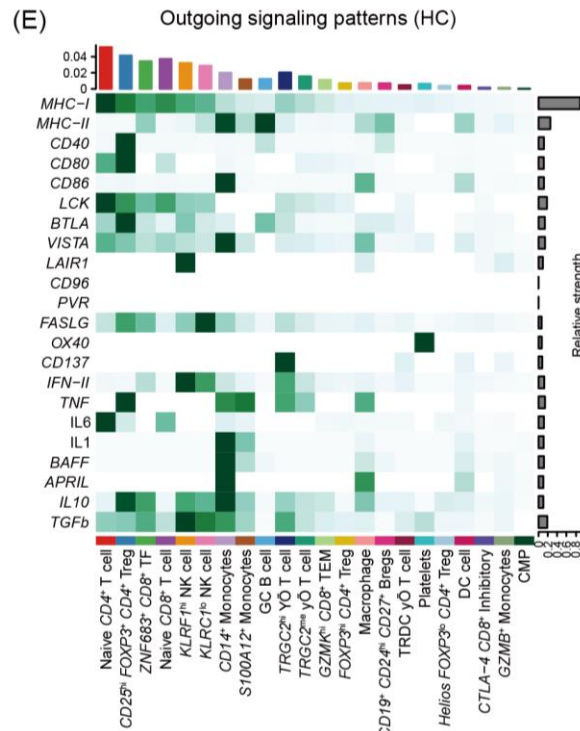
(A)



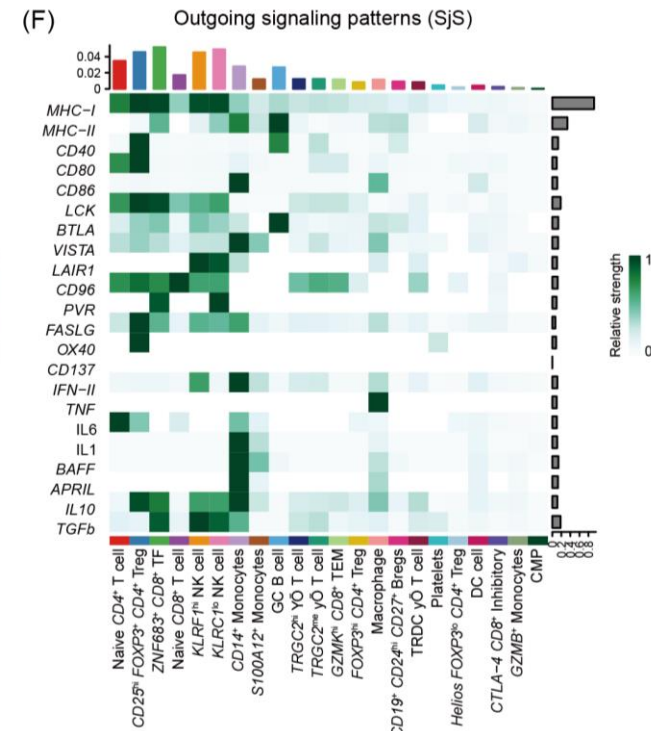
(B)



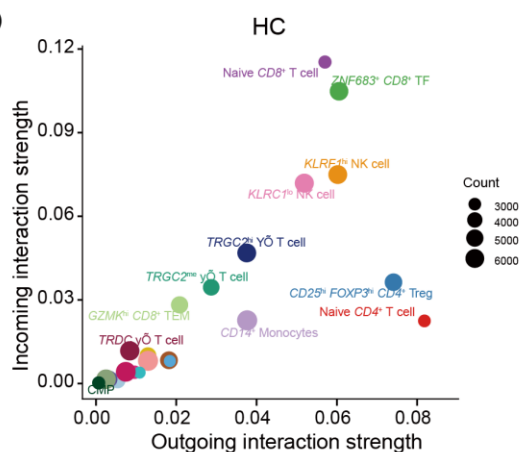
(E)



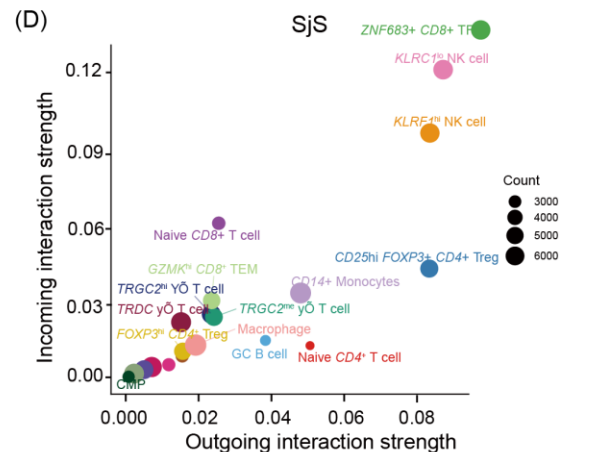
(F)



(C)



(D)



## 1. Significantly Enhanced Overall Interaction Strength

- **Signal Overload:** The overall intercellular communication strength in the Sjs group was significantly higher than that in the healthy control group (HCs).
- **Pathological Background:** Reflects the highly active immune signaling network and the maintenance of the pro-inflammatory microenvironment in disease states.

## 2. Naive CD8<sup>+</sup> T Cells

- **Communication Suppression Function Impairment:** Both afferent and efferent interactions were significantly reduced.
- **Pathway Dissociation:** The MHC-I signaling pathway showed selective attenuation, while MHC-II remained stable, suggesting impaired immune homeostasis and recognition function.

## 3. Active State of Core Regulatory Subpopulations

- **Enhanced Signaling:** The interactions between TRDC γδ T cells and CTLA-4<sup>+</sup> CD8<sup>+</sup> Inhibitory T cells and other cell populations were significantly enhanced.
- **Regulatory Role:** Suggests that these key subsets play a central role in mediating Sjs immune remodeling, chronic inflammation, and tissue damage.



# Summary

- ❑ **Key Immunological Signature Identification:** Deep analysis using scRNA-seq identified Naïve CD8<sup>+</sup> T cells, Helios FOXP3<sup>lo</sup> CD4<sup>+</sup> regulatory T cells, TRDC  $\gamma\delta$  T cells, and CTLA-4<sup>+</sup> CD8<sup>+</sup> inhibitory T cells as key cytological markers distinguishing SjS patients from healthy controls.
- ❑ **Novel Predictive Model Construction:** A novel predictive model, SjS.Sig, comprising 12 core genes, was constructed using machine learning methods. This model demonstrated robustness and excellent diagnostic efficacy in multiple independent clinical cohorts.
- ❑ **Clinical Translation and Application Value:** Flow cytometry, qPCR analysis, and Western blotting (WB) confirmed that SjS is a highly sensitive molecular diagnostic and prognostic tool, providing a theoretical basis for precision immunotherapy.

Shumin Cheng, Ling Yan, Yilin Wu, Yan Li, Ziheng Ye, Zhen Zhang, Yi Zhang, et al. 2026.

Gene expression signatures from single-cell transcriptomics predict Sjögren's syndrome.

*iMetaOmics* 3: e70106. <https://doi.org/10.1002/imo2.70106>

# iMeta: To be top journals in biology and medicine


# WILEY





“**iMeta**” launched in 2022 by iMeta Science Society, **impact factor (IF) 33.2**, ranking **top 65/22400 in the world**. It aims to publish innovative and high-quality papers with broad and diverse audiences. Its scope is similar to *Cell*. The average citation is > 40 in 2025, similar to *Nature* and *Science*. Its unique features include video abstract, bilingual publication, and social media with 600,000 followers. Indexed by [SCIE/ESI](#), [PubMed](#), [Google Scholar](#) etc.

“**iMetaOmics**” launched in 2024, indexed by [ESCI](#), [PubMed](#), [Google Scholar](#), with a **target IF>15**, and its scope is similar to *Nature Communications*, *Science Advances*, *Advanced Science*, *Nucleic Acids Research*, etc.

“**iMetaMed**” launched in 2025, with a target IF>15, similar to *Med*, *Cell Reports Medicine*, *eBioMedicine*, *eClinicalMedicine* etc.

 Society: <http://www.imeta.science>  
 Publisher: <https://wileyonlinelibrary.com/journal/imeta>  
iMeta: <https://wiley.atyponrex.com/journal/IMT2>  
Submission: iMetaOmics: <https://wiley.atyponrex.com/journal/IMO2>  
iMetaMed: <https://wiley.atyponrex.com/journal/IMM3>

 [iMetaScience](#)  
 [iMetaScience](#)

 [office@imeta.science](mailto:office@imeta.science)  
[imetaomics@imeta.science](mailto:imetaomics@imeta.science)  
 [Promotion Video](#)

Update  
2026/3/30