Discover Oncology

Analysis

To explore the molecular mechanisms and shared genetic characteristics of rheumatoid arthritis and cervical cancer based on multiple omics and clinical samples

Xifeng Xu[1] · Yujie Huang[2] · Wu Wei[3] · Zhong Lin[2] · Yongjin Luo[1] · Kaiyi Meng[1]

Received: 24 January 2025 / Accepted: 16 May 2025
Published online: 29 May 2025

Abstract

Objective The potential association between rheumatoid arthritis (RA) and cervical cancer risk is still debated and necessitates additional clinical investigations. This study aimed to explore this relationship using Mendelian randomization and multi-omics analysis, aiming to enhance insights and reduce redundancy.

Methods This study employs a Mendelian randomization approach to investigate potential associations between rheumatoid arthritis (RA) and cervical cancer susceptibility. Single-cell enrichment scores were calculated using RA-specific transcriptome differentially expressed genes, and prognostic models were developed using 10 machine learning algorithms to assess cervical cancer differential gene expression associated with RA scores. Subsequently, differences in clinical characteristics, immune cell infiltration, immunotherapy efficacy, and response to chemotherapeutic drugs were analyzed between high- and low-risk groups of patients. Finally, the expression of model genes was verified by cervical cancer cell lines and clinical fresh samples.

Results Mendelian randomization has revealed an increased incidence of cervical cancer associated with RA. RA enrichment scores show predominant enrichment in the single-cell NK cell subpopulation. Cervical cancer can be distinctly categorized into two subgroups based on RA score-associated prognostic genes, demonstrating significant differences in immune cell infiltration and prognosis between these subgroups. Prognostic modeling indicates that patients in the low-risk group exhibit better prognosis, enhanced immune cell infiltration, and improved response to immunotherapy and drug treatments. Finally, multiple external data confirmed that there were significant differences in the expression of model genes.

Conclusion Rheumatoid arthritis is causally associated with the development of cervical cancer. Patients with rheumatoid arthritis have an increased risk of concurrent cervical cancer. Immune cell pathways, such as NK cell-based, may be important in increasing the risk of cervical cancer in RA patients. The abnormal expression of the model gene may be involved in the progression of cervical cancer patients and the impact of immunotherapy.

Keywords Rheumatoid arthritis · Cervical cancer · Mendelian randomization · Single cell · Multi-omics analysis

Xifeng Xu and Yujie Huang contributed equally to this paper and are co-first authors.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12672-025-02741-7.

✉ Yongjin Luo, luoyongjin951022@163.com; ✉ Kaiyi Meng, mengkaiyi19830224@163.com | [1]Nanning Second People's Hospital, The Third Affiliated Hospital of Guangxi Medical University, Nanning, China. [2]Department of Gynaecology, Guangxi Reproductive Hospital, Nanning 530000, Guangxi, China. [3]Guangxi Zhuang Autonomous Region People's Hospital, Nanning, China.

1 Introduction

Recent research has underscored the pivotal role of the immune system in tumor progression [1]. Rheumatoid arthritis (RA) is a chronic systemic disease primarily characterized by inflammatory synovitis of unknown etiology [2], affecting approximately 0.5–1.0% of the general population [3]. Patients with RA often experience significant comorbidities such as major psychiatric disorders, pulmonary disease, cardiovascular disease, and solid malignancies [4].In women, RA has been associated with an increased risk of cervical cancer [5], but a decreased risk of breast cancer [6]. However, the causal relationship between rheumatoid arthritis and the development of cervical cancer remains uncertain.

In a Swedish cohort study involving 374,944 females, biologic-naïve women with RA were found to have a higher risk of cervical intraepithelial neoplasia (CIN) I (hazard ratio 1.53, 95% CI 1.23–1.89) and CIN II–III (hazard ratio 1.39, 95% CI 1.16 to 1.66) [7]. Infection with high-risk strains of human papillomavirus (HPV) and the integration of HPV genomes into host chromosomes of cervical epithelial cells are crucial early events in the neoplastic progression of cervical lesions [8]. Previous studies suggest that autoimmune diseases like rheumatoid arthritis can induce immunosuppression in patients, potentially increasing their susceptibility to HPV infections [9, 10]. Research has also highlighted a significant association between rheumatoid arthritis and HPV infection, as well as cervical developmental abnormalities [11]. However, the specific mechanism by which rheumatoid arthritis contributes to cervical carcinogenesis remains unclear.

Mendelian randomization (MR) utilizes genetic instrumental variables, providing a novel approach to assess causal relationships between exposure factors and outcomes [12]. MR studies have indicated that RA may elevate the risk of cervical cancer in patients. In this study, we conducted transcriptome data analysis using Mendelian randomization, applied multiple algorithms to compute the expression scores of RA differential genes in single cells of cervical cancer, and subsequently developed prognostic models for cervical cancer patients based on these high and low scoring differential genes. We compared prognostic differences and responses to immunotherapy between different risk groups to investigate shared molecular mechanisms between the two diseases.

2 Method

The overview design of our work was shown in Fig. 1. Data sources for MR and transcriptome data are publicly available online.

2.1 Mendelian randomization analysis

2.1.1 Data sources

Our work is based on aggregated data from the European ancestor Gwas.

I. Estimates of RA effects for snp associated with RA risk were derived from a study that included 3730 cases and 333,429 controls of European ancestry(ukb-a-105).
II. ii.Genetic instrumentation for cervical cancer was extracted from a large GWAS study comprising 1,889 cases and 461,044 controls of European ancestry(ukb-b-8777).

2.1.2 Generation of genetic instruments

All Mendelian randomization (MR) methods rely on three core assumptions to mitigate bias in MR estimation: (i) instrumental variables (IVs) are strongly correlated with RA, (ii) IVs influence cervical cancer solely through their impact on RA, and (iii) IVs are independent of any other confounding factors. Failure to uphold these assumptions can compromise the reliability of conclusions. To ensure optimal IV selection associated with RA, the following steps were implemented:Firstly, single nucleotide polymorphisms (SNPs) closely linked to RA were identified using a statistically significant threshold ($p < 5 \times 10–8$) from published GWAS studies. Secondly, SNPs in high linkage disequilibrium (LD) were assessed to avoid redundancy; those exceeding the threshold (window size = 10,000 kb, R2 < 0.005) were excluded through LD analysis. Based on the established screening criteria, a total of 13 single nucleotide polymorphisms (SNPs) were identified as significantly associated with rheumatoid arthritis (RA). Notably,

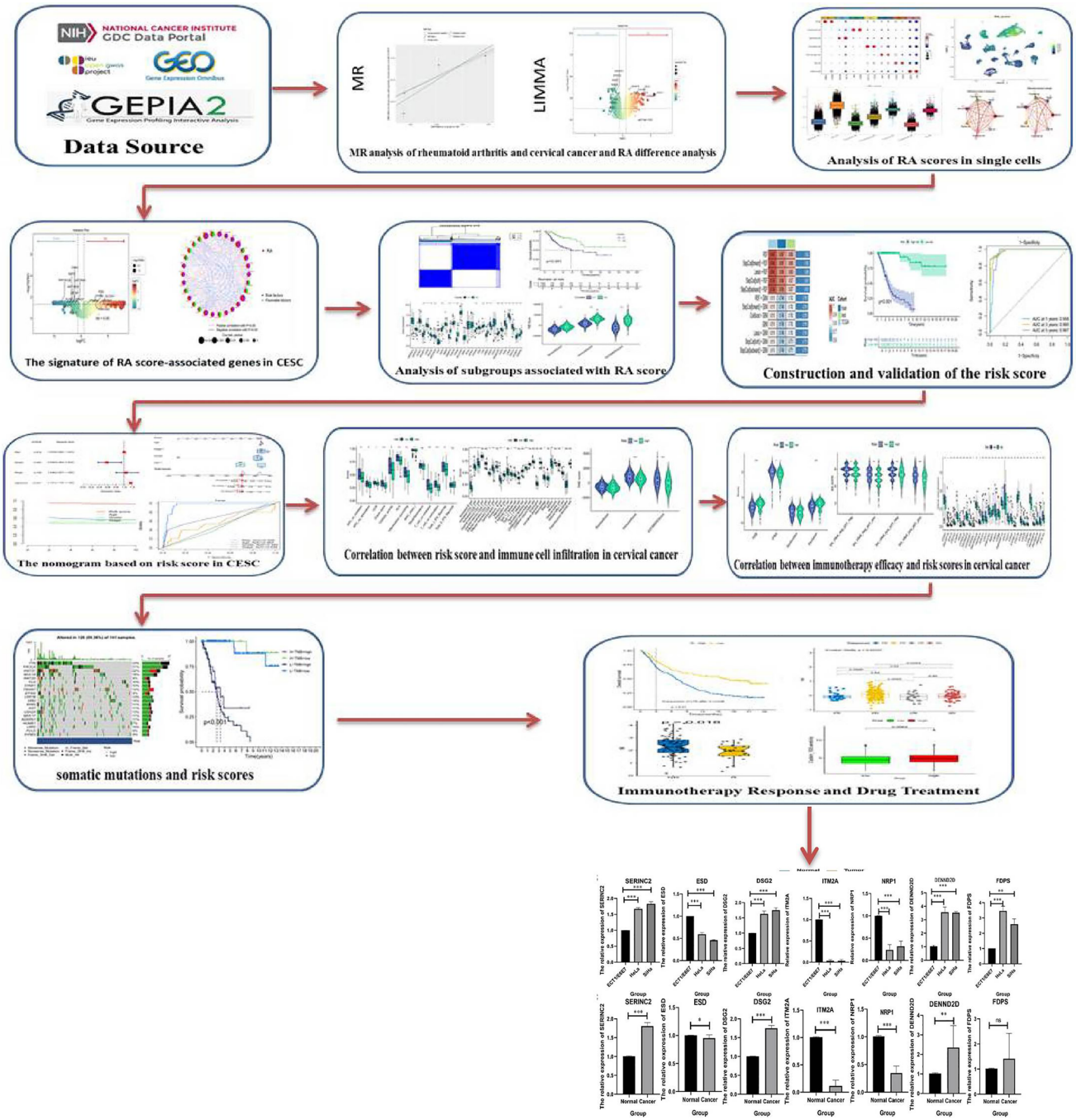

Fig. 1 Research flow chart

all these genetic variants demonstrated F-values exceeding the threshold of 10, indicating strong statistical significance. Subsequently, through comprehensive comparative analysis, four common SNPs were discovered to be shared between RA and cervical cancer, suggesting potential genetic overlap between these two distinct pathological conditions.

2.1.3 Two-sample R analysis

An inverse variance weighted (IVW) meta-analysis of Wald ratio estimates was performed to explore the causal effect of RA on cervical cancer risk. Since the IVW test presents a weighted regression of exposure outcomes with an intercept constrained to zero, the estimates may be biased. In this case, we combined MR-Egger regression and weighted median test to analyze the causal relationship.

2.1.4 Removal of horizontal pleiotropy and sensitivity analyses

Mendelian randomized multiple validity residuals and outliers (MR-Presso) analysis, leave-one-out analysis, and Cochrane's Q-test were used to extensively assess MR results. Notably, the MR-PRESSO method corrects for the presence of horizontal pleiotropy by removing peripheral SNP. Leave-one-out analysis can analyze the effect of outliers. Based on IVW and MR-Egger estimation, Cochran's Q-test identifies the SNP leading to heterogeneity.

2.2 Transcriptomic analyses

2.2.1 Data acquisition

Microarray expression and clinical data of GSE55457 (13 RA cases and 10 controls) were downloaded from Gene expression Omnibus (GEO) database. The RNA sequencing data of 3 normal tissues and 306 cervical cancer samples with corresponding clinical information were obtained from Cancer Genome Atlas (TCGA). The GSE44001 dataset including 300 patients with cervical cancer and the GSE26712 dataset of 185 ovarian cancers were used to understand the prognosis of genes associated with RA scores. Single-cell data were collected from the National Genomics Data Center (access number: PRJCA008573), including five cases of cervical cancer.

2.2.2 ScRNA-seq data preprocessing and analysis

In the RA dataset GSE55457, differential gene analysis was conducted using calibrated criteria of $p < 0.05$ and $|\log FC| > 1$, and results were visualized through volcano plots created with ggplot2. For scRNA-seq datasets processed separately with Seurat v4 [13]. Genes ranging from 500 to 7500 were selected based on their expression profiles, ensuring mitochondrial gene scores were below 15%. Following gene selection, data normalization was performed using "NormalizeData", and the top 2000 highly variable genes (HVGs) were identified to stabilize UMI count variances. Subsequently, data underwent correction for batch effects using Harmony and underwent cell cycle correction. PCA cluster analysis identified seven clusters using PC 40 and a resolution of 0.01, characterized by typical cell type marker genes. Major cell type markers included GNLY and NKG7 for natural killer (NK) cells, EPCAM and KRT19 for epithelial cells, COL1A1 and ACTA2 for fibroblasts, LYZ and APOE for macrophages, IGHG1 and IGKC for plamsa cells, VWF and PLVAP for endothelial cells, and MS4A1 and CD79A for B cells. Bubble plots were used to visualize the expression of marker genes across different cell subpopulations. RA-related genes were scored using Singscore, Modulescore, AUC, and UCell methods. The average of these scores was calculated, and cells were classified into high and low RA score groups based on the median of these averages. Differential gene analysis between these groups was conducted using the FindMarkers function [14].Cellular communication between different subgroups was analyzed using the cellchat package [15] to assess the strength and frequency of interactions across cell types.

2.2.3 Characterization of genes associated with RA score

Differential analysis of RA score-associated genes in cervical cancer was initially conducted using calibrated criteria of $p < 0.05$ and $|\log FC| > 1$, visualized through volcano plots generated with ggplot2. Subsequently, 404 RA score-related genes were subjected to survival analysis among cervical cancer patients using univariate Cox regression ($p < 0.01$), and results were illustrated with forest plots. The interaction between different genes was explored, alongside analysis

of somatic mutation incidence, genetic loci, and copy number variations (CNV) affecting these genes. Finally, univariate Cox regression analysis was performed using GSE44001 cervical cancer data.

2.2.4 Consensus clustering analysis

CESC patients were clustered based on RNA sequences from 11 RA score-related genes in cervical cancer tissue. The consensus clustering utilized the K-means algorithm implemented in the "ConsensusClusterPlus" R package [16], repeated 50 times to ensure robust subtype identification. Optimal subcluster numbers were determined using consensus cumulative distribution function (CDF), delta area, and clustering consensus (CLC) metrics.In the TCGA cohort, CESC patients were stratified into two groups based on consensus RA score-related gene clusters. Overall survival (OS) analysis between these groups was performed using Kaplan–Meier (KM) analysis, facilitated by the "survival" and "survminer" R packages. Additionally, Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were conducted for both subclusters using the "clusterProfiler" R package. These analyses aimed to elucidate potential biological processes and pathways associated with the identified molecular subgroups.

2.2.5 Generation of the RA score prognostic signature

To enhance model accuracy, we conducted univariate Cox regression analyses on differential RA score genes in cervical cancer using GSE44001 and GSE26712 datasets, including genes with $p < 0.05$ in subsequent analyses. Prognostic genes identified in TCGA data and from the aforementioned datasets were intersected separately and deduplicated to derive model genes. Seven genes—SERINC2, ESD, DSG2, ITM2A, NRP1, DENND2D, and FDPS—were selected for inclusion in machine learning models. The screening criterion was p less than 0.05.The cervical cancer data from TCGA was split into training and validation sets in a 1:1 ratio. Ten machine learning methods were employed with 101 algorithm combinations to construct prognostic models simultaneously on the training, validation, and TCGA cohort datasets [17]. These methods included Random Survival Forest (RSF), Elastic Net (Enet), Lasso, Ridge, Stepwise Cox, CoxBoost, Cox Partial Least Squares Regression (plsRcox), Supervised Principal Components (SuperPC), Generalized Augmented Regression Model (GBM), and Survival Support Vector Machine (Surviva-svm).The Harrell consistency index (C-index) was calculated for each model on the training and validation sets as well as the TCGA cohort. The model with the highest average C-index across these datasets was selected as the optimal model. Additionally, the "survival" package was utilized for survival modeling and Kaplan–Meier (KM) analysis, while the "timeROC" package was employed to generate ROC curves.

2.2.6 Nomogram construction and characteristic of risk score model

Initially, we performed univariate and multivariate Cox regression analyses using the "survival" package to examine the relationship between the risk score and clinical characteristics. Subsequently, we integrated prognostic and clinical factors to construct a nomogram using the R package "RMS". The performance of both the risk score model and the nomogram was assessed using calibration curves and time-dependent ROC curves.Following this, we compared clinical characteristics, tumor stemness index, and immune scores between the high-risk and low-risk groups.

2.2.7 Analyses of relation between tumor mutation burden, tumor stemness indices, and risk score and prediction of the effective response of postoperative immunotherapy

The R package "GSVA" was used for ssGSEA [18] and "CIBERSORT [19]" for immune cell infiltration calculations to examine differences in immune cell infiltration levels between the two risk groups. Additionally, the TIDE algorithm [20] was employed to assess response to immune checkpoint blockade therapy (ICB therapy) and predict neoantigen activity (http://tide.dfci.harvard.edu). The TIDE algorithm integrates two primary mechanisms of immune evasion: T-cell dysfunction and T-cell rejection. Higher TIDE scores indicate a greater likelihood of tumor immune escape and poorer response to immunotherapy. Immunotherapy scores for a large sample cohort from the TCGA database were obtained from TCIA (https://tcia.at/home), comparing treatment scores for anti-CTLA4 and anti-PD1 inhibitors across different risk groups. Somatic mutation data (MuTect2 Variant Aggregation and Masking) were downloaded from the TCGA-CESC dataset, and mutation frequencies in specific patients were calculated using the R package "maftools" [21]. TMB, defined as the total number of mutations per million bases of somatic cells, was analyzed using the "maftools" package. Kaplan–Meier (KM) analysis was conducted to reveal overall survival (OS) differences between TMB groups.

2.2.8 Correlation of drug sensitivity and risk score

Survival differences and immunotherapy effects between different risk groups in IMvigor210 were initially compared. Survival differences between different risk groups in GSE78220 [22] and GSE135222 [23] and treatment sensitivity in the GSE91061 [23] dataset were compared. Correlations between risk scores and the IC50 values of commonly used chemotherapeutic drugs such as cisplatin, docetaxel, and paclitaxel, as well as common targeted therapy drugs, were investigated using IC50 data obtained from the Genomics of Drug Sensitivity in Cancer (GDSC).

2.2.9 Validation of model gene expression levels

The expression levels of the seven model genes were validated using the GEPIA 2 online database (http://gepia2.cancer-pku.cn/). Additionally, gene expression in cervical cancer tissues and normal cervical tissues was confirmed using the HPA protein database (https://www.proteinatlas.org/).

2.2.10 Cell culture

Cervical cancer cells (HeLa and SiHa) and normal cervical epithelial cells (ECT1/E6E) were kindly donated to Li Li's group at the Cancer Hospital of Guangxi Swell Medical University. These cells were maintained in RPMI 1640 medium (Sigma—Aldrich; Thermo Fisher Scientific) containing 10% fetal bovine serum (FBS) (FBS; Gibco; Expression Vector; Thermo Fisher Scientific), 100 IU/mL penicillin and 10 µg/mL streptomycin (Thermo Fisher Scientific). All cells were cultured at 37 °C under 5% CO2.

2.2.11 Clinical sample acquisition

A total of 10 cases of fresh cervical cancer tissues and 10 cases of fresh normal cervical tissues were collected in the Nanning Second People's Hospital from 2023 to 2024. Written informed consent was obtained from each patient and approved by the Research Ethics Committee of the Nanning Second People's Hospital. The case selection criteria were established as follows: participants aged between 55 and 60 years with negative human papillomavirus (HPV) status were included. Furthermore, individuals were excluded if they presented with any chronic systemic conditions, including but not limited to hypertension, diabetes mellitus, or previous malignancies. Additionally, cervical tissue samples were obtained from patients undergoing hysterectomy for uterine fibroids or adenomyosis, with histological confirmation showing normal cervical morphology and clear surgical margins. Written informed consent was acquired from all participating individuals following a comprehensive explanation of the study protocol. Participants were explicitly assured regarding the absence of any potential secondary victimization or financial obligations associated with their involvement. Moreover, the research protocol underwent rigorous review and subsequently received formal approval from the Institutional Review Board (IRB) of the Second People's Hospital of Nanning, with the assigned ethical approval number Y2023019. The acquisition and use of human cervical tissue samples in this study were conducted in full compliance with the ethical principles outlined in the Declaration of Helsinki.

2.2.12 qRT-PCR analysis

These RNAs were extracted by TRIzol® (Invitrogen; Thermo Fisher Scientific, Inc) and reverse transcribed by HiScript III RT SuperMix for qPCR (Vazyme Biotech). Gene primer sequences are shown in Table S2. PCR cycling parameters were: 95 °C for 5 min, 40 cycles of 30 s, and 60 °C for 1 min. each sample was performed in 3 separate reactions, and the average of each point was calculated. mRNA expression levels were normalized to GAPDH levels.

2.3 Statistical analysis

In this experiment, all statistical analyses were conducted using R 4.3. For non-parametric data, comparisons between two independent samples and multiple samples were performed using the Wilcoxon test and Kruskal–Wallis test, respectively. Parametric data were analyzed using the t-test and one-way ANOVA. A p-value < 0.05 was considered

statistically significant (* p-value < 0.05; ** $p < 0.01$; *** $p < 0.001$). Relevant R packages, including "ggplot2", "ggpubr", "survival", "survminer", and other R packages, were obtained from Bioconductor or R packages. The threshold for statistical significance was set at $p < 0.05$ for each analysis.

3 Result

3.1 Mendelian randomization analysis of rheumatoid arthritis and cervical cancer

As depicted in Table 1, MR analysis revealed a causal association between RA and cervical cancer in a European cohort (IVW: OR 1.15, 95% CI 1.06–1.24, $p < 0.001$; Weighted median: OR 1.15, 95% CI 1.11–1.20, $p < 0.001$; Weighted mode: OR 1.15, 95% CI 1.08–1.22, $p < 0.05$; Weighted mode: OR 1.15, 95% CI 1.10–1.20, $p < 0.05$) (Table 1 and Fig. 2). First, MR-Egger regression was employed to investigate horizontal pleiotropy, and the results indicated no significant impact of pleiotropy on causality (all p-values > 0.05). Second, the MR Egger test results were consistent with the IVW method without outliers, affirming the reliability of the initial findings. Third, to further explore the relationship between RA and cervical cancer in the European cohort, we conducted leave-one-out analysis and Cochrane Q-test. The leave-one-out analysis did not identify any single SNP driving the causal relationship between RA and cervical cancer (Fig. 2). Cochrane Q-tests showed p-values all less than 0.05 (Q-value of 10.48, $p < 0.05$ for IVW test; Q-value of 14.70, $p < 0.05$ for MR-Egger test), indicating heterogeneity among the SNPs.

3.2 Analysis of RA associated with cervical cancer single cells

Differential analysis of RA sequencing data yielded 312 differential genes, comprising 123 down-regulated and 189 up-regulated genes (Table S1, Fig. 3A). Using single-cell analysis, differential gene enrichments were scored across high-quality transcriptome data obtained from 48,045 cells after rigorous quality assurance and filtering. Debatching of samples using harmony reduced the differences in cell distribution. Subsequent PCA clustering of cells identified seven distinct clusters with minimal inter-cluster influence (Fig. 3B).Cellular annotation based on marker genes categorized the seven clusters into epithelial cells, endothelial cells, fibroblasts, macrophages, B-cells, plamsa cells, and NK cells (Fig. 3C). Bubble plots illustrating marker gene expression in different cell subgroups highlighted distinct expression patterns (Fig. 3D). RA scores derived from differential genes were computed using four scoring methods and mapped onto box plots, revealing predominant enrichment in NK cells across various algorithms (Supplementary Fig. S1). A consolidated scoring outcome based on the average score from the four algorithms demonstrated the cellular distribution of scores in UMAP (Fig. 3E), with high expression particularly noted in NK cells as depicted in Fig. 3F and 3G.Cells were stratified into high and low RA score subpopulations based on median values, and differential gene analysis between these subtypes was conducted. Analysis of cellular communication indicated heightened interactions in the high-score subpopulation (Fig. 3H). Specifically, NK cells exhibited more frequent and intense communication with epithelial cells, fibroblasts, and endothelial cells, with particularly strong interactions observed between NK cells and fibroblasts (Fig. 3I,J).These findings collectively suggest that RA scores influence NK cell dynamics in cervical cancer patients, thereby impacting fibroblasts and epithelial cells (tumor cells).

Table 1 MR analysis revealed a causal association between RA and cervical cancer in a European cohort

id.exposure	id.outcome	method	nsnp	b	se	pval	or	or_lci95	or_uci95
ukb-a-105	ukb-b-8777	MR Egger	4	0.186889	0.070112	0.116626	1.205494	1.050715	1.383074
ukb-a-105	ukb-b-8777	Weighted median	4	0.144097	0.02021	1.00E-12	1.154996	1.11014	1.201665
ukb-a-105	ukb-b-8777	Inverse variance weighted	4	0.136623	0.040813	0.000815	1.146396	1.058264	1.241867
ukb-a-105	ukb-b-8777	Simple mode	4	0.139261	0.031871	0.022172	1.149424	1.079818	1.223516
ukb-a-105	ukb-b-8777	Weighted mode	4	0.139261	0.021726	0.007694	1.149424	1.101505	1.199427

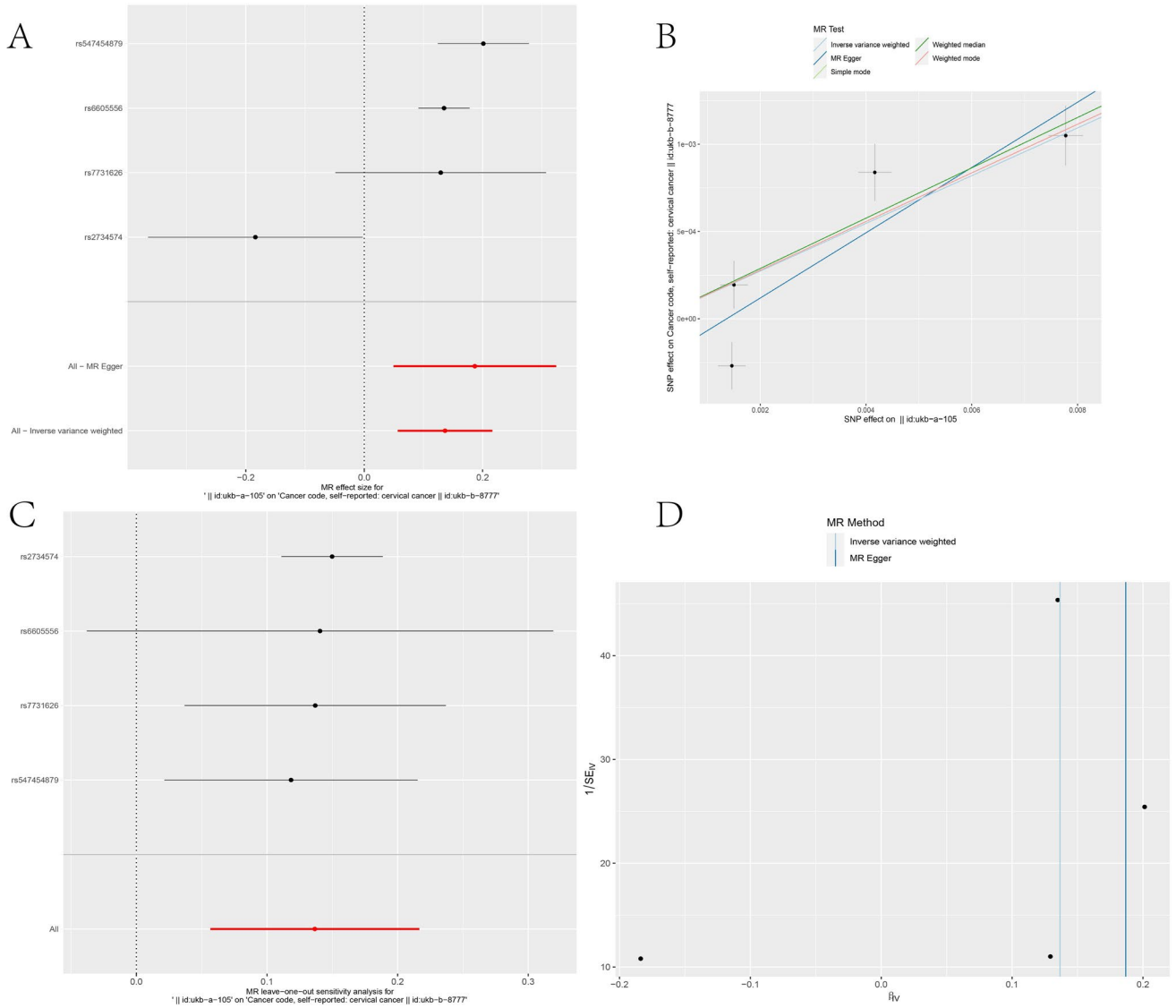

Fig. 2 Mendelian randomization analysis of causal effects. **A**: Distribution of genetic variant effects (n = 4). **B**: Results from Mendelian randomization (MR) methods, including MR test, weighted median, MR Egger, and weighted mean, with effect estimates (β) and confidence intervals shown. Negative values indicate inverse causal relationships. Graphs represent the aggregated analysis. **C**: Leave-one-out sensitivity analysis (forest plot) demonstrating the robustness of causal estimates upon sequential exclusion of individual variants. **D**: Funnel plot assessing potential pleiotropy or bias in the causal relationship between rheumatoid arthritis (RA) and cervical cancer (CC)

3.3 The signature of RA score-associated genes in CESC

Of the 1526 RA score-associated genes, 404 were found to exhibit significant differences in cervical cancer, including 150 down-regulated genes and 254 up-regulated genes, as depicted in Fig. 4A. Univariate Cox analysis conducted on these 404 genes identified 30 genes that were associated with prognosis in cervical cancer (Fig. 4B). To illustrate the intricate relationship between RA score-related genes and their prognostic implications in cervical cancer, a network diagram was constructed (Fig. 4C).The somatic mutation frequencies of these 30 genes in cervical cancer were investigated, with ROCK2 showing the highest prevalence of mutations (up to 15%), while the remaining genes exhibited lower mutation frequencies (Fig. 4D). The genomic locations of these 30 genes are depicted in Fig. 4E. Analysis further revealed prevalent copy number variations (CNVs) among these genes, with TFRC, EFNA1, and LIMD2 displaying amplifications, and ESD showing deletions (Fig. 4F).Comparative analysis indicated that the majority of RA score-related genes were significantly up-regulated in cervical cancer tissues compared to normal tissues (Fig. 4G).

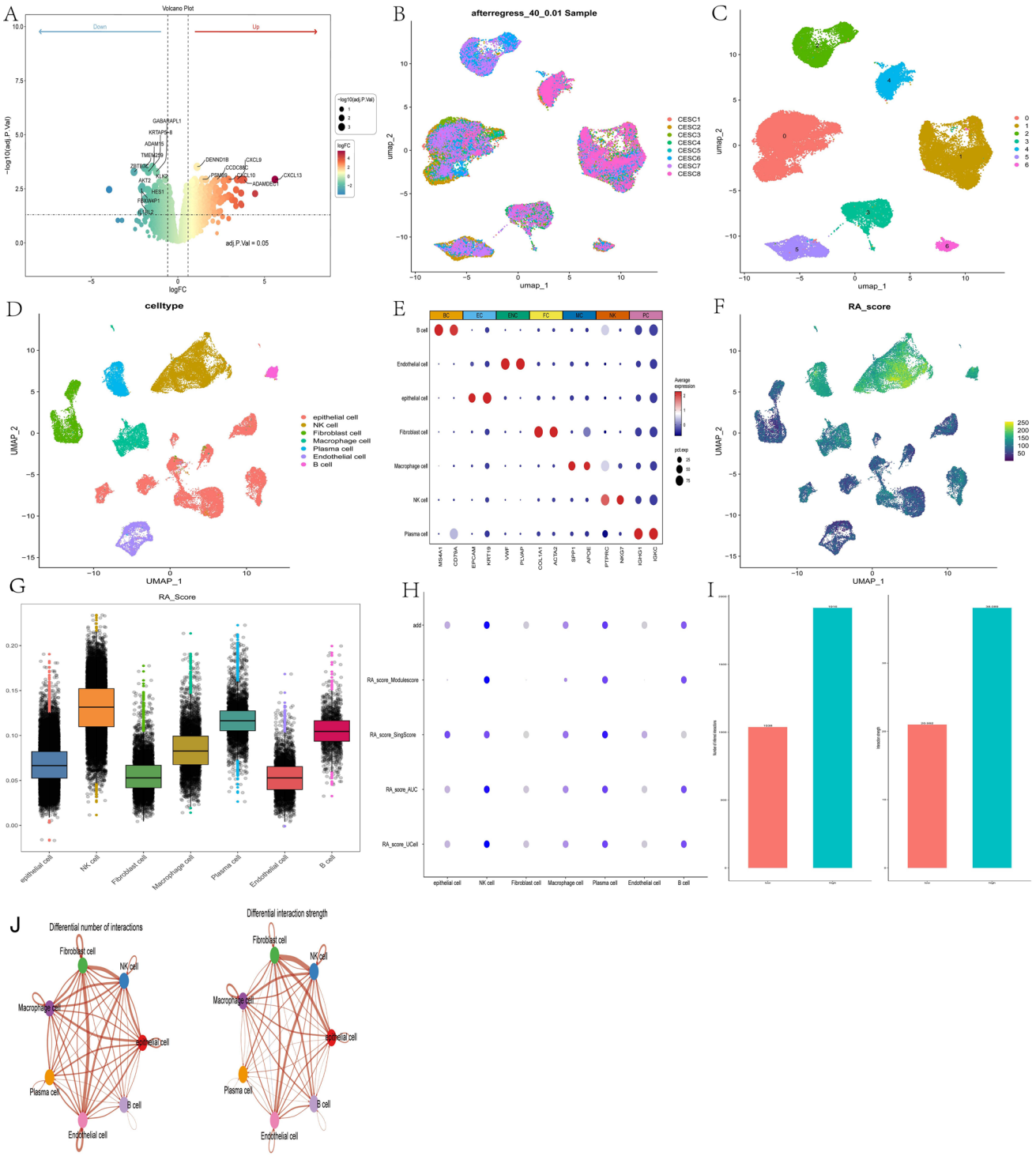

Fig. 3 Single-cell analysis of rheumatoid arthritis (RA)-associated genes in cervical cancer. **A**: Volcano plot of differentially expressed genes (DEGs) associated with RA. **B**: UMAP visualization of single-cell data after batch correction. **C**: Cell clustering analysis identifying 7 distinct clusters. **D**: UMAP plot showing the distribution of different cell subpopulations. **E**: Bubble plot displaying marker genes for each cell subpopulation. **F**: UMAP visualization of RA-associated scores across cell subpopulations. **G**: Boxplot comparing RA-associated scores among different cell subpopulations. **H**: Dot plot showing RA scores calculated by multiple analytical algorithms. **I**–**J**: Cell–cell communication analysis depicting (I) the number and (II) strength of signaling interactions between cell populations

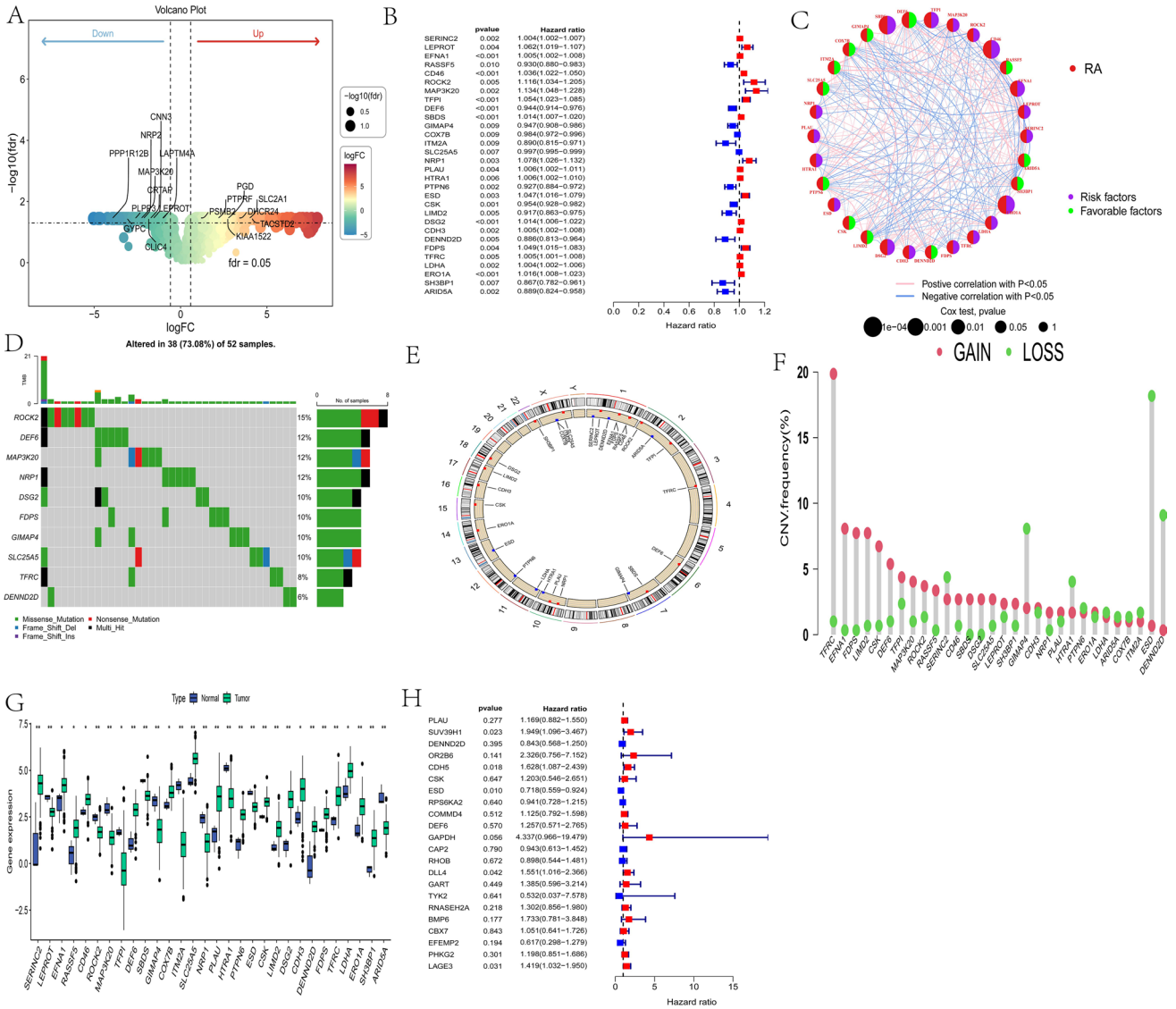

Fig. 4 Prognostic analysis of RA-associated signature genes in cervical cancer. **A**: Volcano plot showing differential expression of RA-associated genes in TCGA-CC dataset. **B**: Forest plot of univariate Cox regression analysis identifying prognostic genes. **C**: Protein–protein interaction network of prognostic genes, with purple nodes indicating risk factors and green nodes representing protective factors. **D**: Mutation waterfall plot of prognostic genes, showing ROCK2 with the highest mutation frequency (15%). **E**: Circos plot displaying chromosomal locations of prognostic genes. **F**: CNV bar plot revealing copy number variations, with TFRC showing the highest gain frequency. **G**: Boxplot demonstrating differential expression patterns of prognostic genes in TCGA (majority showing elevated expression). **H**: Validation forest plot confirming the prognostic value of identified genes in GEO datasets

Assessment of the impact of these 30 genes on overall survival (OS) in cervical cancer patients using GEO dataset showed that the expression levels of DLL4, LAGE3, CDH5, SUV39H1, and ESD were statistically correlated with OS (Fig. 4H).

3.4 Analysis of subgroups associated with RA score

Consistent cluster analysis was utilized to classify 304 cervical cancer patients into two groups, C1 and C2 (Table S2), as illustrated in Fig. 5A–C. The expression differences of the 30 genes between the C1 and C2 clusters are shown in Fig. 5D. Kaplan–Meier (KM) analysis revealed significant differences in survival between patients in the C1 and C2 groups ($p < 0.001$) (Fig. 5E).Regarding the differences in the tumor microenvironment (TME) between the two clusters, C2 exhibited higher abundance of various immune cell infiltrates, including regulatory T cells, CD8 T cells, activated NK cells, and

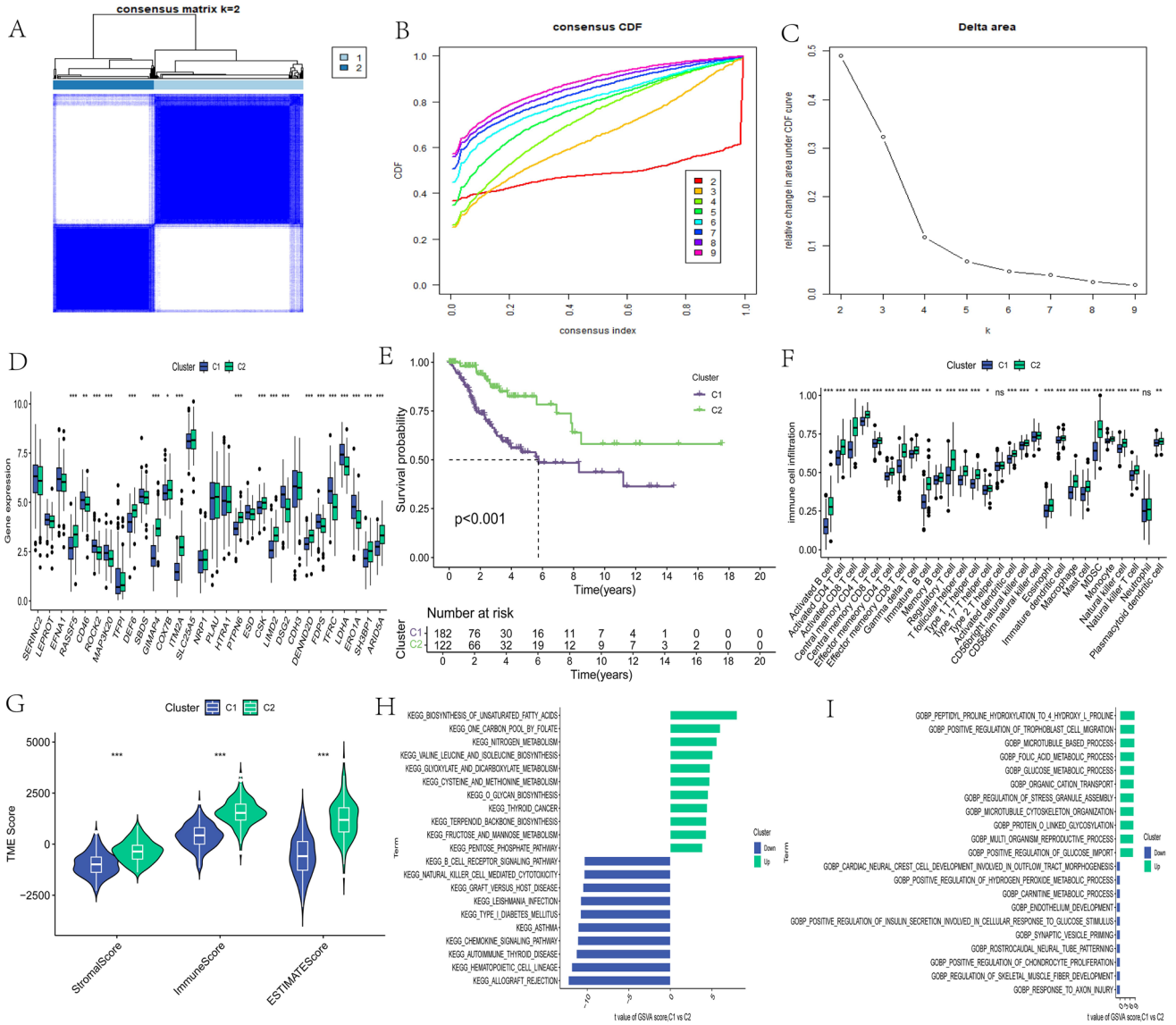

Fig. 5 Molecular subtyping of cervical cancer based on RA-associated signature genes. **A**–**C**: Consensus clustering analysis identified two distinct subtypes (C1 and C2) in TCGA cervical cancer cohort. **D**: Boxplot showing differential expression of prognostic genes between subtypes, with most genes showing significant differences (except LEPROT and few others). **E**: Kaplan–Meier survival analysis demonstrating better prognosis for C2 subtype patients (log-rank $p < 0.05$). **F**: Immune cell infiltration analysis revealing higher infiltration of most immune cell types in C2 subtype. **G**: Comparison of immune scores showing significantly elevated scores in C2 subtype for multiple immune signatures. **H**–**I**: Functional enrichment analysis of differentially expressed genes between subtypes: (**H**) Gene Ontology (GO) terms and (**I**) KEGG pathways

dendritic cells, among others (Fig. 5F). Evaluation of the immune score in cervical cancer tumor tissues indicated that immune score, stromal score, and estimate score were markedly elevated in the C2 subpopulation (Fig. 5G).Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis highlighted significantly activated pathways in C2, such as BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS and ONE CARBON POOL BY FOLATE, among others (Fig. 5H). Furthermore, Gene Ontology (GO) biopathway analysis demonstrated significant enrichment in functional groups such as GLUCOSE_METABOLIC_PROCESS and PEPTIDYL_PROLINE_HYDROXYLATION_TO_4_HYDROXY_L_PROLINE in C2 (Fig. 5I).

3.5 Construction and validation of the risk score

Univariate Cox analysis was conducted using the GSE44001 and GSE26712 datasets, identifying 27 genes associated with patient prognosis in the former and 20 genes in the latter (Supplementary Fig. S2). Subsequently, the prognostic

genes from these datasets were intersected with those from the TCGA dataset, resulting in the identification of seven genes that were consistently associated with prognosis across all datasets (Supplementary Fig. S2 and Table S3).The TCGA dataset was then split into a training set (N = 152) and a validation set (N = 152) at a 1:1 ratio, and the clinical characteristics of these datasets are detailed in Table S4. We employed 101 different combinations of machine learning methods using the leave-one-out cross-validation (LOOCV) framework to construct predictive models on the training set, validation set, and the entire TCGA cohort. The concordance index (c-index) was calculated for each model (Fig. 6A). Ultimately, the random survival forest (RSF) model emerged as the best-performing model, achieving an average c-index of 0.894 (Fig. 6A, bar on the right panel).Using the median risk score, the training set, validation set, and TCGA cohort were stratified into high- and low-risk groups (Table S5). Consistently, patients in the low-risk group exhibited significantly better overall survival (OS) compared to those in the high-risk group across all datasets ($p < 0.001$, Fig. 6B–D). In addition, we further analyzed the relationship between model genes and the prognosis of cervical cancer patients (Supplementary Figure. S3).Finally, ROC curves were analyzed for 1-, 3-, and 5 year survival predictions across different datasets. Notably, the 1-year AUCs were 0.925, 0.872, and 0.958 for the training set, validation set, and TCGA cohort, respectively; the 3 year AUCs were 0.927, 0.916, and 0.980, respectively; and the 5 year AUCs were 0.991, 0.873, and 0.987, respectively (Fig. 6E–G). Comparing the median survival results of the patients in the two risk groups revealed that the median survival time was lower in the high-risk patients (1.48 years) compared to the patients in the low-risk group (2.08 years).

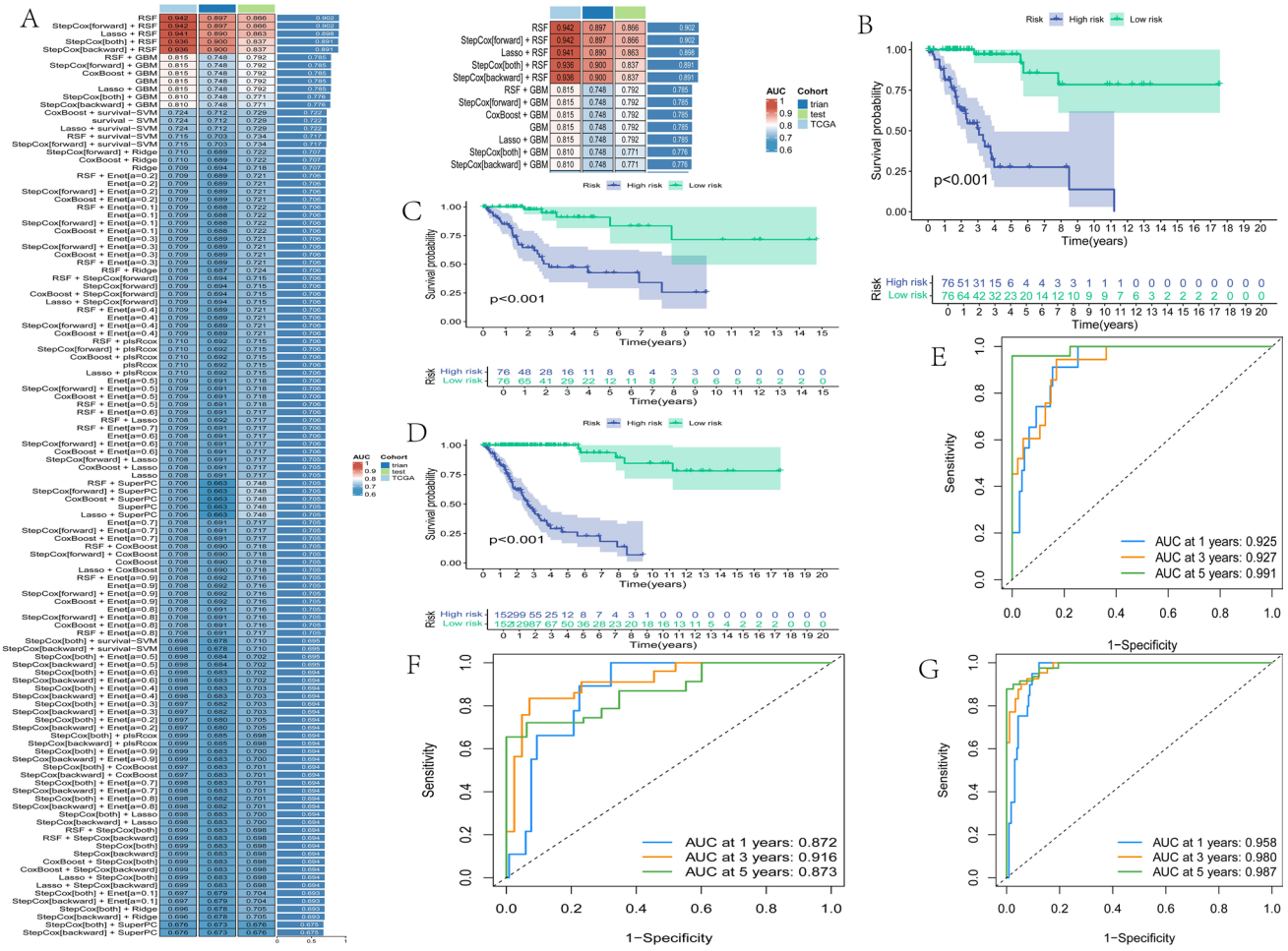

Fig. 6 Construction of a prognostic model based on RA-associated signature genes in CC. **A**: Screening of key prognostic genes using 101 machine learning algorithms. **B**–**D**: Kaplan–Meier survival analysis of high- vs. low-risk patients in **B** the training set, **C** the validation set, and **D** the TCGA cohort, demonstrating significant survival differences (log-rank $p < 0.05$). **E**–**G**: Time-dependent ROC curve analysis showing robust predictive performance (AUC > 0.7) in **E** the training set, **F** the validation set, and **G** the TCGA cohort

3.6 The nomogram based on risk score in CESC

To investigate the independent prognostic role of the risk score in cervical cancer alongside other clinical characteristics such as age, grade, and stage, we performed univariate and multivariate Cox regression analyses. These analyses revealed that the risk score independently predicted overall survival (OS) in cervical cancer, with hazard ratios of 1.153 (95% CI 1.125–1.181) and 1.161 (95% CI 1.130–1.192), respectively (Fig. 7A, B).Nomogram plots were constructed to predict 1-, 3-, and 5 year OS using TCGA data, incorporating risk score, age, grade, and stage as predictive parameters (Fig. 7C). The nomogram demonstrated high predictive accuracy, with survival rates estimated at 99.7%, 98.2%, and 97.4% for 1-, 3-,

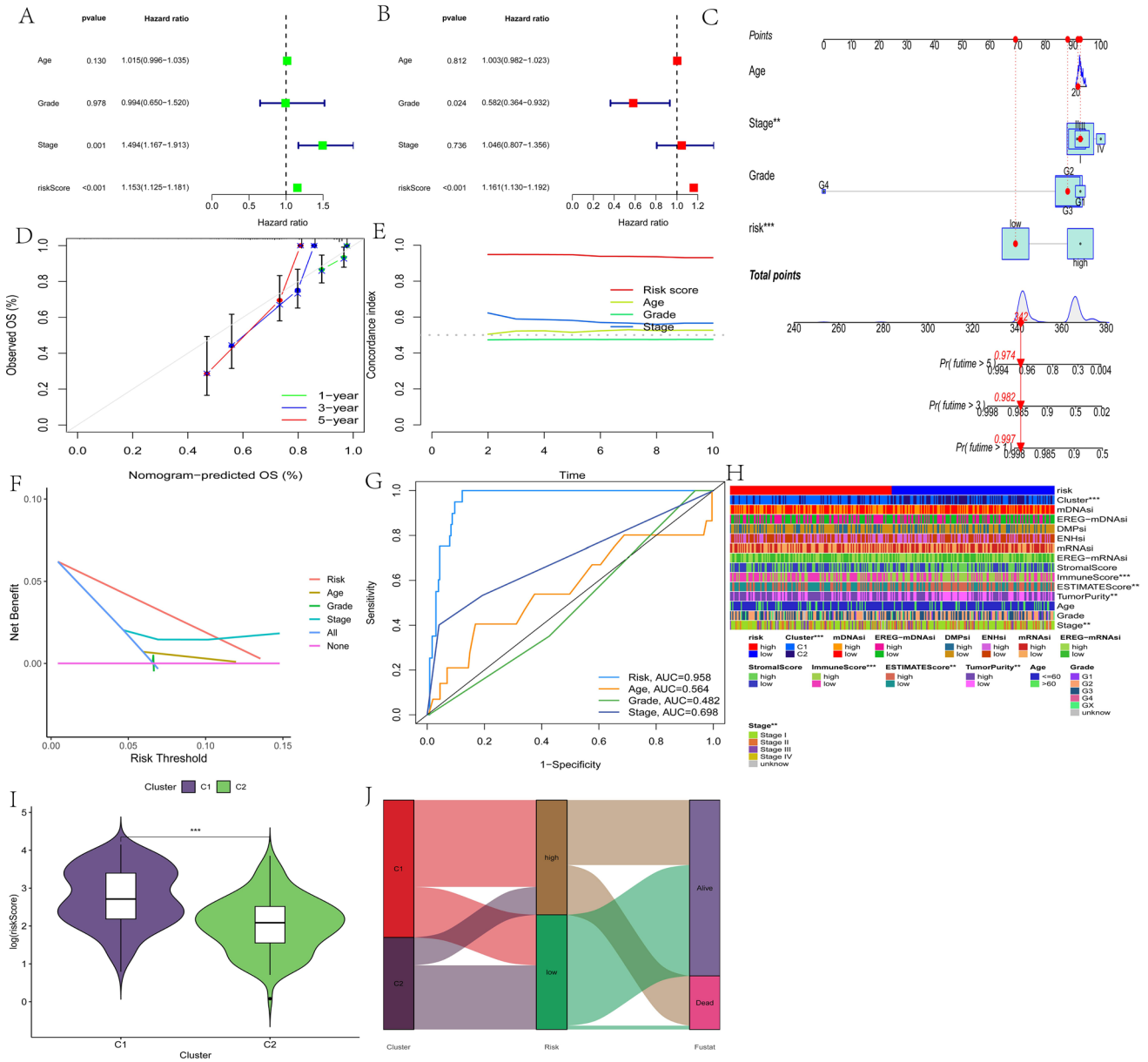

Fig. 7 Independent predictive analysis of the RA-associated prognostic model in cervical cancer. **A**–**B**: Cox regression analyses demonstrating the independent prognostic value of the risk score: **A** univariate and **B** multivariate analyses (incorporating risk score, tumor stage, grade, and age). **C**: Nomogram integrating risk score with clinical characteristics for survival prediction. **D**: Calibration curves showing agreement between predicted and observed 1-, 3-, and 5 year survival probabilities. **E**: C-index comparison of risk score versus other clinical features over time. **F**: Decision curve analysis (DCA) evaluating clinical utility. **G**: ROC curve analysis of model performance. **H**: Heatmap integrating multi-omics data with risk stratification. **I**: Violin plot comparing risk scores across molecular subtypes. **J**: Sankey diagram visualizing relationships between survival status, molecular subtypes, and risk groups

and 5 year periods, respectively. Calibration curves for 1-, 3-, and 5 year OS indicated good agreement between predicted and actual outcomes (Fig. 7D).Furthermore, the risk score exhibited the highest C-index, exceeding 0.9, indicating robust prognostic performance (Fig. 7E). Decision curve analysis (DCA) curves for 1-, 3-, and 5 year predictions reinforced the utility of the risk score as a prognostic tool (Fig. 7F and Supplementary Figure. S4).ROC curves integrating clinical characteristics illustrated that the risk score at 1, 3, and 5 years outperformed other variables, achieving approximately 0.9 AUC (Fig. 7G and Supplementary Fig. S4). A heatmap revealed significant clustering differences in stage, immune score, and stromal score between different risk groups (Fig. 7H).Analyzing risk score differences across clusters with varying RA-related scores, we observed that the C2 subgroup had lower risk scores (Fig. 7I). Additionally, a Sankey diagram depicted that C2, characterized by lower risk scores, corresponded to more favorable survival states (Fig. 7J).

3.7 Correlation between risk score and immune cell infiltration in cervical cancer

In the low-risk group, higher gene expression was observed in most of the HIL families (Fig. 8A). Various methods, including ssGSEA, were employed to enrich and analyze immune cells and their functions. The analysis indicated significant enrichment of almost all immune functions (except type II IFN response) in the low-risk group (Fig. 8B). Figure 8C illustrates that nearly all immune cells were infiltrated more prominently in the low-risk group compared to the high-risk group. Immune scoring also demonstrated higher scores for immune and stromal components in the low-risk group (Fig. 8D). Additionally, results from the CIBERSORT algorithm revealed extensive infiltration of various immune cells, particularly CD8 T cells, in the low-risk group (Fig. 8E).

3.8 Correlation between immunotherapy efficacy and risk scores in cervical cancer

Immunotherapy is emerging as a treatment modality for advanced cervical cancer patients. Our analysis of TIDE results revealed a higher immune response rate in the low-risk group (41%) compared to the high-risk group (33%) (Fig. 9A–B). In addition, the low-risk group showed lower tumor immune dysfunction and exclusion scores, whereas

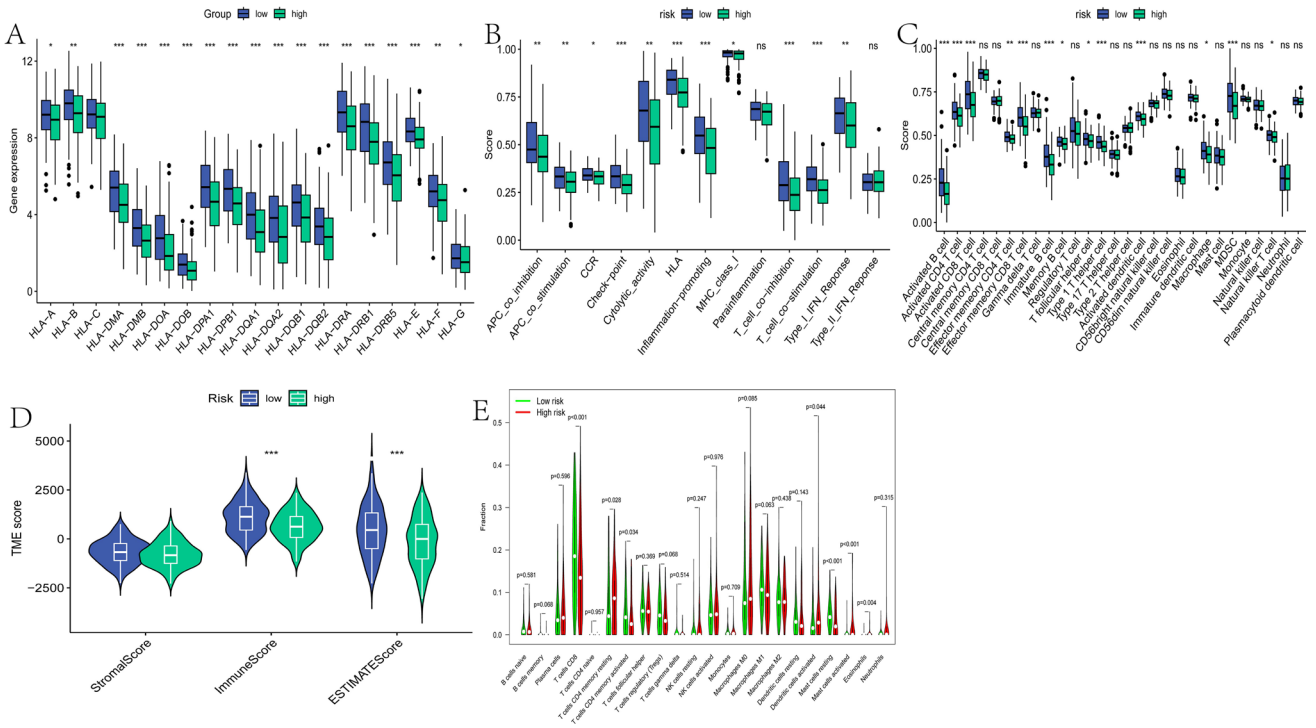

Fig. 8 Immune correlation analysis of the risk score model in cervical cancer. **A**: Boxplot showing differential expression of HLA-related genes between high- and low-risk groups, with most genes significantly upregulated in the low-risk group. **B**: Functional analysis of immune activity, revealing enhanced immune functions (e.g., CCR signaling) in the low-risk group. **C**: ssGSEA analysis demonstrating distinct immune cell infiltration patterns between risk groups. **D**: Comparative analysis of immune scores (e.g., ESTIMATE, stromal, immune scores) across risk groups. **E**: CIBERSORT analysis quantifying differential infiltration levels of specific immune cell subsets between risk groups

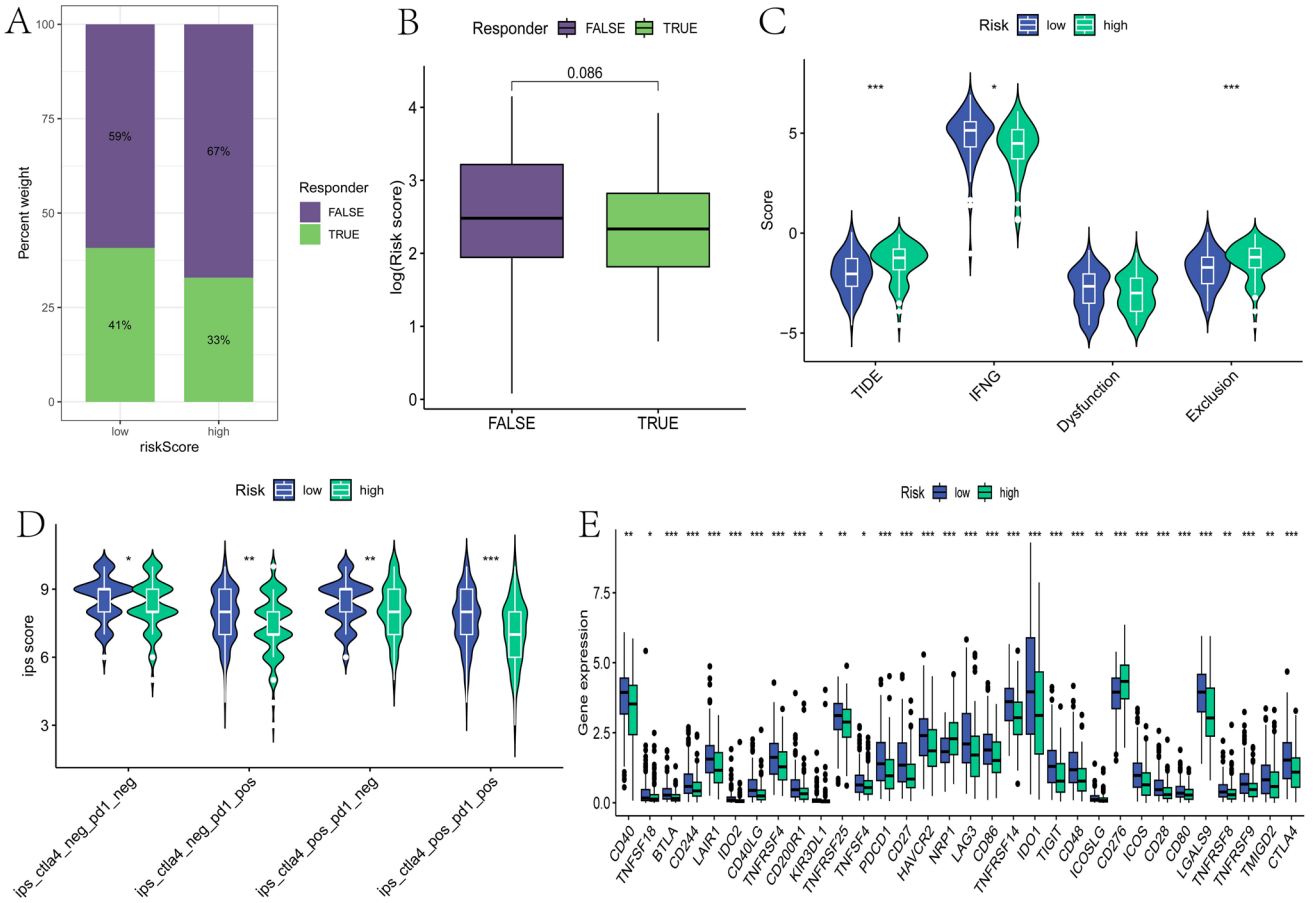

Fig. 9 Evaluation of immunotherapy response in patients with different risk scores. **A**–**C**: Analysis of Tumor Immune Dysfunction and Exclusion (TIDE) scores: **A** Bar plot comparing immunotherapy response rates between high- and low-risk groups; **B** Boxplot showing risk score distribution in responders vs. non-responders; **C** Violin plot comparing TIDE and related functional scores (e.g., dysfunction, exclusion) between risk groups. **D**: Violin plot of Immunophenoscore (IPS) analysis, revealing significantly higher scores in low-risk patients ($p < 0.01$). **E**: Boxplot of immune checkpoint inhibitor-related genes (e.g., PD-1, CTLA-4, LAG-3), with most genes upregulated in the low-risk group

dysfunction scores and interferon gamma scores exhibited reverse trends (Fig. 9C).Given the potential efficacy of immune checkpoint inhibitors (ICIs) in blocking CTLA4/PD-1 interactions to treat certain tumors, IPS scores were used to assess therapeutic potential based on IFNG expression levels. Notably, patients in the low-risk group demonstrated higher IPS scores across multiple assessments (Fig. 9D).Fig. 9E highlights significant differences in expression of checkpoint genes between the two groups, with higher expression observed in the low-risk group. This includes several well-known immunotherapeutic targets such as PDCD1 (PD-1), CD274 (programmed death ligand 1, PD-L1), and CTLA4.

3.9 Correlation of somatic mutations and risk scores in cervical cancer

Tumor mutations play a critical role in affecting patient prognosis. As depicted in Fig. 10A and B, the mutation frequency was higher in the low-risk group (89.36%) compared to the high-risk group (79.72%). Analysis of mutated genes between these subgroups revealed consistency in the top five mutated genes across both groups.Comparing tumor mutation burden (TMB) values between different risk scores, we observed that TMB values tended to be higher in patients with low-risk scores, although this difference did not reach statistical significance (Supplementary Fig. S5). Additionally, Fig. 10C illustrates that patients in the high TMB group exhibited better survival rates compared to those in the low TMB group.Furthermore, Fig. 10D indicates that patients in the high-risk and low-TMB groups had the poorest prognosis.

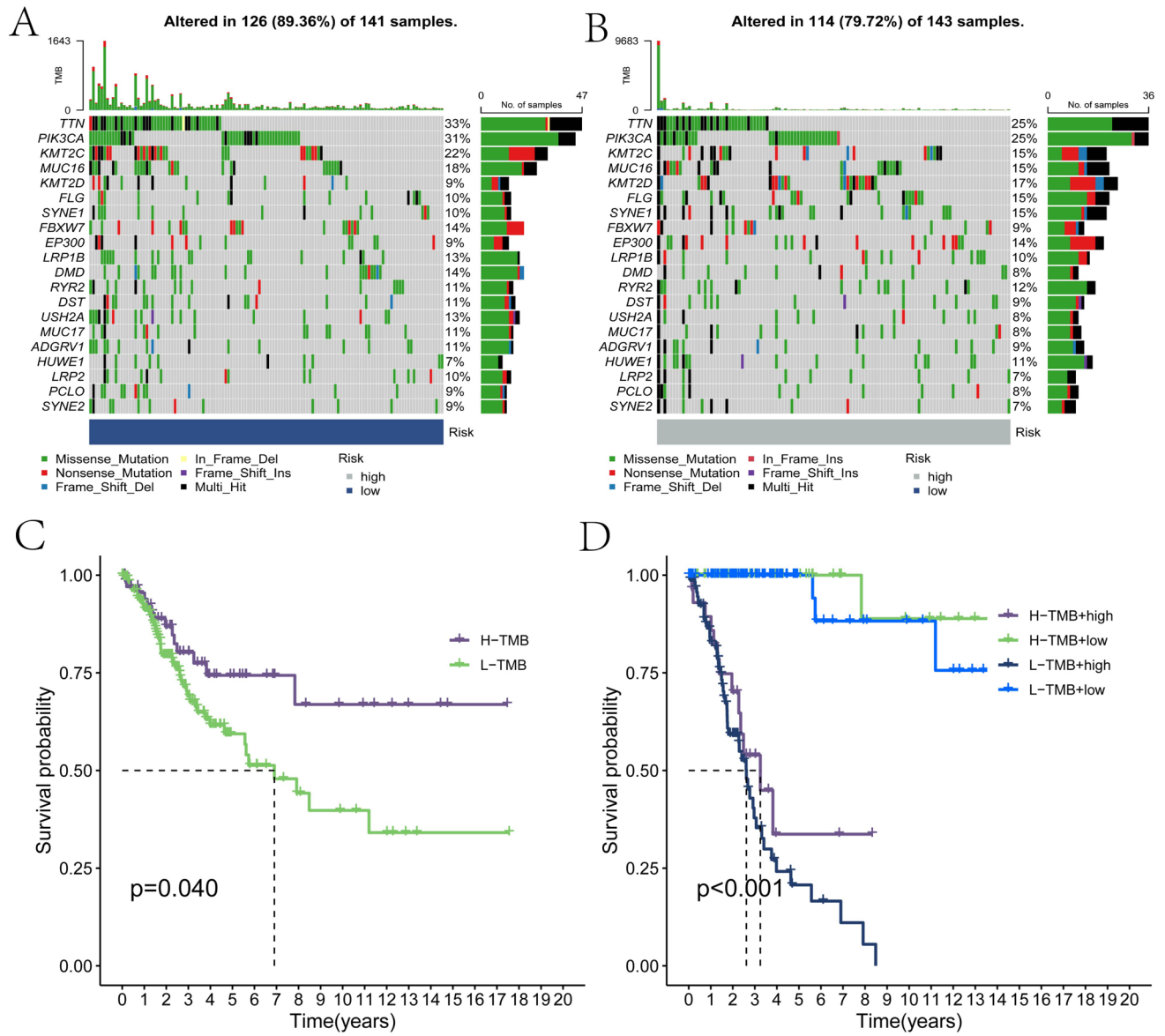

Fig. 10 Tumor mutation burden (TMB) analysis in cervical cancer risk groups. **A**–**B**: Waterfall plots depicting mutation profiles of high-risk **A** and low-risk **B** patients, with low-risk patients showing higher overall mutation rates. **C**–**D**: Kaplan–Meier survival curves stratified by TMB levels (high vs low) and risk groups, demonstrating superior survival in high-TMB/low-risk patients

3.10 Risk scores provide excellent prediction of immunotherapy response and drug treatment

In the IMvigor-210 cohort, we compared the restricted mean survival (RMS) at 6 and 12 months between the two groups to account for the delayed clinical effect of immunotherapy, and assessed the difference in long-term survival after 3 months of treatment ($p < 0.05$; Fig. 11A–B). The results indicated that patients in the low-risk group had better prognosis, suggesting a greater benefit from immunotherapy. Furthermore, the distribution of risk scores among patients with different levels of treatment response showed that the risk scores were significantly lower in the response group (complete remission [24]/partial remission [PR]) compared to the nonresponse group (progressive disease [PD]/stable disease [SD]) ($p < 0.05$; Fig. 11C).Subsequently, we validated these findings across multiple immunotherapy validation cohorts with prognostic information. Consistently, the low-risk groups demonstrated better prognostic outcomes in post-immunotherapy populations in GSE78220 ($p < 0.001$) (Fig. 11D), GSE135222 ($p < 0.001$) (Fig. 11E), and tended to have better immunotherapy outcomes in GSE91061 ($p = 0.018$; Fig. 11F). Finally,

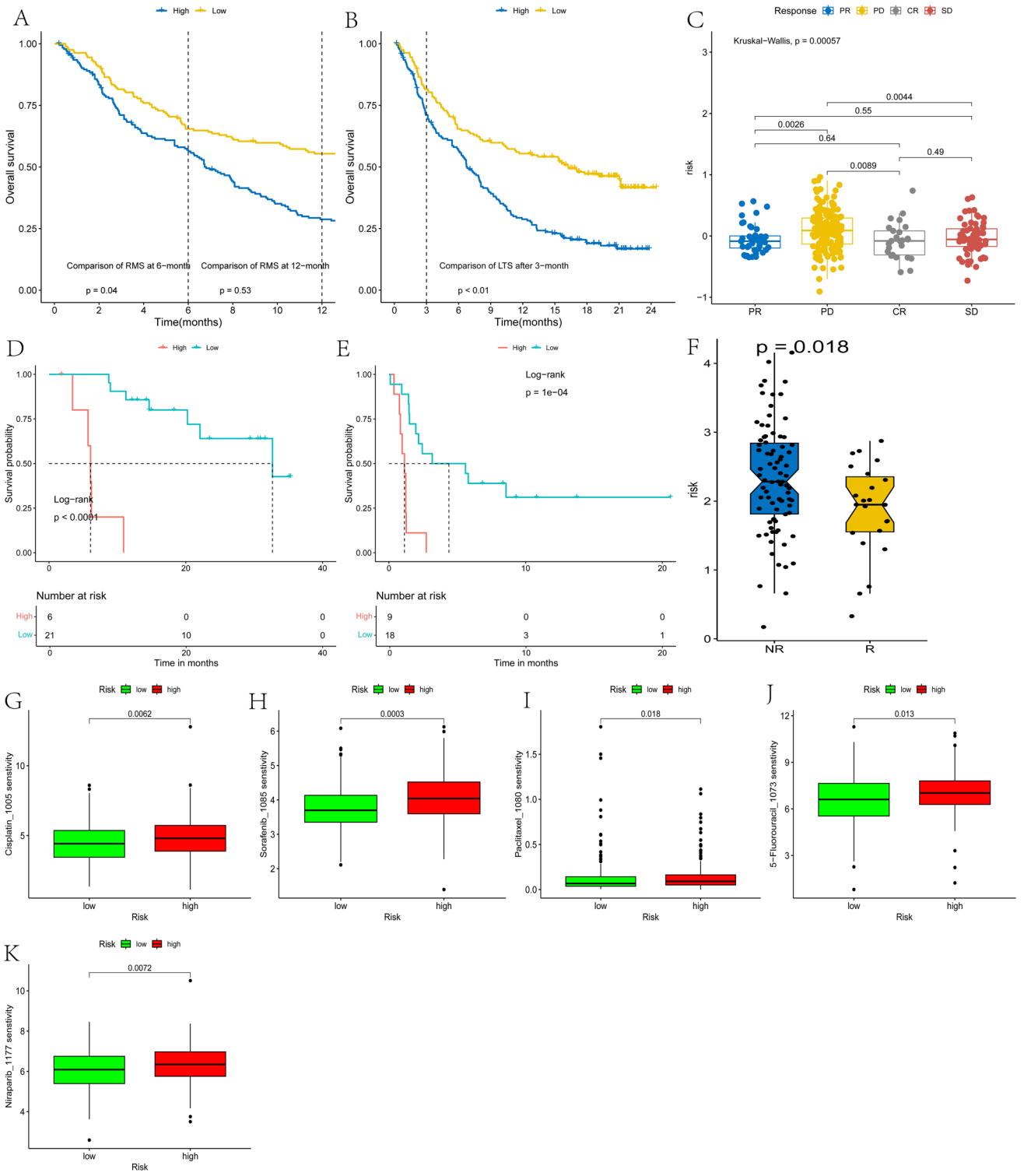

Fig. 11 Multi-dataset validation of risk score associations with immunotherapy and chemotherapy response. **A**–**B**: Immunotherapy survival analysis. **C**: Boxplot of risk scores in patients with differential immune response status (responders vs. non-responders, $p < 0.001$). **D**–**E**: Prognostic relevance of risk scores. Figure 11F: Phase diagram comparing risk scores between immunotherapy-sensitive and -resistant patients ($p < 0.001$). **G**–**K**: Chemotherapeutic drug sensitivity analysis: IC50 values for **G** cisplatin, **H** paclitaxel, **I** 5-FU, **J** doxorubicin, and **K** gemcitabine, with low-risk patients showing significantly lower IC 50 values ($p < 0.05$ for all)

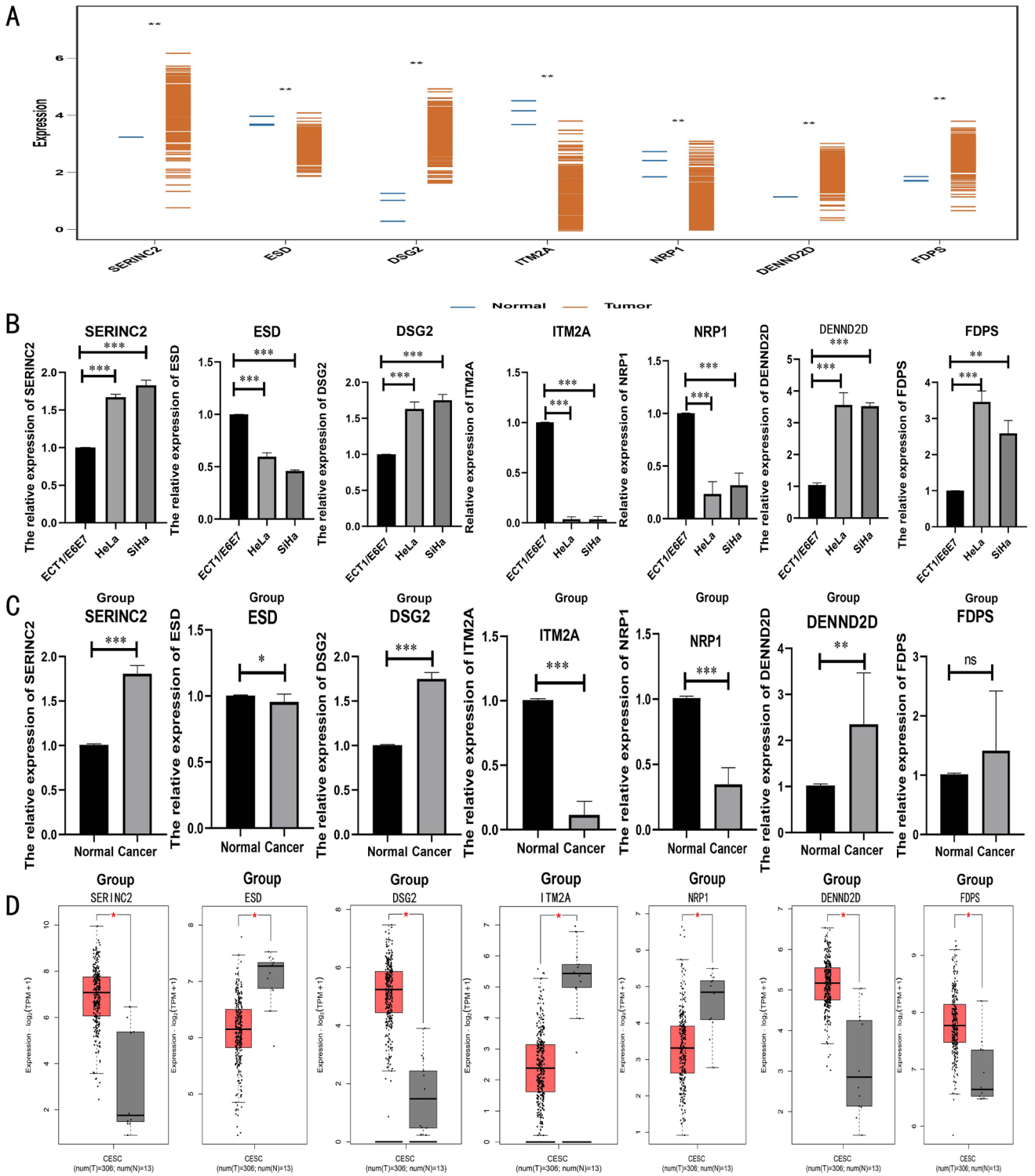

Fig. 12 Experimental and bioinformatic validation of model genes. **A**: Differential expression analysis of model genes in TCGA cervical cancer cohort. **B**: qPCR validation in cervical cancer cell lines (HeLa, SiHa) versus normal cervical epithelial cells, showing consistent overexpression of model genes. **C**: Tissue-level validation via qRT-PCR in 10 paired clinical samples. **D**: Cross-platform validation using an independent online database to demonstrate abnormal expression of cervical cancer model genes

we compared the IC50 values of drugs such as cisplatin, paclitaxel, and niraparib, and consistently found lower IC50

values in the low-risk group, suggesting increased sensitivity to these drugs (Fig. 11G–K).

3.11 Validation of model gene expression in cervical cancer

Figure 12A results showed high expression of SERINC2, DSG2, DENND2D, and FDPS in TCGA cervical cancer data, whereas the expression levels of ESD, NRP1, and ITM2A were relatively low. Then the same trend was validated for cervical cancer cell lines (Fig. 12B). In addition, in clinical tissues, we found that the differences in other genes were consistent with the above results, except for FDPS, which showed no significant differences (Fig. 12C). Validation using the GEPIA 2 online database confirmed high expression of SERINC2, DSG2, DENND2D, and FDPS in cervical cancer, whereas ESD, NRP1 and ITM2A showed contrasting lower expression levels (Fig. 12D). Consistent results were further validated using the HPA protein database (Fig. S6).

4 Discussion

Cervical cancer ranks fourth in cancer-related mortality worldwide and is the fourth most common cancer among women globally [25]. It is predominantly caused by persistent Human Papilloma Virus (HPV) infections [26]. In recent years, increased risks of HPV infection and cervical dysplasia have been reported in patients with systemic inflammatory diseases, including rheumatoid arthritis (RA) [27]. However, the genetic covariation between RA and cervical cancer remains unclear. Our study aims to investigate and uncover potential genetic connections between cervical cancer and RA, shedding light on their potential associations at the genetic level.

MR provides an alternative method to probe causality in epidemiological research by utilizing genetic variants presumed to satisfy instrumental variable (IV) assumptions [28]. After conducting sensitivity analyses such as the Multiple Validity Test [29] and the Cochrane Q-test [30], MR analysis suggested an increased risk of cervical cancer in patients with RA. However, further studies are needed to explore the relationship between gene-level changes and cervical cancer in RA patients.

Therefore, we conducted further functional enrichment analysis on RA-related differential genes to explore their potential biological functions. A total of 43 RA-related genes were found to be up-regulated, while 30 RA-related genes were down-regulated. Using multiple scoring modalities for RA-associated differential genes in single-cell scoring of cervical cancer, we observed that RA-associated scores were predominantly enriched in NK cells. Cellular communication analysis suggested that NK cells exhibit strong signaling interactions with fibroblasts and epithelial cells. Some studies have indicated that NK cells play a crucial role in recognizing and eliminating cervical tumor cells as well as cells infected by HPV virus [24]. NK cells can also influence the expression of E6 and E7 and regulate interferon production via IL-8 [31, 32]. Therefore, we speculate that the increased risk of cervical carcinogenesis in RA patients may be attributed to alterations in NK cells. The genetic interplay between autoimmune diseases and cervical cancer extends beyond RA. Systemic lupus erythematosus (SLE), another chronic autoimmune condition, has been associated with altered HPV clearance mechanisms and increased cervical cancer risk [33]. Unlike RA, which primarily involves Th1-mediated responses, SLE is characterized by dysregulated B-cell activity and type I interferon signatures [34]. This distinction suggests that RA-specific mechanisms might involve unique genetic variants affecting NK cell function and T-cell regulation, rather than the humoral immunity alterations observed in SLE [35].

We further conducted a prognostic analysis based on the differential genes associated with RA-related scores, dividing cervical cancer patients into two subgroups. The C2 subgroup exhibited higher immune scores, stromal scores, and more abundant immune cell infiltration. Importantly, patients in the C2 subgroup showed better prognosis. It is well-established that immune cell infiltration in the tumor microenvironment closely correlates with patient prognosis [36]. Collectively, these findings suggest that the heightened susceptibility to cervical cancer in RA patients might be linked to alterations in their immune cell profiles.

Next, we conducted prognostic modeling of RA score-related genes using 101 combinations of 10 machine learning algorithms. Cervical cancer patients were stratified into high-risk and low-risk groups based on their risk scores. Analysis of patients in different risk groups revealed that those in the low-risk group exhibited better prognosis and richer immune cell infiltration. Furthermore, HLA-related genes and most immune checkpoint inhibitor genes were expressed at higher levels in the low-risk group, which also showed lower TIDE scores. Higher TME scores in the low-risk group suggested a greater potential for triggering an immune response by the tumor [37]. The discovery of immune checkpoints such as CTLA-4 and PD-1 has revolutionized cancer immunotherapy, particularly in solid tumors [38]. Encouragingly, anti-CTLA-4

and anti-PD-1 immune checkpoint inhibitors have demonstrated significant efficacy in cervical cancer [39]. Regulatory T cells' immunosuppressive role contributes to immune evasion in cervical cancer [40]. Recent studies have implicated γδ T cells producing IL-17a in increased angiogenesis during the development of HPV-induced squamous cell carcinoma [41]. In the context of RA, overexpression of IL-18 and a relative deficiency of IL-18 binding protein levels are associated with disease pathogenesis [42]. IL-18, produced by various immune and non-immune cells, aberrantly activates CD4 + T cells, primarily by inducing IFN-γ expression in CD4 + and CD8 + T cells [43, 44]. Elevated tumor mutation burden (TMB) can activate immune cells, potentially enhancing responses to immunotherapy [45]. Combining TMB with risk scores revealed higher TMB levels in the low-risk group. Comparative analysis of overall survival across TMB subgroups indicated better prognosis in patients with high TMB and low-risk scores, consistent with previous findings [46]. Therefore, we propose that the heightened risk of cervical cancer and poorer prognosis in RA patients primarily occur through immune pathways, influencing microenvironmental changes that promote immune evasion and related mechanisms.

The therapeutic landscape of RA, particularly the widespread use of TNF-α inhibitors, may influence cervical cancer risk through immunomodulatory mechanisms. TNF-α plays a dual role in carcinogenesis, exhibiting both tumor-suppressive and tumor-promoting effects depending on the microenvironment [47]. While TNF-α inhibitors effectively control RA disease activity, their impact on cervical cancer risk remains controversial. Some studies suggest that long-term TNF-α inhibition might impair immune surveillance against HPV-infected cells, potentially increasing cervical cancer risk [48]. Conversely, other reports indicate that proper control of systemic inflammation might reduce cancer-promoting cytokine cascades [49]. This dichotomy underscores the need for careful monitoring of RA patients receiving biologic therapies, particularly those with persistent HPV infection.

There are several limitations to this study. First, the database utilized in this research originated from the European region, there may be racial specificity leading to errors in the results, necessitating further clinical validation to ascertain the generalizability of the findings to other populations and ethnicities. Second, the source of single-cell data we applied was only five cases, in which there may be errors in the results due to the small amount of sample data. Finally, due to limited funding, our study could not validate findings on cellular or tissue levels, nor did it validate immune cell markers. It has to be acknowledged that we should follow up with studies targeting the mechanism of action of rheumatoid arthritis affecting cervical carcinogenesis through NK cells. While our study identifies potential genetic connections between RA and cervical cancer and highlights the role of immune pathways, particularly NK cells, in mediating this association, it is important to interpret these findings with caution. The modest effect sizes observed in our MR analyses (e.g., an odds ratio of 1.15) suggest that the genetic contribution to cervical cancer risk in RA patients, while statistically significant, may be small in magnitude. Such modest effects warrant careful consideration, as they could reflect statistical artifacts or confounding factors rather than biologically meaningful associations. Additionally, the gene signatures identified in this study, though informative, require further prospective and functional validation to confirm their relevance in clinical settings.

5 Conclusion

Our study suggests that patients with RA are at increased risk of developing cervical cancer. The increased incidence of cervical cancer in patients with RA is mainly through patient immune alterations, which facilitates strategies to improve individualized follow-up and personalized decision making.

Author contributions Yongjin Luo and Kaiyi Meng: data analysis, methodology, writing. Xifeng Xu: Data Curation, Writing—Original Draft.Yujie Huang, and Wu Wei: Revising the manuscript.Zhong Lin: Funding Acquisition, Resources.

Funding This study was partially funded by the self-funded project of Guangxi Health Commission. (Grant No. Z-A20231188) and Guangxi science and technology base and talent project (Grant no. AC22080002).

Data availability The original data used in this project can be downloaded in the public database GEO (https://www.ncbi.nlm.nih.gov/geo/) and TCGA (https://portal.gdc.cancer.gov/). The scRNA-seq data in the study have been uploaded to the Genome Sequence Archive (GSA) of the National Genomics Data Center (access number: PRJCA008573), including raw data accessible at https://ngdc.cncb.ac.cn/gsa-human and processed data available at https://ngdc.cncb.ac.cn/omix. The data analyzed during the human experiments are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate This study utilized human cervical tissue samples obtained and processed in accordance with the ethical principles of the Declaration of Helsinki. The original clinical sample collection was approved by the Ethics Committee of Nanning Second People's Hospital with proper informed consent procedures. As the current research exclusively employs de-identified data from public databases (specifically [GEO/SRA]), no additional ethical approval was required for this secondary analysis. Clinical trial registration: Not applicable.

Consent for publication Informed consent was obtained from all individual participants included in the study.

Competing interests The authors declare no competing interests.

References

1. Guo L, Hua K. Cervical cancer: emerging immune landscape and treatment. Onco Targets Ther. 2020;13:8037–47.
2. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. The Lancet. 2016;388(10055):2023–38.
3. Sharif K, Sharif A, Jumah F, et al. Rheumatoid arthritis in review: Clinical, anatomical, cellular and molecular points of view. Clin Anat. 2017;31(2):216–23.
4. Dougados M, Soubrier M, Antunez A, et al. Prevalence of comorbidities in rheumatoid arthritis and evaluation of their monitoring: results of an international, cross-sectional study (COMORA). Ann Rheum Dis. 2014;73(1):62–8.
5. Kim SC, Glynn RJ, Giovannucci E, et al. Risk of high-grade cervical dysplasia and cervical cancer in women with systemic inflammatory diseases: a population-based cohort study. Ann Rheum Dis. 2015;74(7):1360–7.
6. Simon TA, Thompson A, Gandhi KK, et al. Incidence of malignancy in adult patients with rheumatoid arthritis: a meta-analysis. Arthr Res Ther. 2015. https://doi.org/10.1186/s13075-015-0728-9.
7. Yuan S, Chen J, Ruan X, et al. Rheumatoid arthritis and risk of site-specific cancers: Mendelian randomization study in European and East Asian populations. Arthr Res Ther. 2022. https://doi.org/10.1186/s13075-022-02970-z.
8. Balasubramaniam SD, Balakrishnan V, Oon CE, et al. Key molecular events in cervical cancer development. Medicina. 2019;55(7):384.
9. *<prevalence and cervical human papil source ginecol obstet mex so 2008 jan 76 1 9 17.pdf>*.
10. Santana IU, et al. Systemic lupus erythematosus, human papillomavirus infection, cervical pre-malignant and malignant lesions: a systematic review. Clin Rheumatol. 2010;30(5):665–72.
11. Rojo-Contreras W, Olivas-Flores EM, Gamez-Nava JI, et al. Cervical human papillomavirus infection in Mexican women with systemic lupus erythematosus or rheumatoid arthritis. Lupus. 2011;21(4):365–72.
12. Bentham J, Morris DL, Graham DSC, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nat Genet. 2015;47(12):1457–64.
13. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–87.
14. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of single-cell data. Cell. 2019;177(7):1888–902.
15. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using cell chat. Nat Commun. 2021;12(1):1088.
16. Zhao S, Wang L, Ding W, et al. Crosstalk of disulfidptosis-related subtypes, establishment of a prognostic signature and immune infiltration characteristics in bladder cancer based on a machine learning survival framework. Front Endocrinol. 2023;14:1180404.
17. Liu Z, Liu L, Weng S, et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. Nat Commun. 2022;13(1):816.
18. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinform. 2013;14:7.
19. Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
20. Jiang P, Gu S, Pan D, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24(10):1550–8.
21. Mayakonda A, Lin DC, Assenov Y, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28(11):1747–56.
22. Hugo W, Zaretsky JM, Sun L, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. Cell. 2016;165(1):35–44.
23. Jung H, Kim HS, Kim JY, et al. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. Nat Commun. 2019;10(1):4278.
24. Gutiérrez-Hoya A, Soto-Cruz I. NK cell regulation in cervical cancer and strategies for immunotherapy. Cells. 2021;10:3104.

25. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63.
26. Johnson CA, James D, Marzan A, et al. Cervical cancer: an overview of pathophysiology and management. Semin Oncol Nurs. 2019;35(2):166–74.
27. Kim SC, Schneeweiss S, Liu J, et al. Biologic disease-modifying antirheumatic drugs and risk of high-grade cervical dysplasia and cervical cancer in rheumatoid arthritis: a cohort study. Arthritis Rheumatol. 2016;68(9):2106–13.
28. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: a review. Res Synth Methods. 2019;10(4):486–96.
29. Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50(5):693–8.
30. Hoaglin DC. Misunderstandings about Q and "Cochran's Q test" in meta-analysis. Stat Med. 2016;35(4):485–95.
31. Koromilas AE, Li S, Matlashewski G. Control of interferon signaling in human papillomavirus infection. Cytokine Growth Factor Rev. 2001;12:157–70.
32. Hoppe-Seyler K, Bossler F, Braun JA, et al. The HPV E6/E7 oncogenes: key factors for viral carcinogenesis and therapeutic targets. Trends Microbiol. 2018;26:158–68.
33. Infante V, Miyaji KT, Soarez PC, et al. Systematic review and meta-analysis of HPV vaccination in women with systemic lupus erythematosus (SLE). Expert Rev Vaccines. 2021;20(3):309–18.
34. Garcia-Carrasco M, Mendoza-Pinto C, Rojas-Villarraga A, et al. Prevalence of cervical HPV infection in women with systemic lupus erythematosus: a systematic review and meta-analysis. Autoimmun Rev. 2019;18(2):184–91.
35. Firestein GS, McInnes IB. Immunopathogenesis of rheumatoid arthritis. Immunity. 2017;46(2):183–96.
36. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017;541(7637):321–30.
37. Leon-Castillo A, Britton H, McConechy MK, et al. Interpretation of somatic POLE mutations in endometrial carcinoma. J Pathol. 2020;250(3):323–35.
38. Balar AV, Weber JS. PD-1 and PD-L1 antibodies in cancer: current status and future directions. Cancer Immunol Immunother. 2017;66(5):551–64.
39. Marabelle A, Fakih M, Lopez J, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. Lancet Oncol. 2020;21(10):1353–65.
40. Ni H, Zhang H, Li L, et al. T cell-intrinsic STING signaling promotes regulatory T cell induction and immunosuppression by upregulating FOXP3 transcription in cervical cancer. J Immunother Cancer. 2022;10(9):501.
41. Van Hede D, Polese B, Humblet C, et al. Human papillomavirus oncoproteins induce a reorganization of epithelial-associated gammadelta T cells promoting tumor formation. Proc Natl Acad Sci USA. 2017;114(43):E9056–65.
42. Verri WA, Cunha TM, Parada CA, et al. IL-15 mediates immune inflammatory hypernociception by triggering a sequential release of IFN-gamma, endothelin, and prostaglandin. Proc Natl Acad Sci USA. 2006;103(25):9721–5.
43. Gableh F, Saeidi M, Hemati S, et al. Combination of the toll like receptor agonist and alpha-galactosylceramide as an efficient adjuvant for cancer vaccine. J Biomed Sci. 2016;23:16.
44. Jing W, Gershan JA, Blitzer GC, et al. Adoptive cell therapy using PD-1(+) myeloma-reactive T cells eliminates established myeloma in mice. J Immunother Cancer. 2017;5:51.
45. Merino DM, McShane LM, Fabrizio D, et al. Establishing guidelines to harmonize tumor mutational burden (TMB): in silico assessment of variation in TMB quantification across diagnostic platforms: phase I of the friends of cancer research TMB harmonization project. J Immunother Cancer. 2020;8(1):206.
46. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet. 2019;51(2):202–6.
47. Balkwill F. TNF-alpha in promotion and progression of cancer. Cancer Metastasis Rev. 2006;25(3):409–16.
48. Curtis JR, Yamaoka K, Chen YH, et al. Malignancy risk with tofacitinib versus TNF inhibitors in rheumatoid arthritis: results from the open-label, randomised controlled ORAL surveillance trial. Ann Rheum Dis. 2023;82(3):331–43.
49. Diesler R, Cottin V. Pulmonary fibrosis associated with rheumatoid arthritis: from pathophysiology to treatment strategies. Expert Rev Respir Med. 2022;16(5):541–53.