Front. Med. 2025, 19(6): 1117−1130
https://doi.org/10.1007/s11684-025-1165-z

RESEARCH ARTICLE

Mutually reinforcing relationship between Sjögren's syndrome and lung adenocarcinoma: insights from Mendelian randomisation, single-cell, and transcriptomic analyses

Kai Xu[1,2,3,*], Manhua Wang[4,*], Zixuan Yang[5,*], Yu Tang[4], Zhen Li[4], Tao Liu[4], Yu Wang (✉)[6], Yuqing Wang (✉)[7], Xiaoqian Zhai (✉)[2,8]

[1]Department of Thoracic Surgery, West China Hospital, Sichuan University, Chengdu 610041, China; [2]Lung Cancer Center/Lung Cancer Institute, West China Hospital, Sichuan University, Chengdu 610041, China; [3]Department of Thoracic Surgery, China-Japan Friendship Hospital, Beijing 100029, China; [4]West China School of Medicine, Sichuan University, Chengdu 610041, China; [5]West China School of Basic Medical Sciences & Forensic Medicine, Sichuan University, Chengdu 610041, China; [6]Department of Mathematics, Louisiana State University, Field House Dr, Baton Rouge, LA 70802, USA; [7]Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY 10466, USA; [8]Department of Medical Oncology, Cancer Centre, West China Hospital, Sichuan University, Chengdu 610041, China

Abstract **Increasing evidence suggests an association between Sjögren's syndrome (SS) and multiple cancers; however, the causal relationships and regulatory mechanisms remain unclear. Using European genome-wide association study data, we employed Mendelian randomisation (MR) and meta-analysis to explore the SS-cancer causality. Bidirectional two-sample MR revealed that SS increased the risk of colorectal (odds ratio (OR) = 1.08) and lung cancers (OR = 1.15), whereas lung (OR = 3.03) and female genital cancers (OR = 6.59) were found to elevate the risk of SS. Co-localisation analysis confirmed bidirectional causality between SS and lung cancer. Summary data-based MR identified *HLA-DPB2* as a hub gene, with single-cell RNA and mRNA analyses suggesting its role in memory B cells regulation via MHC-II ligands and lung carcinogenesis. This study demonstrates a mutually positive association between SS and lung cancer, implicating *HLA-DPB2* as a potential regulatory gene and offering novel insights into the relationship between SS and cancer, especially lung cancer.**

Keywords *HLA-DPB2*; lung adenocarcinoma; Mendelian randomization; single-cell transcriptome; Sjögren's syndrome

Introduction

Sjögren's syndrome (SS) is a systemic autoimmune disease characterized by lymphocytic infiltration of the exocrine glands, leading to dryness of the affected glands [1]. The diagnosis depends on serological tests, especially the presence of SSA and SSB antibodies [2]. Along with systemic lupus erythematosus and progressive systemic sclerosis, SS is one of the most common autoimmune disorders, with an incidence of approximately 4 cases per 1000 individuals every year and a prevalence of approximately 0.5%; however, SS has received much less attention than the other two conditions [3].

Received June 17, 2025; accepted August 4, 2025

Correspondence: Xiaoqian Zhai, xiaoqzhai@scu.edu.cn;
Yuqing Wang, Yuqwang@montefiore.org;
Yu Wang, yuwang@lsu.edu

*These authors contributed equally to this work.

The increased risk of lymphoma in patients with primary Sjögren's syndrome (pSS) has been widely demonstrated, with non-Hodgkin lymphoma (NHL) being the most common malignancy [2,4]. A meta-analysis showed that, except for NHL, patients with pSS also had a higher risk of lung cancer, although the histological subtypes were not explored. Other cancers with elevated risks in patients with pSS include oral and throat, non-melanoma skin, and urinary tract cancer [5]. Recently, a study that utilized two-sample Mendelian randomization (MR) analysis for the first time showed that SS could increase the risk of prostate, endometrial, urinary tract, liver, and bile duct cancers, but did not find a significant causal relationship of SS with lung and breast cancers [6]. Thus, although it has been documented that SS is linked to a higher overall rate of cancers, its correlation with specific tumors remains controversial [7,8].

Immune cells play an important role in mediating both SS and cancer. However, it remains challenging to

determine whether SS has a causal effect on solid cancers or to identify the specific cellular or molecular factors involved in their development. Human leukocyte antigen (HLA) is a cell surface glycoprotein that presents antigens to T cells and plays a key role in immune recognition and responses [9]. A meta-analysis identified that HLA class II is associated with pSS, suggesting that alleles such as DRB1*03:01, DQA1*05:01, and DQB1*02:01 could be risk factors [10]. HLA class II is also considered as a driver of SS variability by regulating the expression level of IFN-α [11]. Additionally, studies have shown that patients with SS primarily present methylation alterations in B cells that are associated with disease development [12].

MR, an epidemiological genetic approach, is a valuable tool for investigating the potential causal relationships and molecular connections between different traits. A study using two-sample MR, which is widely used to explore potential causality while minimising confounding factors often found in observational studies, can reveal the genetic interaction mechanism between the two diseases and elucidate the role of some immune cells [13,14]. In addition, summary data-based Mendelian randomization (SMR) was applied to identify the genes shared between SS and lung adenocarcinoma (LUAD). By utilizing the expression quantitative trait locus (eQTL) and genome-wide association studies (GWAS), SMR can explore shared risk genes through gene expression analysis, offering insights into the comorbidity of different diseases [15,16].

In this study, we investigated the causal relationships between SS and various cancers, with a particular focus on the interactions between SS and LUAD. Using eQTL, bulk tissue RNA sequencing (bulk tissue RNA-seq), and single-cell RNA sequencing (scRNA-seq) data, we explored the potential underlying mechanisms of this interaction.

Materials and methods

Research flowchart introduction

First, GWAS summary statistics, which included three SS datasets and 248 cancer datasets (Fig. 1), were used in the two-sample MR to estimate the causality between SS and cancers through forward and backward MR, with sensitivity and co-localization analyses for validation.

Second, as MR revealed that there may be a mutually reinforcing relationship between lung cancer and SS, we further explored the relationship between SS and LUAD at the cellular and genetic levels. The blood cis-eQTL data from eQTLGen were utilized for SMR analysis on SS and LUAD, and intersection analysis was then conducted to identify the hub genes of the causality between SS and LUAD. Gene expression analysis in The Cancer Genome Atlas (TCGA) data set of LUAD confirmed *HLA-DPB2* as a potential hub gene.

Third, bulk RNA-based and scRNA-based analyses were performed to investigate the potential mechanisms and functions of *HLA-DPB2*. In the bulk RNA-based analysis, weighted correlation network analysis (WGCNA), Gene Ontology (GO), Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment, CIBERSORT, protein–protein interaction (PPI) network and correlation analysis were used, and a potential relationship between the immune and hub genes was revealed. For scRNA-based analysis, differential expression analysis, gene set enrichment analysis (GSEA), Monocle, and Cellchat were used to further explore and validate the potential mechanisms.

Data acquisition

GWAS summary statistics

All traits uploaded to the MRC-IEU online database were downloaded (updated to 2024.05.13, $N = 50\ 044$), and the GWAS summary results for SS and all cancer traits were obtained from the MRC-IEU online database, including 3 data sets of SS in total and 248 data sets for cancer of European origin. Details of the GWAS summary traits are presented in Table S1.

For SMR, the analysis was based on the summary data of whole-blood cis-eQTL summary statistics, which were downloaded from eQTLGen (a meta-analysis of 14 115 individuals).

Bulk tissue RNA-seq and scRNA-seq data

Bulk tissue RNA-seq gene expression data related to lung cancer were obtained from TCGA database via XenaBrowser, while scRNA-seq data were acquired from the Gene Expression Omnibus (GEO) repository (accession: GSE131907). All pSS patient samples were obtained from the GEO database, including two bulk transcriptomic data sets (accession numbers: GSE173670 and GSE66795) and one scRNA-seq data set (accession number: GSE253568). The "SingleR" and "Seurat" R packages were used for processing the scRNA-seq data, cell clustering, and annotation.

Single nucleotide polymorphisms selection and bi-directional MR

To evaluate the causal relationship between SS and different types of cancer, an MR analysis was performed using the R package TwoSampleMR v0.6.3. The instrument variants were mainly selected based on three core assumptions: (1) the variants are associated with the exposure, (2) the variants have no connection with the

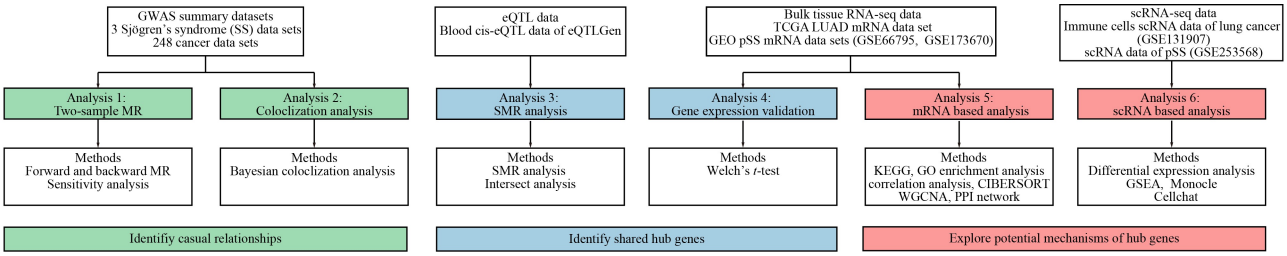

Fig. 1 Overview of this MR-based study. GWAS, genome-wide association studies; eQTL, expression quantitative trait locus; SMR, summary data-based Mendelian randomization; TCGA, The Cancer Genome Atlas; WGCNA, weighted correlation network analysis; GO and KEGG enrichment, Gene Ontology and Kyoto Encyclopaedia of Genes and Genomes enrichment; PPI network, protein−protein interaction network; GSEA, gene set enrichment analysis.

outcome through confounding factors, and (3) the variants do not affect the outcome directly, except by the exposure.

In the forward MR exploring the causal relationship between SS and cancer, all the single nucleotide polymorphisms (SNPs) from different studies are listed in Table S1. We selected genetic variants based on two thresholds for SS exposure: 5e−8 and 5e−6 (Table S2), with a minor allele frequency threshold of 0.01. The threshold of 5e−8 was mainly considered for discovery, and a lenient threshold of 5e−6 was used for validation [17]. Results showing consistent directionality and nominal significance ($P < 0.05$) at both genetic variant thresholds were considered significant. Thus, SNPs meeting the 5e−8 threshold with $P < 0.05$ were used as instruments, and heterogeneity and pleiotropy tests were performed (Table S3). For results showing significant heterogeneity, data were generated using a random-effects model [18].

In the backward MR analysis, which explored the relationship between cancer and SS, SNPs associated with cancer were selected using the same criteria as meaningful (thresholds of 5e−8 and 5e−6), and the MR results using 5e−8 for genetic variants were considered as the main results (Table S4). Results with the same directionality as the MR with $P < 0.05$, using 5e−6 genetic variants for validation, were considered positive (Table S5).

Steiger filtering was performed to ensure that the directionality of the SNPs ($P < 0.05$) was applied to the MR [19]. The inverse variance weighting method was primarily used to assess the causal effect between two phenotypes [20]. When only a single instrumental variable was available, the Wald ratio was used to evaluate the causal effect, and MR-PRESSO was applied to detect outliers [21]. The *F*-statistics of the instrumental variables were calculated by the formula $F = R^2(N - 2)/(1 - R^2)$, the R values were obtained using the "get_r_from_bsen" function in R, and only instruments with $F > 10$ were considered reliable for further analysis [22]. The *F*-statistics of all the instrumental variants ranged from 18.92 to 109.21 for forward MR, and from 20.858 to 455.996 for backward MR (Tables S2 and S4), indicating no significant weak instrument bias [23].

Causal relationships with significant *P*-values ($P < 0.05$) in both thresholds for exposure were considered significant.

Co-localization

Co-localization analysis between LUAD and SS was accomplished using R package coloc v5.2.3 within ± 150 kb of leading SNP of each trait with the default paraments. Results of PPH3 + PPH > 0.8 were deemed significant [24].

Identification and validation of the shared hub gene

SMR was used to infer causal genes using a single eQTL (Table S6) [15]. SMR analysis was performed on the data from over 19 250 known eQTL. Subsequently, *HLA-DPB2* was identified as a shared hub gene by intersection analysis and validation of its expression in bulk RNA-seq data.

The shared genes were validated by comparing their expression levels in the early and advanced disease stages using a *t*-test to further confirm the hub genes.

Bulk RNA-seq based validation and further exploration

Construction of a weighted gene co-expression network

The R package WGCNA was specifically utilized to perform gene co-expression network analysis of tumor tissues to further explore the function of *HLA-DPB2* and the genes co-expressed with it [25,26].

Enrichment analysis

GO enrichment analyses and KEGG pathway were used to identify the function of *HLA-DPB2* co-expression genes [27].

To explore the associations between genes in the green

gene module and *HLA-DPB2*, we analyzed the PPI networks of the 151 identified genes in the green module (Table S7) using the STRING database.

Correlation analysis and immune infiltration profiles

To explore the correlation between *HLA-DPB2* and its parental gene *HLA-DPB1*, Tumour Immune Estimation Resource (TIMER) was applied, and Spearman correlation analysis was used to determine relationships in the TCGA-LUAD data collection to verify the result of TIMER.

To explore the correlation of immune cells with *HLA-DPB1* and *HLA-DPB2*, CIBERSORT was used [28]. Spearman correlation analysis was conducted between *HLA-DPB1* and *HLA-DPB2* and the immune cells.

scRNA-seq based exploration

scRNA-seq data processing

Cells were filtered using gene expression counts below 500 and cells with more than 25% mitochondrial content. After removing low-quality cells, the selected single cells were normalized.

Dimension-reduction, cell clustering and annotation

Principal component analysis (PCA) was applied to reduce the dimensionality of the statistics utilizing the top 2000 most variable genes in the data set by using the "FindVariableFeatures" function in Seurat.

The Uniform Manifold Approximation and Projection (UMAP) was then applied to further reduce the dimensionality of the data set using the "RunPCA" function [29]. Single-cell data were downscaled using UMAP to project the cells onto a two-dimensional space and the cells were clustered using Seurat clusters.

The "SingleR" R package was used to annotate the cell clusters [30]. Highly and specifically expressed genes have been used as markers to identify cell types [31]. Marker genes for each cluster were identified using the "FindAllMarkers" function in Seurat.

scRNA-seq and analysis (pseudotime analysis)

Monocle (V2.30.1) was used to predict the pseudotime of each T cell to explore their differentiation trajectory of T cells [32] and to verify the reliability of T cell re-annotation [32].

Cell-cell interaction analysis

The R package "CellChat" (Version 1.6.1) was applied to infer the potential intercellular communication in the scRNA-seq data [33]. A threshold of 10 cells was used to filter communication.

Results

Forward and backward MR revealed the effects of SS on lung cancer and other cancers

In the set of instrumental variables ($P < 5e{-}8$), forward MR was performed to evaluate the causal effects of SS exposure and different types of cancers as outcomes. As shown in Figs. 2A, 2B and S1, of all the results, the SS is a protective factor in the development of breast cancer (odds ratio (OR) 0.95, 95% confidence intervals (CI) 0.94–0.96), ovarian cancer (OR 0.90, 95% CI 0.80–1.01), esophageal adenocarcinoma and endometrial cancer (OR 0.87, 95% CI 0.83–1.01), with ORs of less than 1. On the other hand, SS is a risk factor for lung cancer (OR 1.15, 95% CI 1.07–1.23), colorectal cancer (OR 1.08, 95% CI 1.04–1.12), malignant lymphoma, bile ducts and liver cancer, with an OR > 1.

Backward MR was used to explore the causal relationships between cancer and SS. Five independent SNPs associated with specific cancers were found to be significantly associated with SS (Fig. 2C and 2D). Meta-analysis of independent SNPs suggested that female genital cancer (OR 6.59, 95% CI 3.89–11.15) and lung cancer (OR 3.03, 95% CI 2.43–3.78) were risk factors for the SS (Fig. S2).

***HLA-DPB2* was identified as the hub gene between SS and lung cancer**

To further explore potential shared genes between SS and LUAD, we conducted a co-localization analysis, which identified a co-localization region near 6p21 with a PPH4 probability of 90.8%, which is considerably higher than 80% (Fig. 2E). This result indicates strong co-localization effects in this area [34]. Analysis of the eQTL data revealed that 10 genes were significantly correlated with these two traits. We identified 6 shared gene expressions in TCGA database, and their corresponding causal effects on SS and LUAD are presented in Fig. 2F. Because all tissues in the TCGA data set were obtained from patients, normal tissues could not be used as controls to validate the causality of lung cancer development. Therefore, we compared tissues from the early and advanced stages of lung cancer to assess progression and validate the 6 shared genes. *HLA-DPB2* was downregulated in the advanced stages of the disease (Fig. 2G), which was consistent with the SMR results (Fig. 2F). Besides, the *HLA-DPB2* is located on chromosome 6 (6p21.32), close to the co-localization region identified in the co-localization analysis. This proximity suggests that this gene is a potential hub linking SS and LUAD. Thus,

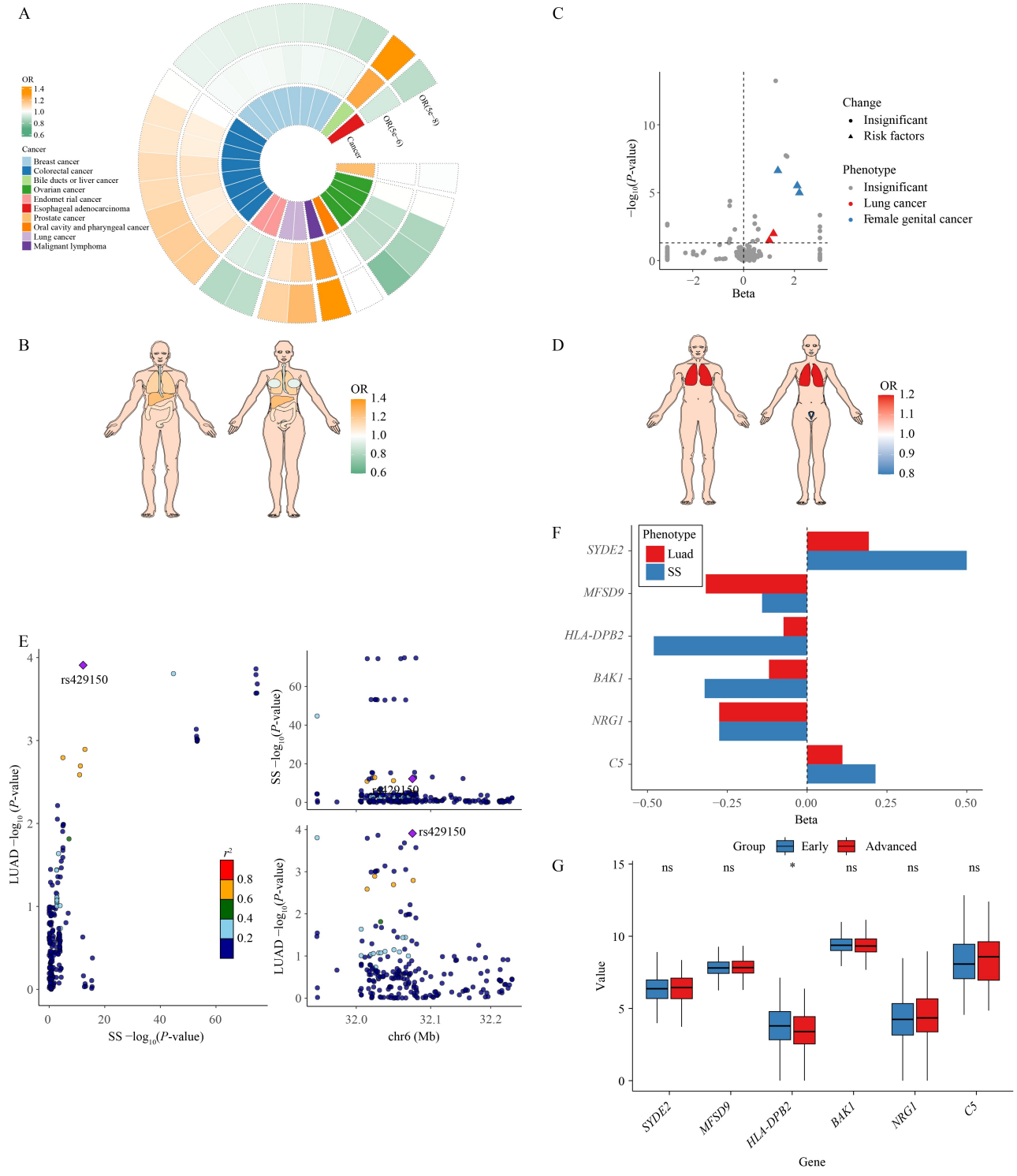

Fig. 2 MR revealed the interaction between lung cancer and SS. (A) Summary of 29 results of OR of SS on 10 types of cancers by forward MR. The smallest loop shows the type of cancer; the medium and largest loops present for the OR in threshold of 5e−6 and 5e−8, respectively. (B) Anatomy map of OR magnitude in specific organs of male (left) and female (right) patients. (C) The causality from cancer to SS identified in reverse MR. Circles represent causalities without causality and triangles represent causalities with causality; blue triangles indicate causalities associated with female genital cancer; and red triangles indicate causalities associated with lung cancer from reverse MR. (D) Shades of color are used to indicate OR magnitude in human anatomical drawings. (E) Co-localization analysis between SS and LUAD. (F) The bar plot of SMR: 6 of the 10 genes identified from the analysis of eQTL data from SS and LUAD using SMR were identified in the TCGA database, the bars represent the beta value in the SMR of each gene, a beta value > 0 suggests that the gene acts as a risk factor, while beta values < 0 indicate a protective effect. (G) Validation of the expression of 6 genes using TCGA.

HLA-DPB2 was considered as a hub gene for further exploration. After confirming directionality using Steiger's test, a bidirectional causality was found between LUAD and SS, suggesting a mutually reinforcing interaction between these two diseases.

HLA-DPB2, the pseudogene of *HLA-DPB1*, is highly related with immune function

To further explore the potential mechanisms of the hub genes, WGCNA and GO were used. WGCNA revealed 5 distinct modules, with the green module (Fig. 3A) comprising 172 genes showing the highest correlation with *HLA-DPB2* and a negative correlation with advanced staging (Fig. 3B; Table S7). *HLA-DPB2* is the pseudogene of *HL-DPB1*, it is a non-functional gene that contains sequences similar to functional gene *HLA-DPB1*.

GO enrichment analysis of the genes in the green module revealed predominant associations with immune-related pathways, particularly the MHC-II pathway (Fig. 3C). Similarly, KEGG pathway analysis demonstrated that the functions of the green module that has highest correlation with *HLA-DPB2* were strongly linked to immune processes, which were primarily enriched in immune-related diseases and pathways, including "Cell Adhesion Molecules" and "Antigen Processing and Presentation" pathways (Fig. 3D).

Previous studies have demonstrated that *HLA-DPB2* promotes *HLA-DPB1* expression in breast cancer, thereby exerting an antitumor effect [35]. In the present study, we also confirmed that *HLA-DPB2* promotes the expression of *HLA-DP1,* as evidenced by a strong correlation between *HLA-DPB1* and *HLA-DPB2* in both the TIMER and TCGA database (rho = 0.627, P = 1.29e−57, Fig. S3A; R = 0.612, P < 2.2e−16, Fig. S3B).

Given that the functions of the green module are closely related to immunity, we explored the correlations of *HLA-DPB1* and *HLA-DPB2* with immune cells. Both genes positively correlated with macrophages, memory B cells (MemB), resting dendritic cells, monocytes, and $CD8^+$ T cells (Fig. S3C). In addition, PPI analysis of the green module genes showed that the seven core genes of the module related to *HLA-DPB2*, including MNDA [36] and TLR8 [37] were highly correlated with immunity

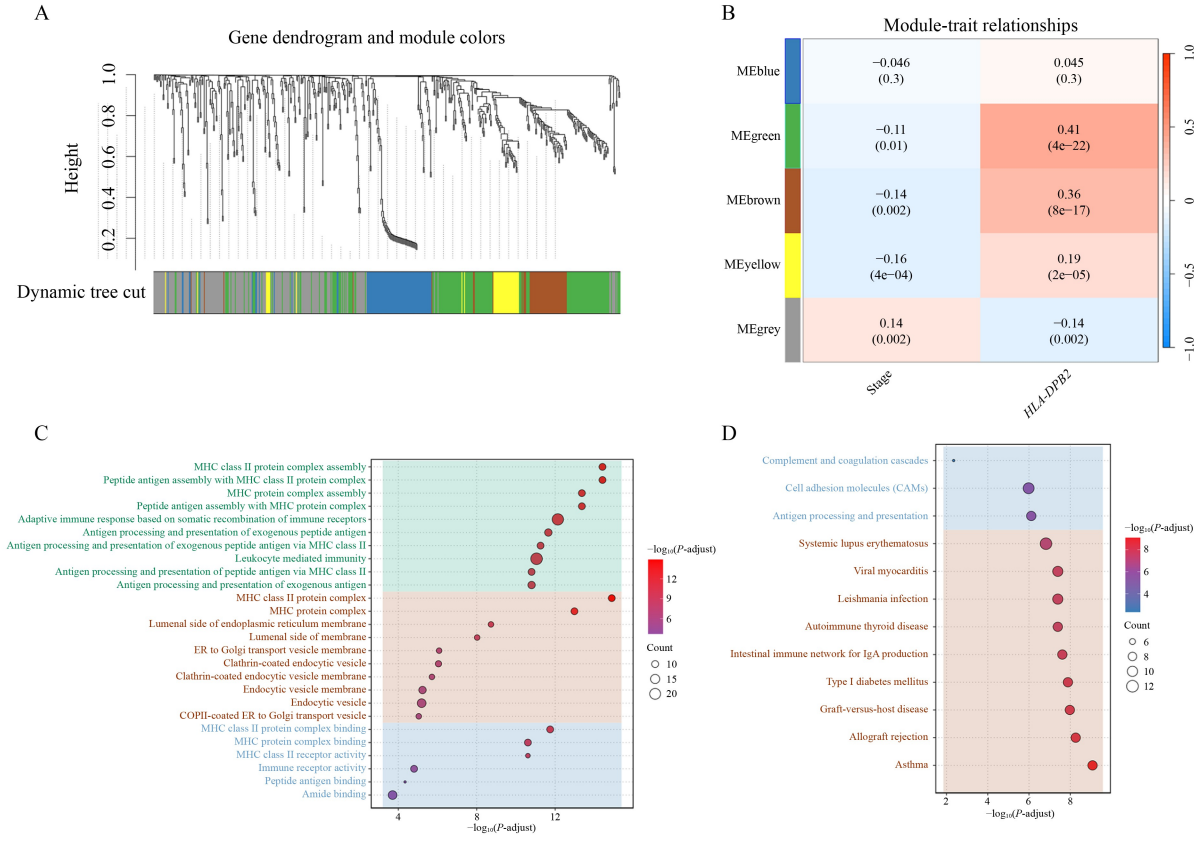

Fig. 3 Gene co-expression of *HLA-DPB2* and functional enrichment analysis in LUAD. (A) In LUAD, hierarchical clustering dendrograms illustrate the co-expressed genes organized into modules. Each row with a distinct color corresponds to a module, representing a cluster of tightly interconnected genes. (B) Heatmap illustrating the correlation among module eigengenes, stages, and *HLA-DPB2* expression of LUAD. (C) The green module GO enrichment results. Biological process (BP) is shown in green, cellular component (CC) in brown, and molecular function (MF) in blue. (D) In the figure, pathways related to other immune diseases are shown in brown, and immune-related pathways in blue.

(Fig. S4). Given that *HLA-DPB2* is a pseudogene strongly correlated with *HLA-DPB1*, but direct expression data for *HLA-DPB2* are generally unavailable in single-cell data sets, we hypothesized that proteins strongly interacting with *HLA-DPB1* may also be relevant to *HLA-DPB2*. Therefore, we analyzed *HLA-DPB1* in single-cell tumor data set as a proxy to infer the potential mechanisms by which *HLA-DPB2* might suppress tumorigenesis and progression.

HLA-DPB1 expression is decreased in tumor tissues

First of all, we demonstrated a close association between *HLA-DPB2* and *HLA-DPB1* in the transcriptome. To avoid the potential limitations of using a single method, we performed co-localization analysis using eQTL data for *HLA-DPB2* and *HLA-DPB1*, which revealed a strong co-localization effect between these two genes (PPH3 + PPH4 = 1.00 > 0.8, Fig. S3D). Using scRNA-seq analysis, we first validated *HLA-DPB1* expression in the immune cells of normal and tumor tissues. The results showed a significant decline in *HLA-DPB1* expression in immune cells from the tumor tissues (Fig. 4A). Next, we explored immune cells in early and advanced tumor tissues according to our bulk RNA-seq results. To ensure reliable results, the distribution of immune cells was explored at different cancer stages using UMAP (Fig. S5).

HLA-DPB1 is predominantly expressed in B cells, macrophages and monocytes

To identify the primary immune cell types expressing *HLA-DPB1*, we pre-processed the scRNA-seq data set GSE131907 of lung cancer immune cells using stringent quality control metrics [29] and visualized them using the UMAP method. We categorised the cells into 25 cell subpopulations using Seurat clusters (Fig. S6) and annotated them into 10 cell types using the singleR package (Figs. 4B and S7). UMAP plots and differential expression analyses revealed that *HLA-DPB1* was highly expressed in B cell, macrophages, and monocytes, suggesting that it plays a significant role in these immune cells (Figs. 4C and S8).

HLA-DPB1 expression is downregulated in MemB with cancer progression

The results of the SMR and downregulated expression in lung cancer patients compared to normal individuals (Fig. 4A) suggest that *HLA-DPB1* acts as a protective factor. *HLA-DPB1* was higher in B cells in the early stage of cancer than in the late stage but showed no difference in macrophages and monocytes (Fig. 5A). To refine the B cell analysis, we re-annotated them into two subsets, MemB and germinal center (GC) B cells (Figs. 5B and S9). To explore B cell subpopulations with differential *HLA-DPB1* expression, we performed differential analysis in GC B cells and MemB and found that *HLA-DPB1* was only different in MemB (Fig. 5C). The GSEA of MemB indicated the upregulation of three immune-related pathways in the progression of LUAD, including the intrinsic component of the plasma membrane, side of the membrane, and antigen binding (Fig. S10).

Significant alterations in the MHC-II signaling pathway during the progression of lung cancer

In the immune system, anti-tumor effects mainly rely on T cells [38]. In the scRNA analysis, since SingleR annotation did not differentiate T cell subpopulations such as depleted T cells, we conducted a detailed subpopulation delineation and annotation of T cells (Figs. S11–S13) and validated it by pseudotime analysis (Fig. S14A and S14B).

To investigate the differences in the interactions

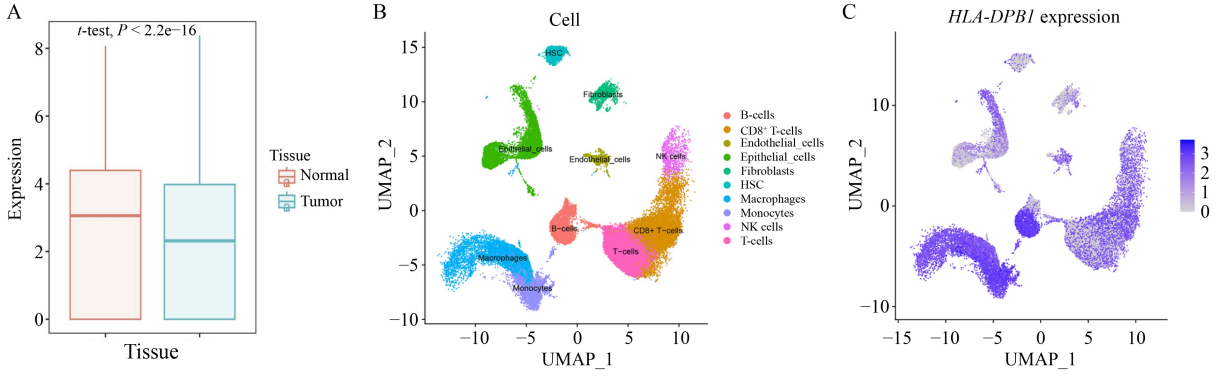

Fig. 4 Identification of cell clusters and annotation. (A) *HLA-DPB1* expression in different cancer stages. (B) UMAP plot of lung cancer immune cell scRNA-seq data set GSE131907. The 25 cell clusters are labeled, and 10 cells are annotated using the singleR R package. (C) Expression of the *HLA-DPB1* gene on the UMAP plot. Darker colors indicate higher levels of *HLA-DPB1* expression.

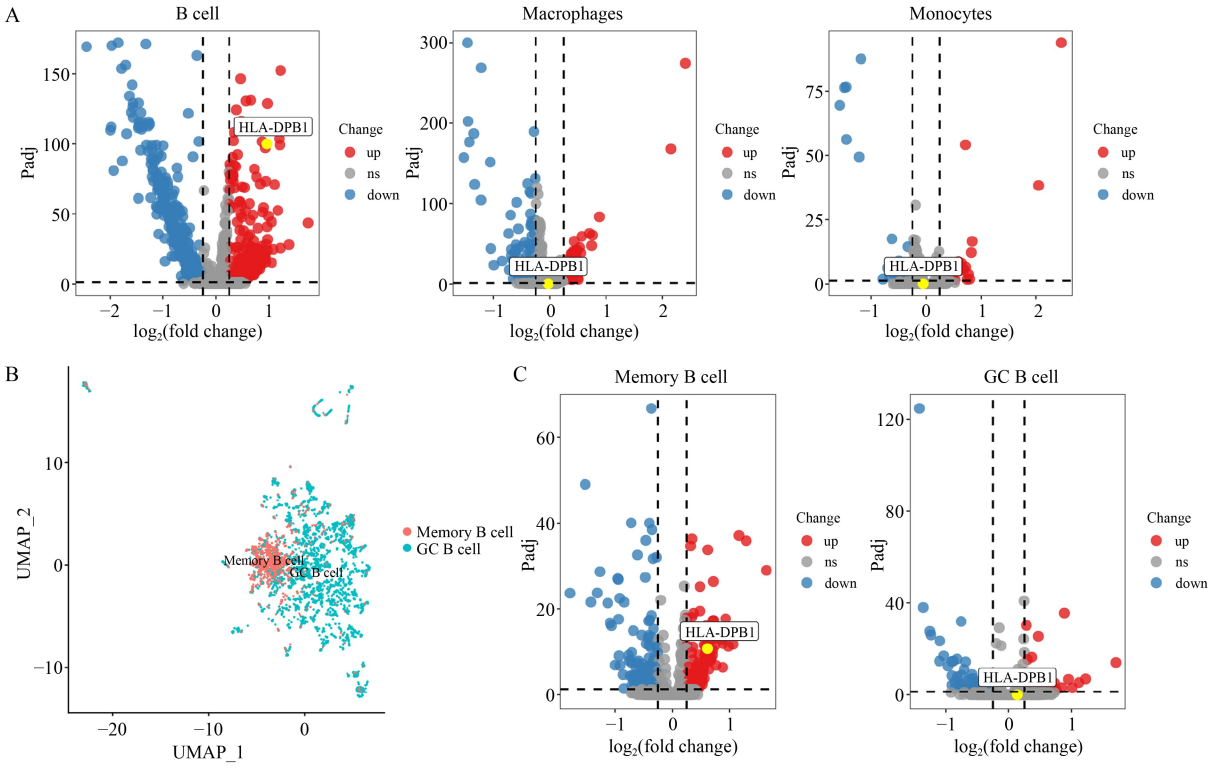

Fig. 5 Cells differential expression analysis and B cells reannotated. (A and C) Volcano plot of immune cells. *HLA-DPB1* is highlighted in yellow. (B) UMAP plot of B cells subsets. UMAP, uniform manifold approximation and projection.

between MemB and other immune cells in the early and advanced stages, we analyzed cellular interactions using CellChat and found that MemB primarily interacts with effector T cells, macrophages, and monocytes (Fig. 6A), with stronger communication observed at early cancer stages (Fig. 6B and 6C). Ligand-receptor pairs from MemB to effector T cells, macrophages, and monocytes, including MIF, MHC-II, ICAM, CD99, ANNEXIN, IL16, CLEC, and UGRP1, were significantly downregulated during the progression of lung cancer (Fig. 6D). Interactions based on the MHC-II signaling pathway between MemB and exhausted T cells, such as macrophages, monocytes, and B GC cells, increased (Fig. 6E). Since the effector T cells also presented obvious interactions with macrophages and monocytes, we investigated their communication and found that the GALECTIN, CLEC, CD99, MIF, RESISTIN, and ALCAM pairs showed statistically significant variation at different stages (Fig. S15). These scRNA-based findings suggest that elevated MHC-II-mediated communication with immune cells is a key pathway through which *HLA-DPB2* contributes to LUAD progression.

Validation of *HLA-DPB2* expression in pSS patients

Having established that *HLA-DPB2* is significantly associated with lung cancer progression and inferred (through analysis of its parental gene, *HLA-DPB1*) that MemB are likely to mediate this involvement, we subsequently validated these findings using pSS patient data sets. Given that pSS is a systemic autoimmune disorder, we used peripheral blood data sets to validate the association between *HLA-DPB2* expression and the disease pathology. We found a correlation between *HLA-DPB2* expression levels and both clinical disease severity (GSE66795, Fig. 7A) and progression (GSE173670, Fig. 7B).

Elevated *HLA-DPB1* in MemB correlates with immune suppression in healthy individuals

To elucidate the role of *HLA-DPB2* in B cells and investigate the potential mechanisms for suppressing pSS pathogenesis, we analyzed the GSE253568 single-cell data set. Similar to the lung cancer scRNA-seq data, *HLA-DPB2* expression was not directly detectable in this data set. Therefore, we used the parental gene *HLA-DPB1* as a proxy. Given the absence of detailed clinical severity data for this cohort, we performed the analyses using only healthy controls. These samples were stratified into high- and low-expression groups based on B cell *HLA-DPB1* levels (Fig. 7C) to identify the functional differences that may explain *HLA-DPB1*'s putative protective effects against pSS development. Differential expression analysis

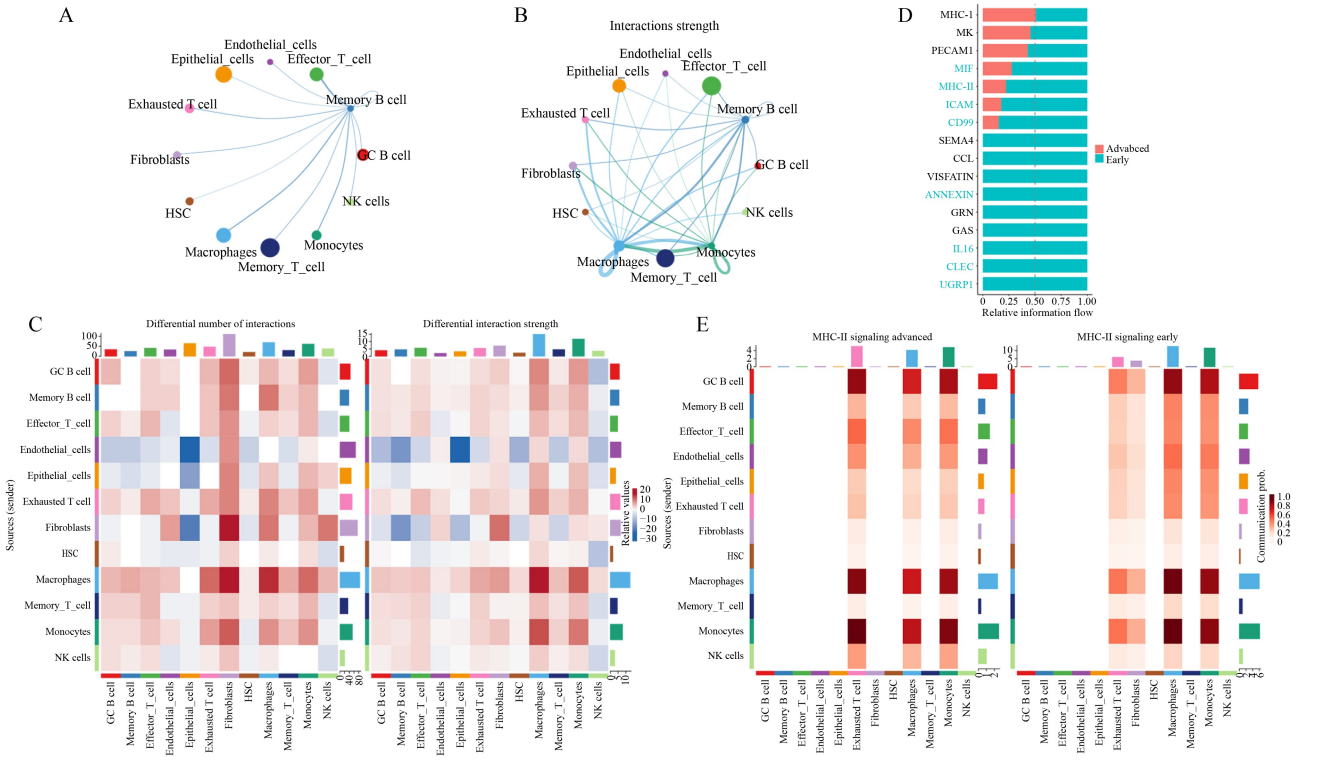

Fig. 6 Intercellular communication among immune cells in the early- and advanced-stages of lung cancer. (A) Ligand-receptor signaling pathway networks originating from B cells in a complete sample set. The line thickness refers to the strength of ligand-receptor pairs. (B) Changes of signaling pathway originating from B cells, macrophages and monocytes by contrasting early stage to advanced stage samples. The line thickness refers to the variation of strength of ligand-receptor pairs. (C) Heatmap showing specific differences in the number (left) and strength (right) of cell-cell interactions in the early and advanced stages. (D) Variations of MemB to macrophages and monocytes of classic ligand-receptor pairs by contrasting the strength in the early and advanced stages. The names of ligand-receptor pairs in blue are statistically significant. (E) Heatmap showing variations in the strength of MHC-II based cellular communications among immune cells in the early (right) and advanced (left) stages.

revealed significant transcriptional differences between *HLA-DPB1* high and *HLA-DPB1* low MemB, with 79 upregulated and 107 downregulated genes in the high-expression group (false discovery rate (FDR) < 0.05, $|\log_2 FC| > 1$) (Fig. 7D). GO enrichment analysis of genes downregulated in *HLA-DPB1* high MemB revealed significant suppression of pro-inflammatory pathways (FDR < 0.05), including the immune response-regulating signaling pathway, immune response-regulating cell surface receptor signaling pathway, and immune response-activating signaling pathway (Fig. 7E, Table S10). KEGG pathway analysis corroborated the GO enrichment results, demonstrating the coordinated downregulation of immune-related and cancer-associated pathways in *HLA-DPB1* high MemB (Fig. 7F, Table S11). Functional enrichment analysis of the 79 upregulated genes yielded limited results, with only two mitochondria-related pathways identified in the GO analysis (Table S12). No significant pathways were detected in the KEGG analysis (FDR > 0.1). These findings demonstrate that decreased *HLA-DPB2* expression is significantly associated with pSS disease progression, corroborating the MR results. Mechanistically, high expression of its parental gene, *HLA-DPB1*, in healthy B memory cells was found to suppress multiple pathogenic pathways, including pro-inflammatory responses, immune activation, and tumor-associated processes. These results suggest that the *HLA-DPB2/DPB1* axis confers protection against both pSS and lung cancer through the coordinated downregulation of shared pathological pathways.

Discussion

In our study, we employed bidirectional MR to investigate the potential reciprocal promotion effect between SS and cancer, uncovering a mutually reinforcing causal relationship between SS and lung cancer. Additionally, MR can prevent the influence of potential bias in previous studies on the correlation between SS and malignancy [39]. Through SMR analysis of the eQTL data, we further identified *HLA-DPB2* as a putative gene, and subsequent multi-omics investigations revealed its dual protective role in both lung cancer and

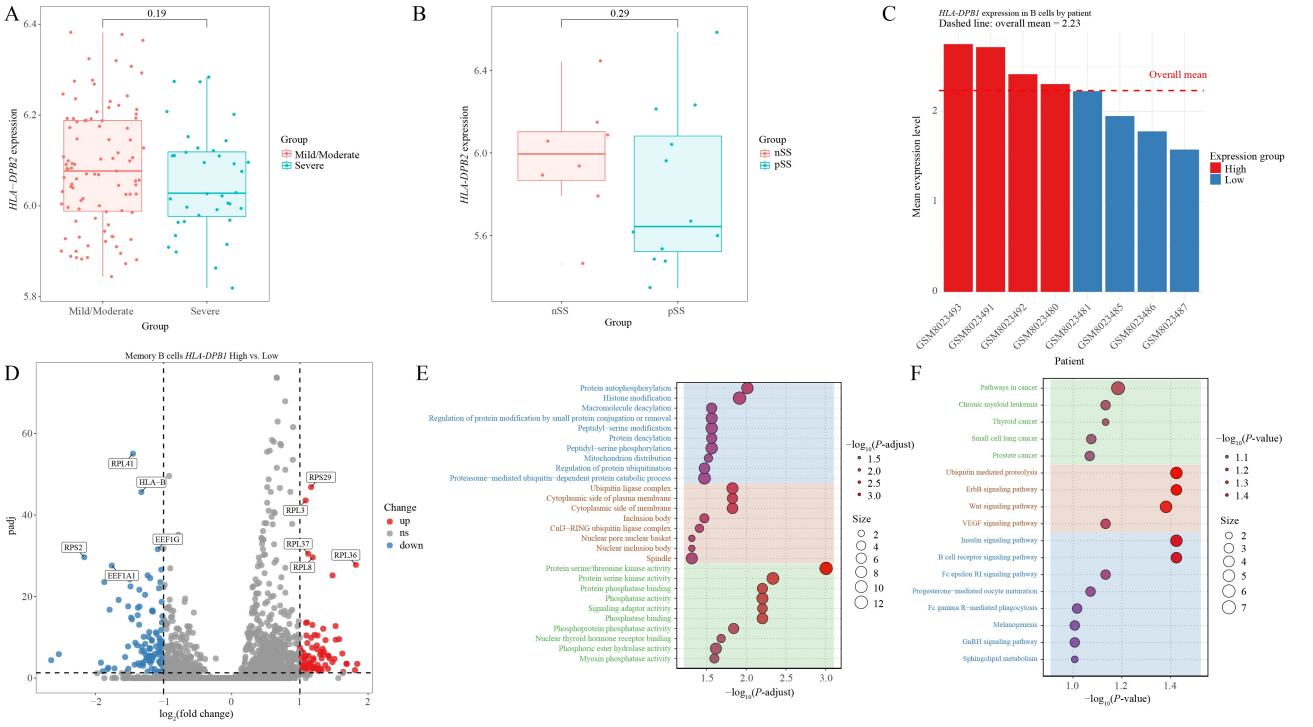

Fig. 7 Validation of *HLA-DPB2* expression in patients with pSS and elevated *HLA-DPB1* in MemB correlates with immune suppression in healthy individuals. (A) Box plot (with *P*-values from *t*-tests) comparing *HLA-DPB2* expression between mild/moderate and severe pSS patients in the GSE66795 bulk transcriptomic data set. (B) Box plot demonstrating *HLA-DPB2* expression differences across SS disease progression stages (nSS vs. pSS) in the GSE173670 transcriptomic data set, compared using a two-tailed *t*-test. (C) Stratification of healthy controls from GSE253568 scRNA-seq data by B cell *HLA-DPB1* expression levels (high vs. low), visualized as stacked bar charts. (D) Volcano plot of DEGs between *HLA-DPB1* high and *HLA-DPB1* low MemB groups in healthy controls (GSE253568). (E) GO enrichment of downregulated DEGs from Panel D, displaying top terms: biological process (blue), molecular function (brown), cellular component (green). (F) KEGG pathway analysis of downregulated DEGs. Tumour-related pathways (green), signal processing (brown), and organismal systems (blue). DEGs, differentially expressed genes; GO, gene ontology; KEGG, Kyoto Encyclopaedia of Genes and Genomes; nSS, non-Sjögren's syndrome; pSS, primary Sjögren's syndrome; scRNA-seq, single-cell RNA sequencing; SS, Sjögren's syndrome.

pSS. In lung cancer, transcriptomic analysis demonstrated a strong correlation between *HLA-DPB2* and immune function, and its involvement in immune response pathways, whereas single-cell RNA sequencing-based analyses indicated that *HLA-DPB2* predominantly exerted its effects on MemB, suggesting a potential mechanism by which these cells modulate tumorigenesis and tumor progression via macrophage/monocyte-dependent regulation of T cell-MHC-II interactions. Parallel studies in patients with pSS revealed an inverse correlation between *HLA-DPB2* expression and disease severity, and functional validation showed that high expression of its parental gene, *HLA-DPB1*, in B cells significantly downregulated pro-inflammatory, immune activation, and oncogenic signaling pathways. This study established a causal relationship between SS and lung cancer, and based on this, we identified the *HLA-DPB2*/DPB1 axis as a conserved immunoregulatory mechanism.

In our study, we identified a consistent causal relationship between SS and lung cancer using both forward and reverse MR analyses. This finding was further supported by co-localization analysis, suggesting a potential reciprocal relationship between SS and lung cancer. Several studies support this association. Studies have shown that up to 20% of SS patients exhibit lung involvement that can be found in imaging [40,41]. Another study revealed an incidence of lung cancer of 0.477% in patients with pSS, which is higher than that in the normal population (91.36 and 58.18 cases per 100 000 individuals for male and female individuals, respectively) [42]. Adenocarcinoma was identified as the most common subtype [43]. However, these findings were limited by the small sample size (10 patients with SS and lung cancer) and the lack of a large lung cancer control group, which may have introduced bias. However, conflicting perspectives exist on this topic. A two-sample MR study revealed no significant causal relationship between SS and lung cancer (including LUAD and squamous cell lung cancer) [6]. However, this study relied solely on GWAS data, which may lead to bias owing to a lack of validation. Moreover, the SNPs identified in this study did not completely represent variants of SS [6]. Differences in data sources and lung cancer subtypes between their study and ours may explain

these discrepancies.

As our study revealed a possible mutually reinforcing causal relationship between SS and lung cancer, we combined eQTL, transcriptomic, and single-cell data analyses to investigate the cellular and molecular mechanisms, which further revealed the *HLA-DPB2/DPB1* axis as a key regulatory mechanism and the role of MemB. We identified *HLA-DPB2* as a hub gene influencing the development of SS and LUAD, and its expression decreased in advanced stages of lung cancer, as validated by single-cell-based analysis. Although *HLA-DPB2* is a pseudogene that does not encode proteins, studies have shown that pseudogenes can act as long noncoding RNAs to regulate parent or unrelated protein-coding genes, influencing tumor development by acting as microRNA decoys [44]; moreover, the study revealed a robust association between the *HLA-DPB2/DPB1* axis [35], and we proved the robust associations between the *HLA-DPB2*/*DPB1* axis by correlation analysis and co-localization analysis, thereby supporting the rationale of using *HLA-DPB1* as a proxy to investigate *HLA-DPB2*'s biological functions. Several studies have shown an association between *HLA-DPB2* and various cancers including cervical [45], breast [35], ovarian [46] and rectal cancers [47]. In breast cancer, expression of the *HLA-DPB2*/*DPB1* axis is closely associated with T cells during disease progression, indicating an anti-tumor effect through immune cell recruitment, in accordance with our results [35]. Previous studies have shown that HLA class II alleles can increase the risk of SS in HLA-DR by promoting B-lymphocyte survival and activation, whereas genes on HLA-DQ can also influence progression. However, the role of HLA-DP in SS development remains unclear [10,11,48]. Notably, although *HLA-DPB2* expression has been implicated in the severity of rheumatoid arthritis [49], its potential causal role in the development of that condition remains unclear. Similarly, the effect of *HLA-DPB2* expression on SS susceptibility and progression represents a significant knowledge gap. Our study provides the first evidence that *HLA-DPB2* expression may serve as a protective factor against pSS and lung cancer development, with higher expression levels correlating with less severe disease progression under both conditions. Mechanistically, we propose that this protective effect is mediated by B cells, particularly the MemB population. However, further investigation at the genetic and molecular levels is required.

In our study, the expression of *HLA-DPB1* was predominantly found in MemB and was higher in early samples. Immune-related pathways were significantly upregulated in the early stage compared to advanced stages, suggesting that pathways associated with *HLA-DPB2* may influence lung cancer development through MemB. By combining genomic, transcriptomic, and single-cell analyses, we provided evidence for the role of the *HLA-DPB2*/*DPB1* axis in lung cancer development. This is consistent with the findings of previous studies. A multi-omics study showed that the immune infiltration of MemB differed between normal and tumor tissues, with MemB being highly enriched in tumor tissues [50,51] and showing a significant increase in patients with LUAD [52]. Another study focusing on MemB subtypes found that increased CD27 expression in switched MemB and $IgD^{+}CD24^{+}$ B cells may be associated with the development of lung cancer [53]. However, the roles of MemBs in lung cancer remain unclear. A study using a metagenomic approach (CIBERSORT) suggested that the lack of MemB is associated with poor prognosis in early clinical LUAD, often accompanied by an increase in the number of macrophages [54]. Lung cancers with abundant MemB infiltration respond better to anti-PD-1 therapy [55]. In addition, MemB was correlated with a positive treatment outcome following neoadjuvant chemoimmunotherapy [56], which was positively associated with a low risk of developing tumors [57]. This study provides important preliminary evidence that implicates MHC-II-related pathways and MemB in tumor progression. However, validation in patients with SS remains limited owing to insufficient clinical progression data in the available SS samples. Future studies are needed to investigate the molecular mechanisms of MHC-II pathways in lung cancer pathogenesis and to conduct large-scale sequencing studies with well-documented disease progression metrics in SS cohorts.

Our study had several limitations. First, the GWAS data used in the bidirectional MR analysis were primarily derived from individuals of European descent, limiting their generalisability to other ethnic groups. Additionally, confounding factors such as age and sex, as well as other environmental variables, exerted a certain influence on the MR analysis. Future studies with larger sample sizes and diverse populations are needed. Second, the association between *HLA-DPB2* expression and pSS severity in the transcriptomic data sets did not reach statistical significance, possibly because of insufficient sample size. Larger pSS cohorts with standardised clinical phenotypes are required to clarify this relationship. Third, although our single-cell analysis of *HLA-DPB1* in healthy donor B cells provided mechanistic insights, future studies should examine pSS patient-derived samples with well-characterized disease severity to confirm their translational relevance. Finally, our study was conducted based on existing research data, and further *in vivo* and *in vitro* experiments are necessary to confirm the correlation between SS and lung cancer and to clarify the functions of *HLA-DPB2* and MemB in disease progression.

To our knowledge, this is the first study that integrates Mendelian randomization with single-cell and

transcriptome analyses to investigate the causal relationship between SS and malignancy and the potential underlying mechanisms. Our study revealed a mutually reinforcing causal relationship between SS and lung cancer, with *HLA-DPB2* playing a key role. Single-cell analysis further revealed that the hub gene may affect MemB, influencing tumor occurrence and development through MHC-II ligands on monocytes, macrophages, and effector T cells. Parallel pSS single-cell data confirmed the MemB-mediated immunomodulation of shared pathogenic pathways. These findings collectively establish the *HLA-DPB2*/*DPB1* axis as a novel protective mechanism in both pSS and lung cancer pathogenesis, mediated through the MemB-dependent regulation of immune homeostasis.

Acknowledgements

We would like to thank the China Postdoctoral Science Foundation (No. 2023M742488), Sichuan Provincial Natural Science Fund (No. 24NSFSC6690), Postdoctoral Fund of West China Hospital (No. 2023HXBH004), and the "From 0 to 1" Innovative Research Project of Sichuan University (No. 2023SCUH0031). We thank the individuals/organizations that made the databases publicly available for research.

Compliance with ethics guidelines

Conflicts of interest Kai Xu, Manhua Wang, Zixuan Yang, Yu Tang, Zhen Li, Tao Liu, Yu Wang, Yuqing Wang, and Xiaoqian Zhai declare that they have no conflicts of interest.

The requirement for ethics approval was waived because the data were obtained from open access databases.

Data availability and compliance statement

The authors declare that the acquisition and subsequent use of all data presented in this manuscript comply fully with all relevant local, national, and international laws, regulations, ethical guidelines, and the terms of use associated with the original data sources.

The authors bear full legal responsibility for ensuring the legality of data acquisition and all subsequent uses.

All the GWAS summary statistical data used in this study are publicly available. SS and LUSC data sets were obtained from the IEU OpenGWAS project.

All the software and packages used are publicly available. The codes used in this study can be found at the websites of TwoSampleMR, coloc, and SingleR.

Electronic supplementary material Supplementary material is available in the online version of this article at https://doi.org/10.1007/s11684-025-1165-z and is accessible for authorized users.

References

1. Fox RI. Sjögren's syndrome. Lancet 2005; 366(9482): 321–331
2. Brito-Zerón P, Baldini C, Bootsma H, Bowman SJ, Jonsson R, Mariette X, Sivils K, Theander E, Tzioufas A, Ramos-Casals M. Sjögren syndrome. Nat Rev Dis Primers 2016; 2(1): 16047
3. Patel R, Shahane A. The epidemiology of Sjögren's syndrome. Clin Epidemiol 2014; 6: 247–255
4. Liang Y, Yang Z, Qin B, Zhong R. Primary Sjogren's syndrome and malignancy risk: a systematic review and meta-analysis. Ann Rheum Dis 2014; 73(6): 1151–1156
5. Zhong H, Liu S, Wang Y, Xu D, Li M, Zhao Y, Zeng X. Primary Sjögren's syndrome is associated with increased risk of malignancies besides lymphoma: a systematic review and meta-analysis. Autoimmun Rev 2022; 21(5): 103084
6. Jia Y, Yao P, Li J, Wei X, Liu X, Wu H, Wang W, Feng C, Li C, Zhang Y, Cai Y, Zhang S, Ma X. Causal associations of Sjögren's syndrome with cancers: a two-sample Mendelian randomization study. Arthritis Res Ther 2023; 25(1): 171
7. Goulabchand R, Malafaye N, Jacot W, Witkowski Durand Viel P, Morel J, Lukas C, Rozier P, Lamure S, Noel D, Molinari N, Mura T, Guilpain P. Cancer incidence in primary Sjögren's syndrome: data from the French hospitalization database. Autoimmun Rev 2021; 20(12): 102987
8. Weng MY, Huang YT, Liu MF, Lu TH. Incidence of cancer in a nationwide population cohort of 7852 patients with primary Sjogren's syndrome in Taiwan. Ann Rheum Dis 2012; 71(4): 524–527
9. Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. Nat Rev Immunol 2018; 18(5): 325–339
10. Cruz-Tapias P, Rojas-Villarraga A, Maier-Moore S, Anaya JM. HLA and Sjögren's syndrome susceptibility. A meta-analysis of worldwide studies. Autoimmun Rev 2012; 11(4): 281–287
11. Trutschel D, Bost P, Mariette X, Bondet V, Llibre A, Posseme C, Charbit B, Thorball CW, Jonsson R, Lessard CJ, Felten R, Ng WF, Chatenoud L, Dumortier H, Sibilia J, Fellay J, Brokstad KA, Appel S, Tarn JR, Quintana-Murci L, Mingueneau M, Meyer N, Duffy D, Schwikowski B, Gottenberg JE. Variability of primary Sjögren's syndrome is driven by interferon-α and interferon-α blood levels

are associated with the class II HLA-DQ locus. Arthritis Rheumatol 2022; 74(12): 1991–2002
12. Miceli-Richard C, Wang-Renault SF, Boudaoud S, Busato F, Lallemand C, Bethune K, Belkhir R, Nocturne G, Mariette X, Tost J. Overlap between differentially methylated DNA regions in blood B lymphocytes and genetic at-risk loci in primary Sjögren's syndrome. Ann Rheum Dis 2016; 75(5): 933–940
13. Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. JAMA 2017; 318(19): 1925–1926
14. Boef AG, Dekkers OM, le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. Int J Epidemiol 2015; 44(2): 496–511
15. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 2016; 48(5): 481–487
16. Wu Y, Zeng J, Zhang F, Zhu Z, Qi T, Zheng Z, Lloyd-Jones LR, Marioni RE, Martin NG, Montgomery GW, Deary IJ, Wray NR, Visscher PM, McRae AF, Yang J. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun 2018; 9(1): 918
17. Xiao L, Liu S, Wu Y, Huang Y, Tao S, Liu Y, Tang Y, Xie M, Ma Q, Yin Y, Dai M, Zhang M, Llamocca E, Gui H, Wang Q. The interactions between host genome and gut microbiome increase the risk of psychiatric disorders: Mendelian randomization and biological annotation. Brain Behav Immun 2023; 113: 389–400
18. van Aert RCM, Schmid CH, Svensson D, Jackson D. Study specific prediction intervals for random-effects meta-analysis: a tutorial: prediction intervals in meta-analysis. Res Synth Methods 2021; 12(4): 429–447
19. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet 2017; 13(11): e1007081
20. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, Palmer T, Schooling CM, Wallace C, Zhao Q, Smith GD. Mendelian randomization. Nat Rev Methods Primers 2022; 2: 6
21. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018; 50(5): 693–698
22. Li B, Martin EB. An approximation to the F distribution using the chi-square distribution. Comput Stat Data Anal 2002; 40(1): 21–26
23. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol 2016; 45(6): 1961–1974
24. Su WM, Gu XJ, Dou M, Duan QQ, Jiang Z, Yin KF, Cai WC, Cao B, Wang Y, Chen YP. Systematic druggable genome-wide Mendelian randomisation identifies therapeutic targets for Alzheimer's disease. J Neurol Neurosurg Psychiatry 2023; 94(11): 954–961
25. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008; 9(1): 559
26. Langfelder P, Horvath S. Fast R functions for robust correlations and hierarchical clustering. J Stat Softw 2012; 46(11): i11
27. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28(1): 27–30
28. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015; 12(5): 453–457
29. Yang S, Guo J, Kong Z, Deng M, Da J, Lin X, Peng S, Fu J, Luo T, Ma J, Yin H, Liu L, Liu J, Zha Y, Tan Y, Zhang J. Causal effects of gut microbiota on sepsis and sepsis-related death: insights from genome-wide Mendelian randomization, single-cell RNA, bulk RNA sequencing, and network pharmacology. J Transl Med 2024; 22(1): 10
30. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, Butte AJ, Bhattacharya M. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 2019; 20(2): 163–172
31. He M, He Q, Cai X, Liu J, Deng H, Li F, Zhong R, Lu Y, Peng H, Wu X, Chen Z, Lao S, Li C, Li J, He J, Liang W. Intratumoral tertiary lymphoid structure (TLS) maturation is influenced by draining lymph nodes of lung cancer. J Immunother Cancer 2023; 11(4): e005539
32. Fan F, Gao J, Zhao Y, Wang J, Meng L, Ma J, Li T, Han H, Lai J, Gao Z, Li X, Guo R, Cao Z, Zhang Y, Zhang X, Chen H. Elevated mast cell abundance is associated with enrichment of $CCR2^+$ cytotoxic T cells and favorable prognosis in lung adenocarcinoma. Cancer Res 2023; 83(16): 2690–2703
33. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using CellChat. Nat Commun 2021; 12(1): 1088
34. Ou YN, Yang YX, Deng YT, Zhang C, Hu H, Wu BS, Liu Y, Wang YJ, Zhu Y, Suckling J, Tan L, Yu JT. Identification of novel drug targets for Alzheimer's disease by integrating genetics and proteomes from brain and blood. Mol Psychiatry 2021; 26(10): 6065–6073
35. Lyu L, Yao J, Wang M, Zheng Y, Xu P, Wang S, Zhang D, Deng Y, Wu Y, Yang S, Lyu J, Guan F, Dai Z. Overexpressed pseudogene HLA-DPB2 promotes tumor immune infiltrates by regulating HLA-DPB1 and indicates a better prognosis in breast cancer. Front Oncol 2020; 10: 1245
36. Bottardi S, Layne T, Ramòn AC, Quansah N, Wurtele H, Affar EB, Milot E. MNDA, a PYHIN factor involved in transcriptional regulation and apoptosis control in leukocytes. Front Immunol 2024; 15: 1395035
37. Gane EJ, Dunbar PR, Brooks AE, Zhang F, Chen D, Wallin JJ, van Buuren N, Arora P, Fletcher SP, Tan SK, Yang JC, Gaggar A, Kottilil S, Tang L. Safety and efficacy of the oral TLR8 agonist selgantolimod in individuals with chronic hepatitis B under viral suppression. J Hepatol 2023; 78(3): 513–523
38. Tay C, Tanaka A, Sakaguchi S. Tumor-infiltrating regulatory T cells as targets of cancer immunotherapy. Cancer Cell 2023; 41(3): 450–465
39. Larsson SC, Butterworth AS, Burgess S. Mendelian randomization for cardiovascular diseases: principles and applications. Eur Heart J 2023; 44(47): 4913–4924
40. Papiris SA, Maniati M, Constantopoulos SH, Roussos C, Moutsopoulos HM, Skopouli FN. Lung involvement in primary Sjögren's syndrome is mainly related to the small airway disease.

Ann Rheum Dis 1999; 58(1): 61–64
41. Yoo H, Hino T, Hwang J, Franks TJ, Han J, Im Y, Lee HY, Chung MP, Hatabu H, Lee KS. Connective tissue disease-related interstitial lung disease (CTD-ILD) and interstitial lung abnormality (ILA): evolving concept of CT findings, pathology and management. Eur J Radiol Open 2022; 9: 100419
42. Han B, Zheng R, Zeng H, Wang S, Sun K, Chen R, Li L, Wei W, He J. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024; 4(1): 47–53
43. Xu Y, Fei Y, Zhong W, Zhang L, Zhao J, Li L, Wang M. The prevalence and clinical characteristics of primary Sjogren's syndrome patients with lung cancer: an analysis of ten cases in China and literature review. Thorac Cancer 2015; 6(4): 475–479
44. Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DR. Pseudogenes: pseudo-functional or key regulators in health and disease? RNA 2011; 17(5): 792–798
45. Shi Y, Li L, Hu Z, Li S, Wang S, Liu J, Wu C, He L, Zhou J, Li Z, Hu T, Chen Y, Jia Y, Wang S, Wu L, Cheng X, Yang Z, Yang R, Li X, Huang K, Zhang Q, Zhou H, Tang F, Chen Z, Shen J, Jiang J, Ding H, Xing H, Zhang S, Qu P, Song X, Lin Z, Deng D, Xi L, Lv W, Han X, Tao G, Yan L, Han Z, Li Z, Miao X, Pan S, Shen Y, Wang H, Liu D, Gong E, Li Z, Zhou L, Luan X, Wang C, Song Q, Wu S, Xu H, Shen J, Qiang F, Ma G, Liu L, Chen X, Liu J, Wu J, Shen Y, Wen Y, Chu M, Yu J, Hu X, Fan Y, He H, Jiang Y, Lei Z, Liu C, Chen J, Zhang Y, Yi C, Chen S, Li W, Wang D, Wang Z, Di W, Shen K, Lin D, Shen H, Feng Y, Xie X, Ma D. A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. Nat Genet 2013; 45(8): 918–922
46. Li N, Li B, Zhan X. Comprehensive analysis of tumor microenvironment identified prognostic immune-related gene signature in ovarian cancer. Front Genet 2021; 12: 616073
47. Palma P, Cuadros M, Conde-Muíño R, Olmedo C, Cano C, Segura-Jiménez I, Blanco A, Bueno P, Ferrón JA, Medina P. Microarray profiling of mononuclear peripheral blood cells identifies novel candidate genes related to chemoradiation response in rectal cancer. PLoS One 2013; 8(9): e74034
48. Rivière E, Pascaud J, Tchitchek N, Boudaoud S, Paoletti A, Ly B, Dupré A, Chen H, Thai A, Allaire N, Jagla B, Mingueneau M, Nocturne G, Mariette X. Salivary gland epithelial cells from patients with Sjögren's syndrome induce B-lymphocyte survival and activation. Ann Rheum Dis 2020; 79(11): 1468–1477
49. Goldmann K, Spiliopoulou A, Iakovliev A, Plant D, Nair N, Cubuk C, McKeigue P, Barnes MR, Barton A, Pitzalis C, Lewis MJ. Expression quantitative trait loci analysis in rheumatoid arthritis identifies tissue specific variants associated with severity and outcome. Ann Rheum Dis 2024; 83(3): 288–299
50. Wang S, Wang Q, Fan B, Gong J, Sun L, Hu B, Wang D. Machine learning-based screening of the diagnostic genes and their relationship with immune-cell infiltration in patients with lung adenocarcinoma. J Thorac Dis 2022; 14(3): 699–711
51. Hao D, Han G, Sinjab A, Gomez-Bolanos LI, Lazcano R, Serrano A, Hernandez SD, Dai E, Cao X, Hu J, Dang M, Wang R, Chu Y, Song X, Zhang J, Parra ER, Wargo JA, Swisher SG, Cascone T, Sepesi B, Futreal AP, Li M, Dubinett SM, Fujimoto J, Solis Soto LM, Wistuba II, Stevenson CS, Spira A, Shalapour S, Kadara H, Wang L. The single-cell immunogenomic landscape of B and plasma cells in early-stage lung adenocarcinoma. Cancer Discov 2022; 12(11): 2626–2645
52. Lu X, Ma L, Yin X, Ji H, Qian Y, Zhong S, Yan A, Zhang Y. The impact of tobacco exposure on tumor microenvironment and prognosis in lung adenocarcinoma by integrative analysis of multi-omics data. Int Immunopharmacol 2021; 101(Pt B): 108253
53. Xu M, Li C, Xiang L, Chen S, Chen L, Ling G, Hu Y, Yang L, Yuan X, Xia X, Zhang H. Assessing the causal relationship between 731 immunophenotypes and the risk of lung cancer: a bidirectional mendelian randomization study. BMC Cancer 2024; 24(1): 270
54. Liu X, Wu S, Yang Y, Zhao M, Zhu G, Hou Z. The prognostic landscape of tumor-infiltrating immune cell and immunomodulators in lung cancer. Biomed Pharmacother 2017; 95: 55–61
55. Jang HJ, Lee HS, Ramos D, Park IK, Kang CH, Burt BM, Kim YT. Transcriptome-based molecular subtyping of non-small cell lung cancer may predict response to immune checkpoint inhibitors. J Thorac Cardiovasc Surg 2020; 159(4): 1598–1610.e1593
56. Hou L, Zhang S, Yu W, Yang X, Shen M, Hao X, Ren X, Sun Q. Single-cell transcriptomics reveals tumor-infiltrating B cell function after neoadjuvant pembrolizumab and chemotherapy in non-small cell lung cancer. J Leukoc Biol 2024; 116(3): 555–564
57. Bakkila BF, Marks VA, Kerekes D, Kunstman JW, Salem RR, Billingsley KG, Ahuja N, Laurans M, Olino K, Khan SA. Impact of COVID-19 on the gastrointestinal surgical oncology patient population. Heliyon 2023; 9(8): e18459