Discover Oncology

Analysis

Identification of therapeutic targets in lung adenocarcinoma using Mendelian randomization and multi-omics

Yue Li[1] · Keru Ma[2] · Hao Wang[2] · Zongying Liu[3] · Zhuying Li[1]

Received: 27 February 2025 / Accepted: 27 May 2025
Published online: 07 June 2025

Abstract

Background Lung adenocarcinoma (LUAD) remains associated with limited effective pharmacological treatment options. This study aimed to identify potential therapeutic targets for LUAD through the integration and analysis of multi-omics datasets.

Methods A meta-analysis was conducted using two extensive proteomics datasets, the UK Biobank Proteomics Project (UKB-PPP) and the Fenland study, to identify disease-associated targets for LUAD through the Summary-Data-Based Mendelian Randomization method. Sensitivity analysis, including heterogeneity tests for dependent instruments, were conducted to validate the findings. The prognostic relevance of the identified candidate targets was assessed using transcriptomic data. Functional interactions were explored via protein–protein interaction network analysis, while single-cell analyses were employed to determine cell-specific expression patterns and differentiation trajectories. Potential side effects and therapeutic indications of these targets were evaluated using phenome-wide association studies and pharmacological data mining.

Results Following meta-analysis, a primary significant target, intercellular adhesion molecule 5 (ICAM5), along with potential targets FUT8 and KLK13, were identified as therapeutic candidates for LUAD. FUT8 demonstrated a positive association with LUAD risk (OR = 1.02, $p = 0.049$), while ICAM5 (OR = 0.88, $p = 0.002$) and KLK13 (OR = 0.85, $p = 0.021$) exhibited negative associations. ICAM5 was further identified as an independent prognostic factor for patient survival (HR: 0.788, 95% CI: 0.663–0.936, $p = 0.007$) and revealed significant diagnostic and prognostic utility in LUAD. ICAM5 expression correlated with various immune infiltration patterns, suggesting potential modulation of the tumor immune microenvironment. Single-cell analysis revealed that ICAM5 did not directly impact LUAD cell differentiation, though its downstream target, MUC1, may contribute to differentiation processes, particularly in KRAS-mutated LUAD. Furthermore, phenome-wide association studies did not reveal substantial evidence of adverse phenotypes linked to ICAM5, supporting its safety profile for drug development.

Conclusion ICAM5 emerges as a promising biological marker with significant prognostic and therapeutic potential in LUAD.

Keywords Drug target · ICAM5 · LUAD · Mendelian randomization · Prognosis

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12672-025-02835-2.

✉ Zhuying Li, lizhuying_6808@163.com | [1]Department of Respiratory Medicine, First Affiliated Hospital of Heilongjiang University of Chinese Medicine, No. 26 of Heping Road, Xiangfang District, Harbin 150040, China. [2]Department of Breast Surgery, Harbin Medical University Cancer Hospital, Harbin 150081, China. [3]Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin 150081, China.

Abbreviations

LUAD	Lung adenocarcinoma
UKB-PPP	UK Biobank Pharma Proteomics Project
SMR	Summary-data-based Mendelian Randomization
HEIDI	Heterogeneity In Dependent Instruments
PPI	Protein-protein interaction
PheWAS	Phenome-wide association study
MR	Mendelian randomization
GWAS	Genome-Wide Association Studies
FDR correction	False discovery rate correction
pQTL	The protein quantitative trait loci
MAF	Minor Allele Frequency
SNP	Single nucleotide polymorphism
eQTL	Expression quantitative trait Loci
TCGA	The Cancer Genome Atlas
HR	Hazard ratios
KM	Kaplan-Meier
OS	Overall survival
DCA	Decision curve analysis
c-index	The concordance index
time-ROC	Time-dependent receiver operating characteristic
KEGG	Kyoto Encyclopedia of Genes and Genomes
GSEA	Gene Set Enrichment Analysis
GO	Gene Ontology
BP	Biological process
CC	Cellular component
MF	Molecular function
BC	The betweenness centrality
AUC	Area Under Curve
ICAM5	Intercellular adhesion molecule-5
MUC1	Mucins 1
CD8 + T cells	Cytotoxic T Lymphocytes
FUT8	Fucosyltransferase 8provided by HGNC
KLK13	Kallikrein related peptidase 13
KACs	Lysine actylation
SFTPB	Surfactant protein B
KDELC2	KDEL (Lys-Asp-Glu-Leu) containing 2
CLDN4	Claudin-4
MYC	The v-myc avian myelocytomatosis viral oncogene homolog
EMT	Epithelial-mesenchymal transition
KRAS	Kirsten rat sarcoma viral oncogene homolog
NSCLC	Non-small cell lung cancer

1 Introduction

Lung adenocarcinoma (LUAD) represents one of the most prevalent pathological subtypes of non-small cell lung cancer (NSCLC) [1]. Despite surgical intervention being the cornerstone of treatment, LUAD continues to pose a substantial global health challenge, with a five-year survival rate of merely 22% [2, 3]. This underscores the critical need for effective therapeutic strategies beyond surgery. Moreover, the advanced stages of LUAD are often characterized by treatment resistance, further complicating disease management [4]. Identifying key molecular drivers of LUAD is therefore essential for improving early detection and therapeutic interventions.

Advancements in bioinformatics and molecular database development have significantly enhanced the potential for target discovery in various diseases, providing valuable insights into diagnosis and treatment. While single-omics approaches, such as single-cell analysis and weighted network analysis, have contributed to elucidating the pathogenesis of LUAD, the complex and heterogeneous molecular characteristics of the disease necessitate a more integrative approach [5–7]. Single-omics studies often fail to fully capture the multifaceted molecular mechanisms underlying LUAD. In contrast, recent multi-omics investigations have provided comprehensive prognostic insights by integrating gene expression, protein abundance, and methylation data [8]. Such approaches have proven effective in cancer research, including studies on breast and colorectal cancers [9, 10].

Mendelian randomization (MR) is a robust analytical technique that uses genetic variations to explore causal relationships between exposures and outcomes. Summary data-based Mendelian randomization (SMR) and heterogeneity tests for heterogeneity in dependent instruments (HEIDI) offer enhanced statistical power and have been extensively applied in drug target development [11, 12]. Emerging studies have highlighted the regulatory effects of gene expression loci on lung cancer risk [13]. However, the phenotypic regulatory influence of genes is often overshadowed by changes in protein abundance. Given the significant role of protein levels in LUAD pathogenesis and the limited research on multi-omics target exploration, MR offers a promising avenue for identifying disease-driving targets in LUAD.

This study utilized two extensive proteomic datasets, the UK Biobank Proteomics Project (UKB-PPP) and the Fenland study, as exposure variables to explore potential disease-driving targets in LUAD using the SMR method. Furthermore, multi-omics data were integrated to provide a comprehensive evaluation of the identified targets, elucidating their roles in the development and prognosis of LUAD.

2 Materials and methods

2.1 Research process

Protein quantitative trait locus (pQTL) data were obtained from two large-scale proteomics studies, the UKB-PPP and the Fenland study. A meta-analysis was performed using genome-wide association study (GWAS) data from these datasets to identify significant proteins associated with LUAD. Identified proteins were subject to false discovery rate (FDR) correction to account for multiple testing and reduce the likelihood of false positives. Subsequently, transcriptomic analysis was conducted on all significant targets to comprehensively assess their impact on prognosis.

2.2 Data source

pQTL data were derived from the UKB-PPP and the Fenland study [14, 15]. These studies conducted proteomic analyses on plasma samples from 54,219 and 10,708 participants of European ancestry, respectively, generating datasets comprising of 2923 and 4775 proteins [14, 15]. GWAS data for LUAD were obtained from the Transdisciplinary Research in Cancer of the Lung consortium, encompassing 11,245 LUAD cases and 54,619 controls, with a total of 10,345,176 single nucleotide polymorphisms (SNPs). RNA sequencing data for LUAD were sourced from TCGA (https://www.cancer.gov/ccg/research/genome-sequencing/tcga).

2.3 SMR analysis

For proteins whose cis-regions did not yield significant top SNPs, the top SNP selection method from the original UKB-PPP and Fenland studies was used, as outlined in prior reports [14, 15]. This method enabled the identification of significant top SNPs for all proteins.

In the SMR analysis, SNPs with a minor allele frequency below 0.01 were excluded. Additional filters included a maximum allele frequency difference threshold of 0.1 and a requirement that no more than 5% of SNPs exceeded this threshold. In the HEIDI test, highly correlated SNPs ($r^2 > 0.9$) with the primary QTL were excluded, as were SNPs with minimal correlation ($r^2 < 0.05$) to the primary eQTL. Default settings in SMR version v1.3.1 (https://yanglab.westlake.edu.cn/software/smr/#SMR) were applied for all analyses.

In the sensitivity analysis, we first performed a meta-analysis on the UKB-PPP and Fenland datasets, and then applied the Benjamini–Hochberg method to adjust the p-values and minimize false positive results, controlling the false discovery

rate (FDR) in multiple hypothesis testing. Secondly, considering the HEIDI test, a p-value greater than 0.01 indicates that there is no heterogeneity in the current results. Therefore, the criteria for selecting significant targets were defined as FDR meta *P*-value < 0.05 and HEIDI *P*-value > 0.01.

2.4 Prognostic evaluation

After identifying significant targets using the SMR method, these targets were hypothesized to potentially influence the risk of LUAD. Survival analysis was subsequently conducted on the identified targets using data from TCGA.

Cox regression models were applied to assess the prognostic impact of these candidate targets on LUAD, with HR and 95% CI used to quantify their effects on prognosis. Kaplan–Meier survival curves were generated to examine the association between the expression levels of these targets and OS.

To further analyze and visualize prognostic models, the rms package was used for nomogram development. Calibration curves were constructed to compare predicted probabilities with observed outcomes, providing an assessment of model accuracy. Additionally, decision curve analysis was conducted to assess the net clinical benefit of the prognostic model.

2.5 Gene expression analysis

For the core targets identified, differential expression between tumor and normal tissues was analyzed using data from TCGA database. Restricted cubic splines were used to visualize the relationship between the expression levels of candidate targets and the risk of mortality in LUAD [16]. Spearman's rank correlation coefficient was used to assess the correlations among continuous variables.

2.6 Construction of prognostic models for core targets

Core targets were identified by integrating findings from SMR and prognostic analyses. For these core targets, multivariate Cox regression analysis was conducted. A nomogram was developed, incorporating both the core targets and relevant clinical pathological features based on the results of the multivariate Cox model.

A risk score was generated using the coefficients from the Cox model, allowing for individualized risk assessment. The performance of the model was evaluated using the concordance index (c-index), calibration plots, and time-dependent receiver operating characteristic (time-ROC) curves [17].

2.7 Functional enrichment

Functional enrichment analysis was performed for the candidate targets to elucidate their biological roles and interactions. Friends analysis was conducted using the GOSemSim package to assess the importance of each target based on network topology parameters. Candidate targets were ranked and visualized according to their gene similarity scores. Differential expression analysis of the core targets was conducted using the limma package (version 3.40.6) to identify differentially expressed genes, with thresholds set at an adjusted *p*-value < 0.05 and a fold change > 2.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed to annotate the core targets, encompassing the GO categories of biological processes (BP), cellular components (CC), and molecular functions (MF). To further elucidate the biological relevance of the core targets, Gene Set Enrichment Analysis (GSEA) for KEGG pathways was conducted. Additionally, Disease ontology (DO) enrichment analysis was carried out using the DOSE package to identify associations between the core targets and specific diseases. Structural domain analysis was conducted to examine the domains of the core proteins.

Protein–protein interaction (PPI) network analysis for the differentially expressed genes of the core targets was conducted using the STRING database (https://cn.string-db.org/). The betweenness centrality (BC) of the core genes within the PPI network was calculated using the cytoNCA module in Cytoscape (version 3.10.1) to identify key genes.

2.8 Immunoprofiling

The correlation between the core targets and immune infiltration was investigated using multiple algorithms, including xCell, MCP-counter, QuanTIseq, and CIBERSORT [18–21]. The relationships between the core targets and the Immune Score, Stromal Score, and ESTIMATE Score were further evaluated using the ESTIMATE algorithm [22].

Given the relevance of immune checkpoint therapies in LUAD, the association between the core targets and key immune checkpoint genes was analyzed. These included *CD274*, *CTLA4, HAVCR2*, *LAG3*, *PDCD1*, *PDCD1LG2*, *SIGLEC15*, and *TIGIT* [23].

2.9 Single-cell analysis

Single-cell RNA sequencing (scRNA-seq) data from Jiang et al. (GSE189487) were used for single-cell analysis [24]. The raw scRNA-seq data were processed using the Seurat package (v5.1.0). Cells were retained based on the following criteria: more than 2000 unique molecular identifiers, more than 500 discovered genes, and less than 5% mitochondrial gene content. Cluster annotations were performed with reference to prior studies and the CellMarker database (http://xteam.xbio.top/CellMarker/) [25].

In LUAD development, the primary focus was on identifying benign and malignant cell populations, including alveolar intermediate cells (KRT8 + , KACs) and malignant cells [25]. Pseudotime analysis was subsequently conducted to assess the potential role of candidate targets in cell differentiation, using Monocle3 (v1.3.7).

2.10 Phenome-wide Mendelian Randomization (pheWAS) analysis

To investigate causal relationships between identified targets and other diseases, thereby assessing potential side effects or therapeutic indications, the methodology described by Zhou et al. was used. This approach employed SAIGE (Scalable and Accurate Implementation of Generalized Mixed Models) to analyze over 1400 binary phenotypes derived from 408,961 participants of European descent in the UK Biobank [26]. FDR correction was applied to adjust the significance thresholds for all results, ensuring the robustness of the findings.

2.11 Pharmacological search

A pharmacological search was conducted for the candidate targets using traditional Chinese medicine databases, including the Herb database (http://herb.ac.cn/) and the TCMBank database (https://tcmbank.cn/). The objective of this approach was to comprehensively explore the potential involvement of these targets in drug development and identify relevant compounds for therapeutic applications.

2.12 Cell lines

The cell lines HCC827 (purchased from PnoScience, catalog number CL-0094) and NCI-H1975 (purchased from Cyagen, catalog number iCell-h156) were used in this experiment. HCC827 cells were cultured in RPMI 1640 medium (Corning, catalog number 10–040-CVB), while NCI-H1975 cells were cultured in DMEM (Corning, catalog number 10-013-CVR), both supplemented with 10% fetal bovine serum (Ausbian, catalog number VS500T) and 1% Penicillin–Streptomycin (Gibco, catalog number 15140122). The cells were cultured at 37 °C in a 5% CO2 incubator, with regular medium changes and monitoring of cell growth. When the cells reached 70%-80% confluency, they were passaged. All cell cultures were maintained in a 37 °C, 5% CO2 humidified incubator, and mycoplasma contamination was regularly tested.

2.13 qPCR

Total RNA was extracted from cells using Trizol reagent (Sigma, Cat# T9424-100m) following the manufacturer's instructions. Briefly, cells were lysed with Trizol, and after phase separation with chloroform, the RNA was precipitated with isopropanol, washed with 75% ethanol, and resuspended in RNase-free water. RNA concentration and quality were assessed using a Nanodrop 100 spectrophotometer (Thermo, Cat# 2000/2000C). For cDNA synthesis, 1 µg of RNA was

reverse transcribed using Hiscript QRT Supermix for qPCR (+ gDNA WIPER) (Vazyme, Cat# R123-01). The cDNA was stored at – 80 °C. Real-time PCR was performed using AceQ qPCR SYBR Green Master Mix (Vazyme, Cat# Q111-02) with 5 µL mastermix, 0.25 µL of each primer (10 µM), 0.2 µL Dye2, 2 µL cDNA, and 2.3 µL RNase-free H2O. The following primer sequences were used:

ICAM5: (Upstream primer: GGGGCAGATGGTGACAGTAAC; Downstream primer: GTCGTTCTCGGTGGCATTTAG; Amplified fragment size: 121 bp). β-actin: (Upstream primer: CAAAGTTCACAATGTGGCCGAGGA; Downstream primer: GGGACTTCC TGTAACAACGCATCT; Amplified fragment size: 93 bp). PCR was conducted with an initial denaturation at 95 °C for 15 min, followed by 40 cycles of 95 °C for 15 s, 60 °C for 15 s, and 72 °C for 30 s, with a melting curve analysis. Relative gene expression was calculated using the 2^(– △△Ct) method, where △Ct is the difference between the target gene and reference gene Ct values, and – △△Ct represents the difference between the NC group and each sample.

2.14 Cell proliferation assay

For the cell viability assay using the CCK-8 kit, LUAD cells were seeded into 96-well plates at a density of 5 × 10^3 cells per well. After 2–4 h of incubation at 37 °C and 5% CO2, the culture medium was replaced with fresh medium containing 10% CCK-8 reagent at various time points. The cells were incubated with the CCK-8 reagent for 2 h, and the absorbance at 450 nm was measured using a microplate reader. A standard curve was constructed using cell counts and corresponding OD values. To assess cell proliferation, the cell viability was calculated using the formula: **Cell viability (%) = [(A(experiment)—A(blank)) / (A(control)—A(blank))] × 100, where A(experiment) is the absorbance of wells containing cells, CCK-8 solution, and drug solution; A(blank) is the absorbance of wells containing medium and CCK-8 solution but no cells; and A(control) is the absorbance of wells containing cells and CCK-8 solution but no drug solution.

2.15 Cell migration assay

The Transwell migration assay was conducted to assess cell migration. A 24-well Transwell plate was used, and 100 µL of cell suspension was added to the upper chamber, containing approximately 1×10^5 cells. After a brief incubation to allow cell attachment, the cells were resuspended in serum-free medium, and 600 µL of medium containing 30% FBS was added to the lower chamber to induce migration. The cells were allowed to migrate for 24 h, and after incubation, cells on the lower surface of the membrane were stained with a dye solution for 5 min. Following several washes with water, the Transwell was air-dried, and migrated cells were observed under a microscope.

2.16 Colony-formation assay

Cells were first prepared for culturing by resuspending them in a complete medium and counting the cell number. For each experiment, 400–1000 cells were seeded into each well of a 6-well plate. The plate was incubated for 14 days or until a sufficient number of colonies were formed, and the medium was refreshed every 3 days. At the end of the incubation period, the colonies were fixed with 1 mL of 4% paraformaldehyde for 30–60 min and washed with PBS. After fixation, the colonies were stained with 500 µL of GIEMSA stain solution for 10–20 min and washed again with PBS. Colonies were then counted under a microscope to assess the cell proliferation ability. The experiment utilized standard reagents including DMEM (Corning), fetal bovine serum (FBS, Ausbian), GIEMSA stain (Shanghai Jinbei Biological Technology Co., Ltd.), and paraformaldehyde (Sigma-Aldrich).

3 Results

3.1 Pathogenic targets of LUAD

In the UKB-PPP and Fenland cohorts, a meta-analysis was conducted following the SMR analysis. After applying FDR correction, eight significant targets were identified (as depicted in Table S1). Among these, FUT8 (OR = 1.02, $p = 0.049$), RNASET2 (OR = 1.08, $p = 0.068$), SELL (OR = 1.05, $p = 0.022$), and TGFBI (OR = 1.04, $p = 0.022$) revealed a positive association with LUAD risk. Conversely, ICAM5 (OR = 0.88, $p = 0.002$), INHBB (OR = 0.82, $p = 0.022$), KLK13 (OR = 0.85, $p = 0.021$), and WFIKKN2 (OR = 0.90, $p = 0.049$) were negatively associated with LUAD risk (as depicted in Fig. 1A).

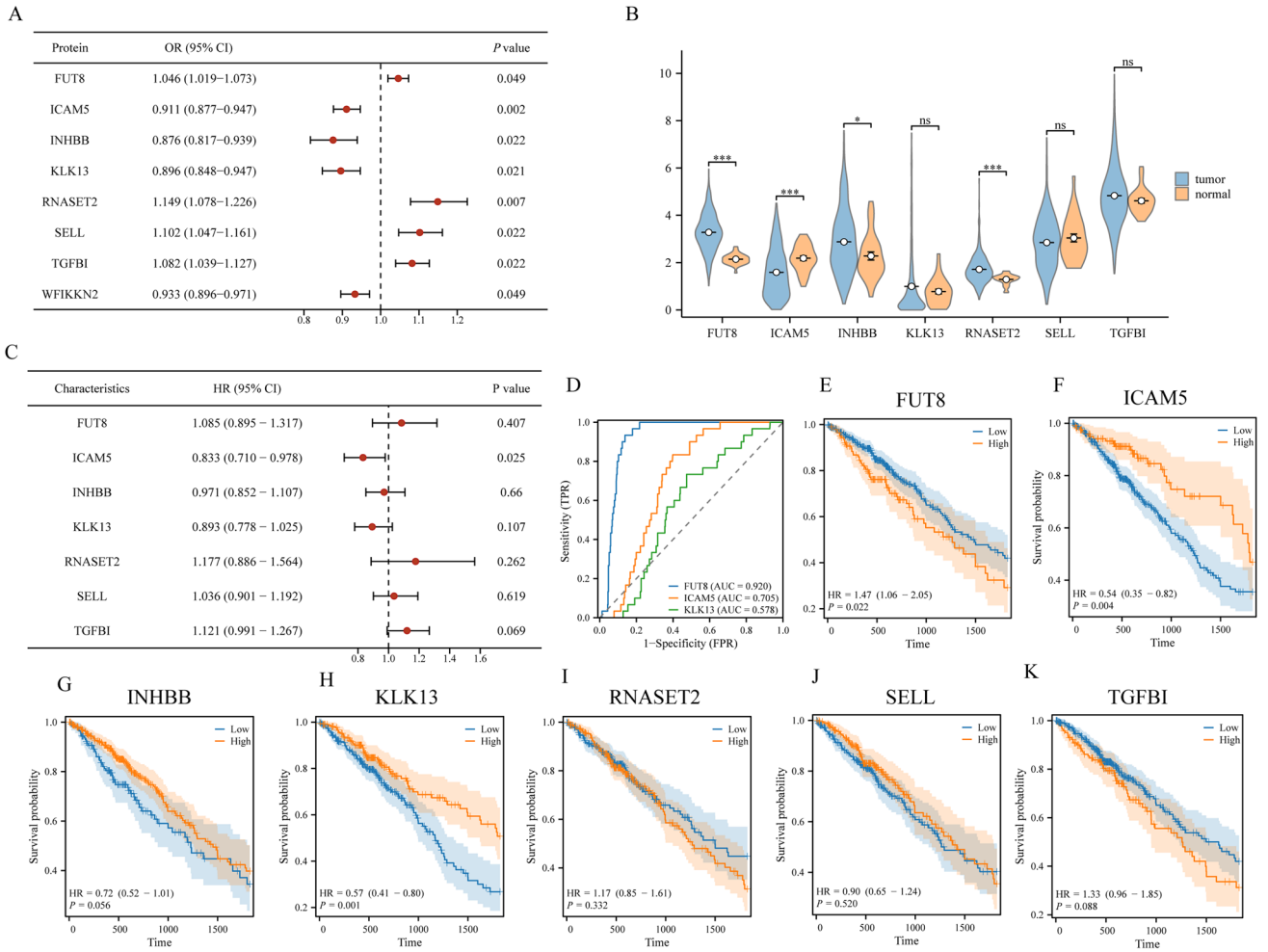


Fig. 1 Expression of Candidate Targets. **A** SMR analysis depicting the pathogenic significance of candidate targets. **B** Differential expression of candidate targets in LUAD tissues compared to normal tissues. **C** Prognostic evaluation of candidate targets for LUAD. **D** Diagnostic performance of candidate targets, including Kaplan–Meier survival curve analysis. **D**–**K** The prognostic impact of significant candidate targets on LUAD. *LUAD* lung adenocarcinoma, *SMR* summary data-based Mendelian randomization, *K-M* Kaplan–Meier

HEIDI testing was conducted for these eight significant targets, with all targets passing the test ($p > 0.05$) (as depicted in Table S1). This indicates that the SMR analysis results are unlikely to be substantially influenced by heterogeneity.

3.2 Expression of candidate targets in LUAD

The mRNA expression levels of the candidate targets were validated in LUAD using the TCGA database. The analysis revealed that FUT8, INHBB, and RNASET2 were significantly upregulated in tumor tissues ($p < 0.05$), whereas ICAM5 was significantly downregulated ($p < 0.05$). Expression data for WFIKKN2 were not available in the TCGA database (Fig. 1B).

3.3 Prognosis of candidate targets for LUAD

The prognostic value of the candidate targets for LUAD was assessed through univariate analysis (as depicted in Table S2). The findings indicated that ICAM5 significantly influenced LUAD prognosis, with a HR of 0.833 (95% CI: 0.710–0.978, $p = 0.025$) (as depicted in Fig. 1C).

Kaplan–Meier (K-M) survival curve analysis demonstrated significant effects of FUT8, ICAM5, and KLK13 on long-term survival in LUAD. Increased expression of FUT8 was associated with poor prognosis ($p = 0.022$), whereas decreased expression of KLK13 ($p = 0.001$) and ICAM5 ($p = 0.004$) correlated with worse prognosis (as depicted in Fig. 1).

To further assess their prognostic use, ROC analysis was performed for ICAM5, FUT8, and KLK13. Among these, FUT8 demonstrated the highest AUC at 0.920 for predicting LUAD occurrence, followed by ICAM5 with an AUC of 0.705. In contrast, KLK13 did not exhibit high diagnostic value (as depicted in Fig. 1D).

3.4 Prognostic models of candidate targets

Given the potential impact of FUT8, ICAM5, and KLK13 on the occurrence and prognosis of LUAD, a predictive model for LUAD prognosis was developed by integrating these three targets with clinical pathological features. A nomogram was constructed to visualize the relationship between these prognostic parameters and 5-year survival rates (as depicted in Fig. 2).

A risk score model was subsequently developed based on the three targets, and calibration plots demonstrated that the nomogram accurately predicted 3-year and 5-year survival outcomes (as depicted in Fig. 2C). RCS analysis further revealed nonlinear relationships between FUT8, ICAM5, KLK13, and patient prognosis (as depicted in Fig. 2D–F).

3.5 Clinical association of ICAM5

Since ICAM5 exhibited significant results in the univariate Cox analysis, its correlation with clinical pathological features was further evaluated (as depicted in Table S2). The analysis revealed that ICAM5 expression was significantly upregulated in stage T1 tumors ($p < 0.05$) (as depicted in Fig. 3). However, no significant associations were observed with N stage, M stage, TNM staging, or patient age (as depicted in Fig. 3B–F).

In multivariate analysis, ICAM5 was identified as an independent prognostic factor associated with patient outcomes (HR: 0.788, 95% CI: 0.663–0.936, $p = 0.007$) (as depicted in Fig. 3G). A nomogram was constructed to predict patient prognosis based on these findings (as depicted in Fig. 3I). The risk score in the Cox model was calculated using the formula:

$$\text{Risk score} = 0.390 + \text{ICAM5} \times (-0.238) + \text{TNM II} \times (0.778) + \text{TNM III} \times (1.422) + \text{TNM IV} \times (1.498).$$

A higher risk score was associated with poorer prognosis (as depicted in Fig. 3H).

The predictive performance of the model was further validated using a calibration plot, time-dependent ROC curves, and the concordance index (c-index = 0.700), all of which demonstrated that the model exhibits good predictive accuracy.

3.6 Functional evaluation of ICAM5

In the Friend analysis, ICAM5 was identified as a potential key gene (as depicted in Fig. 4B). The structural domains of ICAM5 were analyzed, revealing important protein domains (as depicted in Fig. 4C, D). Differential expression analysis using the limma package (as depicted in Table S3), along with GO and KEGG analyses, indicated that ICAM5 is associated with functions such as MHC activity, phagosome formation, and antigen presentation (as depicted in Fig. 4E, Table S4).

GSEA further linked ICAM5 to processes such as the cell cycle and immune response (as depicted in Fig. 4F, Table S5). DO analysis identified associations between ICAM5 and conditions including benign cell tumors, ovarian malignancies, and breast cancer (as depicted in Fig. 4G, Table S6).

In the PPI network, MUC1 occupied a central position, indicating its potential role as a core protein within the network (as depicted in Fig. 4). This highlights the potential interplay between ICAM5 and MUC1 in key biological pathways.

3.7 Relationship between ICAM5 and immune infiltration

Analysis using various immune algorithms revealed that ICAM5 is significantly associated with multiple immune cell types. Notably, ICAM5 demonstrated a negative correlation with CD8 + T cells (as depicted in Fig. 5A–C) and M1 macrophages (as depicted in Fig. 5E). However, no significant correlation was observed between ICAM5 and Immune, Stromal, or ESTIMATE Scores as determined by the ESTIMATE algorithm (as depicted in Fig. 5).

In the analysis of eight immune checkpoint genes, ICAM5 was positively correlated with CD274 (as depicted in Fig. 5F). These findings indicate that ICAM5 is involved in various aspects of immune infiltration and contributes to the development and progression of LUAD.

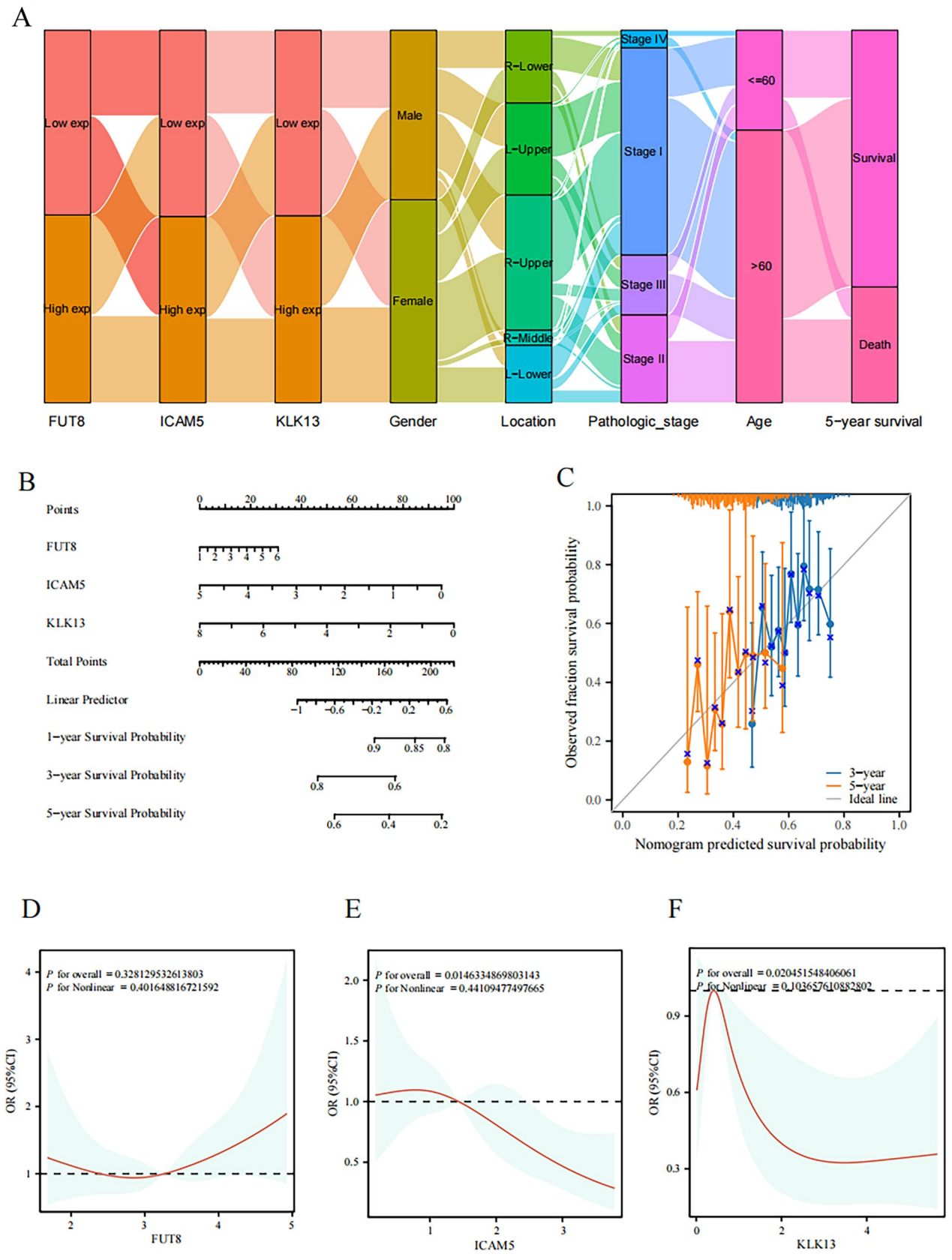


Fig. 2 Predictive models of candidate targets and clinical evaluation. **A** Mulberry plot depicting the relationship between candidate targets and prognosis. **B** Nomogram integrating candidate targets and clinical features. **C** Calibration plot demonstrating the accuracy of the nomogram. **D**–**F** Restricted cubic spline analysis exploring the nonlinear associations of candidate targets with prognosis

Fig. 3 Clinical Association of ICAM5. **A** Differences in T stage expression of ICAM5. **B** Differences in N stage expression of ICAM5. **C** Differences in M stage expression of ICAM5. **D** Differences in TNM stage expression of ICAM5. **E** Variations in tumor location when compared to ICAM5 expression. **F** Relationship between ICAM5 expression and patient age. **G** Multivariate analysis evaluating ICAM5 as an independent prognostic factor. **H** Survival analysis of the predictive model. **I** Nomogram constructed using ICAM5 and clinicopathological features. **J** Calibration plot of the nomogram. **K** Time-dependent ROC curve analysis evaluating the predictive performance of the model

3.8 Single-cell analysis

Single-cell analysis was conducted for all target genes, revealing the availability of expression data for *FUT8*, *ICAM5*, *KLK13*, and *MUC1* (as depicted in Fig. 6). In LUAD pathogenesis, pseudotime analysis was conducted to examine the differentiation process, revealing that KACs differentiate into LUAD-KRAS mutant cells.

Among these genes, the expression levels of *FUT8*, *ICAM5*, and *KLK13* did not exhibit significant changes during the cell differentiation process (as depicted in Fig. 6E). However, *MUC1* displayed dynamic expression changes during the differentiation of KACs into LUAD-KRAS cells, indicating its pivotal role in this differentiation process (as depicted in Fig. 6E). These findings indicate that, while ICAM5 may not directly regulate the differentiation of KACs, it could potentially influence the differentiation trajectory of LUAD-KRAS cells indirectly through its interaction with MUC1 (see Fig. 6E).

3.9 PheWAS

In the PheWAS analysis, no significant associations between ICAM5 and the full spectrum of phenotypes were identified after FDR correction (as depicted in Table S7). However, suggestive associations were observed, indicating that ICAM5 is positively correlated with an increased risk of Inflammatory Bowel Disease (beta = 0.077) and Intestinal Infection (beta = 0.048). Conversely, negative correlations were noted with the risk of Herpes Simplex (beta = – 0.33) and Other Nonspecific Findings on Examination of Urine (beta = – 0.27) (as depicted in Table S7).

Additionally, in drug target searches, ICAM5 was identified as a potential target for several traditional Chinese medicines, including FENG FANG and GUAN HUA DANG SHEN (as depicted in Table 1). These findings highlight potential therapeutic implications of ICAM5 in both modern and traditional medicinal contexts.

3.10 ICAM5 inhibits LUAD invasion in vitro

Given the potential significance of ICAM5 in LUAD, previous studies have also reported that ICAM5 may have an impact on the development of lung cancer (OR = 0.91; 95% CI 0.88–0.95; $P = 2.94 \times 10^{-5}$) [27]. Therefore, we conducted cell functional experiments to further validate the effect of ICAM5 on LUAD. To further determine the effect of ICAM5 on LUAD, we performed in vitro validation (as depicted in Fig. 7). We established an ICAM5 knockdown cell line. In the CCK-8, colony formation, and Transwell assays, we found that the reduction of ICAM5 significantly increased LUAD's colony formation, proliferation, and migration.

4 Discussion

LUAD exhibits significant heterogeneity and is often characterized by an absence of distinct clinical symptoms, complicating early detection. In this study, two major proteomic datasets were used to apply the SMR method for the first time to identify potential pathogenic targets of LUAD, including FUT8, KLK13, and ICAM5. Additionally, multi-omics data were used to comprehensively explore the molecular mechanisms of these targets in LUAD pathogenesis, their prognostic significance, and their associations with the broader phenome.

ICAM5 was identified as a potential primary pathogenic target for LUAD. Notably, a recent study identified SFTPB and KDELC2 as pathogenic targets for LUAD, differing from the findings of this study [28]. This discrepancy may be attributed to differences in the proteomic datasets analyzed and the inclusion of the SMR method. Although similar methods have been applied in eQTL studies for LUAD, no significant association between ICAM5 and LUAD pathogenesis was observed in prior research [13]. A deCODE-based study proposed an association between ICAM5 (OR = 0.95)

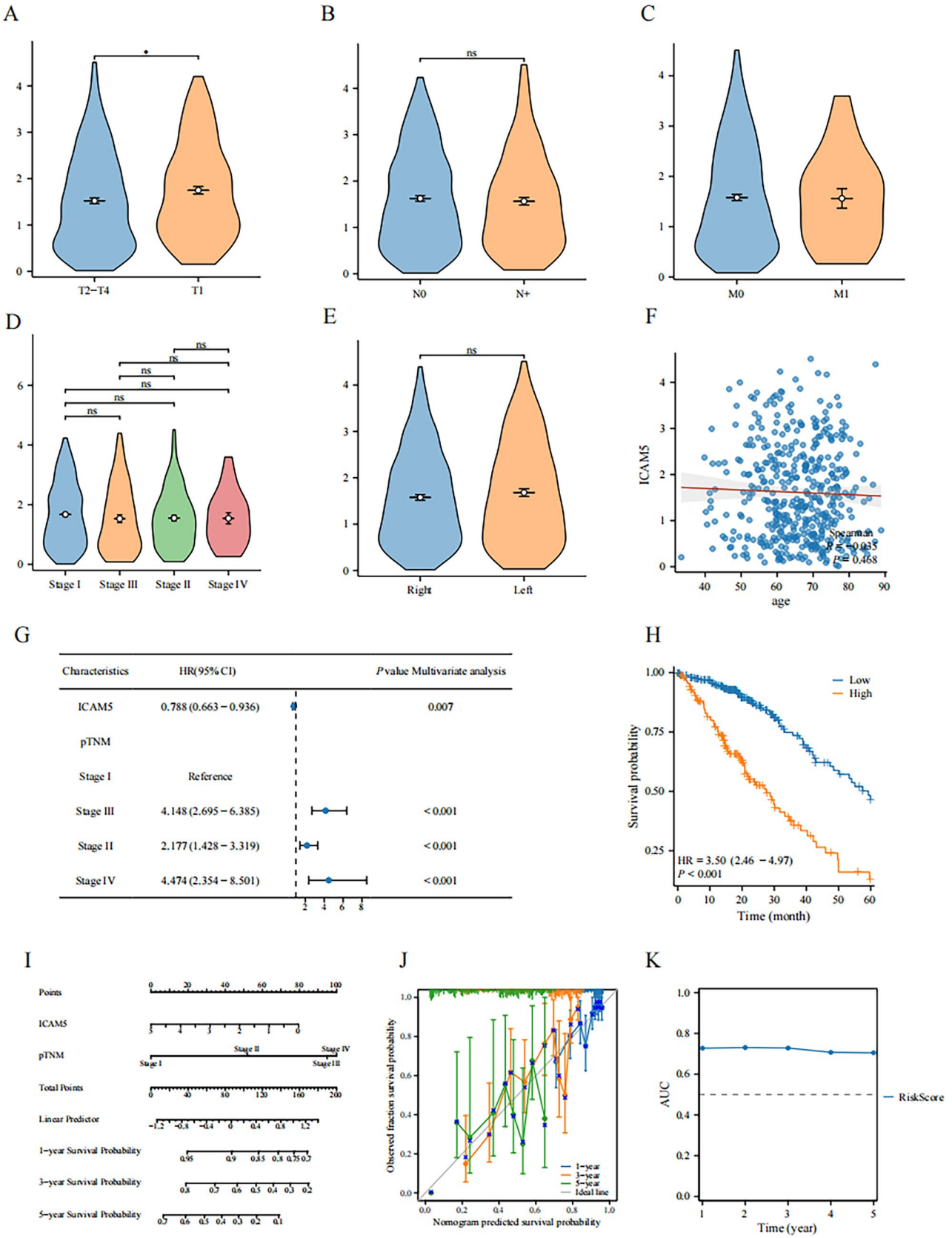

A
B
C
D
E
F
T2−T4
T1
N0
N+
M0
M1
Stage I
Stage III
Stage II
Stage IV
Right
Left
ICAM5
age
Spearman
R = −0.035
P = 0.468
G
Characteristics
HR(95% CI)
P value Multivariate analysis
ICAM5
0.788 (0.663 − 0.936)
0.007
pTNM
Stage I
Reference
Stage III
4.148 (2.695 − 6.385)
< 0.001
Stage II
2.177 (1.428 − 3.319)
< 0.001
Stage IV
4.474 (2.354 − 8.501)
< 0.001
H
Low
High
Survival probability
HR = 3.50 (2.46 − 4.97)
P < 0.001
Time (month)
I
Points
ICAM5
pTNM
Total Points
Linear Predictor
1−year Survival Probability
3−year Survival Probability
5−year Survival Probability
J
Observed fraction survival probability
Nomogram predicted survival probability
1−year
3−year
5−year
Ideal line
K
AUC
RiskScore
Time (year)

Fig. 4 Functional evaluation of ICAM5. **A** PPI network analysis depicting the interactions of ICAM5 with other proteins. **B** Friend analysis ranking candidate targets based on network topology parameters. **C**, **D** Structural domain analysis highlighting the protein domains of ICAM5. **E** GO and KEGG pathway analysis depicting the biological functions and pathways associated with ICAM5. **F** GSEA analysis revealing the involvement of ICAM5 in cell cycle regulation and immune responses. **G** DO analysis linking ICAM5 to specific diseases, including benign and malignant conditions

and lung cancer risk; however, this study relied on a traditional Mendelian randomization design and did not differentiate between pathological subtypes [29]. These findings underscore the heterogeneity of LUAD pathogenesis.

Given the direct role of proteins in phenotypic regulation, proteomic analyses often provide more immediate insights. In this study, the use of meta-analysis across two proteomic datasets enhanced the robustness of the evidence. Furthermore, candidate targets were validated through multi-omics analysis, providing a comprehensive proteomic perspective on the heterogeneous mechanisms underlying LUAD. This work highlights the potential of SMR in advancing LUAD research.

Genetic changes are widely recognized as key contributors to tumor development and progression. In this study, ICAM5 was identified as playing a significant role in the risk of LUAD. This finding aligns with data from the TCGA, where ICAM5 was significantly upregulated in normal tissues, and lower ICAM5 expression was associated with poorer prognosis, consistent with previous research [30]. Among immune targets, a positive correlation was observed between ICAM5 expression and CD274, indicating that patients with LUAD having elevated ICAM5 levels may benefit from immunotherapy.

ICAM5 is predominantly expressed in specific regions of the brain and is implicated in processes such as dendritic growth and microglial cell proliferation [31–33]. However, studies on ICAM5 in cancer remain limited. Current evidence indicates that ICAM5 may reduce the risk of LUAD [29, 34]. The SMR analysis used in this study, based on a cis-design, offers higher precision and efficacy in identifying causal relationships in gene expression regulation. Furthermore, HEIDI test results indicated no significant heterogeneity in the genetic effect of ICAM5 on LUAD, effectively ruling out the possibility of false-positive results caused by linkage disequilibrium.

Although the cis-design focuses on top SNPs within the cis-regulatory range of ICAM5, the broader genetic impact of these top SNPs on LUAD remains unclear [29, 34]. Notably, the rs281439 variant of ICAM5 was identified as having the most significant association with LUAD, indicating that it is the strongest genetic contributor to LUAD risk. Additionally, this genetic polymorphism has been linked to an increased risk of breast cancer, highlighting the heterogeneous regulatory effects of genetic variations on phenotypes [35].

In TCGA, ICAM5 exhibited a high AUC value of 0.705 in predicting LUAD occurrence, providing robust evidence for its potential as a candidate target in LUAD pathogenesis.

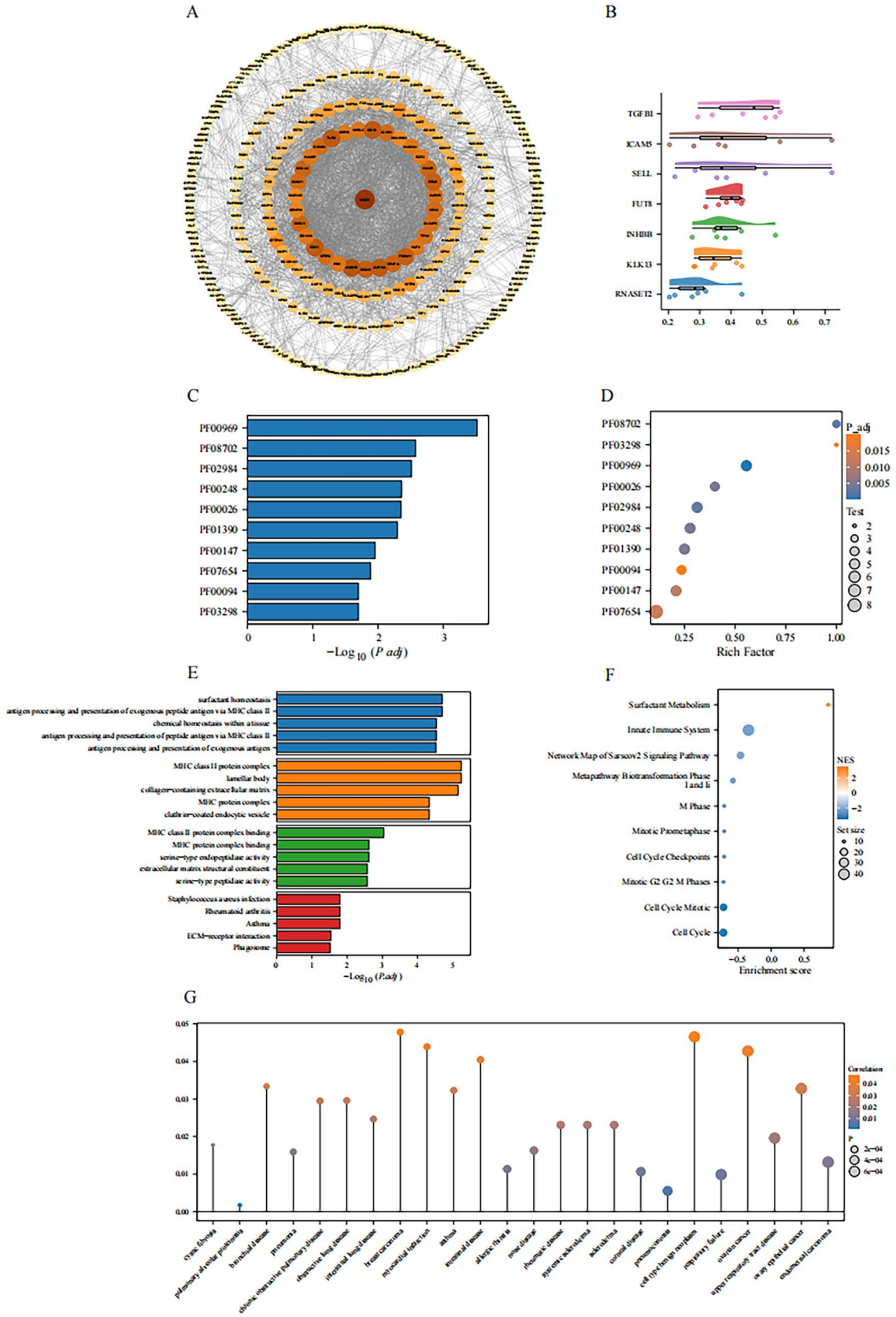
The tumor microenvironment is a pivotal factor in the initiation and progression of LUAD [36]. Analysis revealed that ICAM5 is correlated with various immune cell types, including a positive association with $CD4^+$ T cells and memory B cells. $CD4^+$ T cells play essential roles in antigen presentation and orchestrating immune responses. Elevated infiltration of $CD4^+$ T cells has been linked to improved prognosis in cancer patients [37]. Similarly, memory B cells, which are crucial for immune activation, are key contributors to anti-tumor immunity [38]. Although B cells are traditionally recognized for their role in humoral immunity, emerging evidence indicates their involvement in the favorable response to immune therapies [39]. For instance, Xia et al. reported that B cell infiltration is associated with improved outcomes in NSCLC patients undergoing anti-PD-1 therapy, highlighting their potential positive impact on patient prognosis [40].

Enrichment analysis of ICAM5 further supports its close association with immune system pathways. Additionally, immune target correlation analysis demonstrated a significant positive correlation between ICAM5 and CD274, a critical immune checkpoint molecule. This relationship proposes that ICAM5 may contribute to the benefits observed with immune therapy.

In summary, these findings indicate that ICAM5 plays a role in regulating the immune microenvironment, which may influence the initiation and progression of LUAD.

In single-cell analysis, while ICAM5 expression was discovered in both KACs and malignant cells, ICAM5 does not directly regulate the differentiation process from KACs to malignant cells. KACs, as relatively undifferentiated alveolar progenitor cells, exhibit increased expression of the tumor marker CLDN4, indicating their potential to rapidly differentiate into KRAS-driven malignant cells [25].

Previous research identified MUC1 as a diagnostic marker for small cell lung cancer [41]. A study by Shinjiro et al. demonstrated that MUC1 is strongly associated with tumor differentiation, with significant expression in LUAD [42]. Given that KRAS-mutant solid tumors are linked to MYC signaling, and in KRAS-mutant NSCLC cells, MUC1 has been implicated


A
B
TGFBI
ICAM5
SELL
FUT8
INHBB
KLK13
RNASET2
0.2
0.3
0.4
0.5
0.6
0.7
C
PF00969
PF08702
PF02984
PF00248
PF00026
PF01390
PF00147
PF07654
PF00094
PF03298
0
1
2
3
−Log 10 (P adj)
D
PF08702
PF03298
PF00969
PF00026
PF02984
PF00248
PF01390
PF00094
PF00147
PF07654
P_adj
0.015
0.010
0.005
Test
2
3
4
5
6
7
8
0.25
0.50
0.75
1.00
Rich Factor
E
surfactant homeostasis
antigen processing and presentation of exogenous peptide antigen via MHC class II
chemical homeostasis within a tissue
antigen processing and presentation of peptide antigen via MHC class II
antigen processing and presentation of exogenous antigen
MHC class II protein complex
lamellar body
collagen-containing extracellular matrix
MHC protein complex
clathrin-coated endocytic vesicle
MHC class II protein complex binding
MHC protein complex binding
serine-type endopeptidase activity
extracellular matrix structural constituent
serine-type peptidase activity
Staphylococcus aureus infection
Rheumatoid arthritis
Asthma
ECM-receptor interaction
Phagosome
0
1
2
3
4
5
−Log10 (P.adj)
F
Surfactant Metabolism
Innate Immune System
Network Map of Sarscov2 Signaling Pathway
Metapathway Biotransformation Phase I and II
M Phase
Mitotic Prometaphase
Cell Cycle Checkpoints
Mitotic G2 G2 M Phases
Cell Cycle Mitotic
Cell Cycle
NES
2
0
−2
Set size
10
20
30
40
−0.5
0.0
0.5
Enrichment score
G
0.05
0.04
0.03
0.02
0.01
0.00
Correlation
0.04
0.03
0.02
0.01
P
2e−04
4e−04
6e−04
cystic fibrosis
pulmonary alveolar proteinosis
bronchial disease
pneumonia
chronic obstructive pulmonary disease
obstructive lung disease
interstitial lung disease
bronchocarcinoma
mycobacterial infection
asthma
mucosal disease
allergic disease
nose disease
rheumatic disease
squamous adenoma
adenoma
corneal disease
pneumonitis
cell type benign neoplasm
respiratory failure
ovarian cancer
upper respiratory tract disease
ovary epithelial cancer
endometrial carcinoma

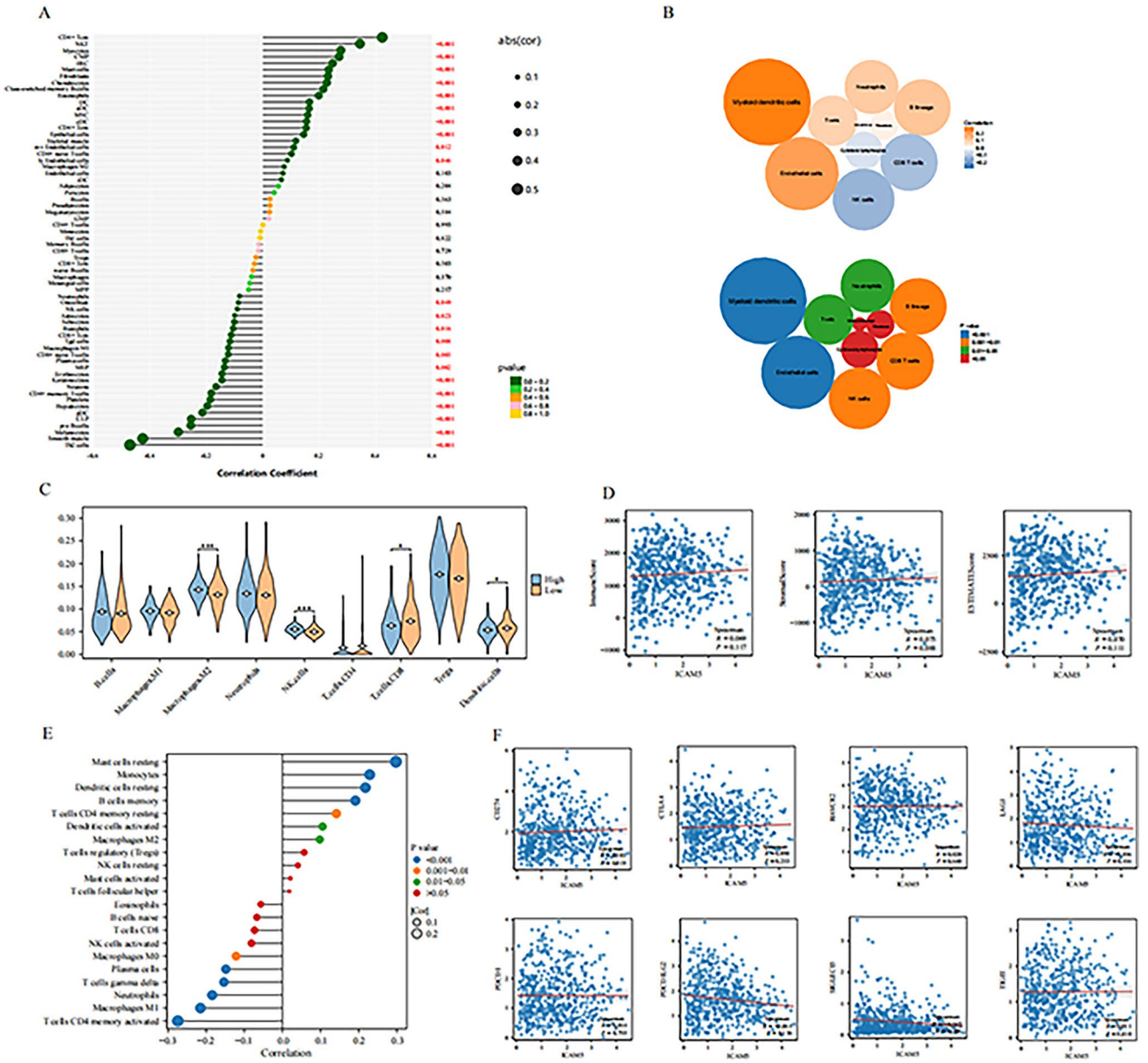


Fig. 5 Relationship between ICAM5 and immune infiltration. **A** xCell analysis depicting the relationship of ICAM5 with various immune cell types. **B** MCP-counter analysis showing immune infiltration patterns associated with ICAM5. **C** QuanTIseq analysis evaluating immune cell compositions correlated with ICAM5. **D** ESTIMATE analysis indicating the relationship between ICAM5 and Immune, Stromal, and ESTIMATE Scores. **E** CIBERSORT analysis detailing the correlation of ICAM5 with specific immune cell subtypes. **F** Correlation analysis revealing the associations of ICAM5 with multiple immune infiltration targets, including CD274

in promoting epithelial-mesenchymal transition and enhancing cellular stemness, MUC1 is as a critical player in tumor progression [43, 44]. Associations between MUC1 and KRAS mutations have been observed in other cancers, including pancreatic and ovarian cancers [45, 46]. Furthermore, recent studies indicate that inhibition of MUC1 can suppress MYC expression, thereby reducing the growth of KRAS-mutant LUAD [47].

These findings strongly propose that MUC1 plays a key role in LUAD pathogenesis. We hypothesize that MUC1 may function as a downstream target of ICAM5, potentially mediating LUAD differentiation. However, further studies are necessary to confirm this hypothesis and elucidate the mechanistic link between ICAM5 and MUC1 in LUAD development.

In addition, previous studies on lung cancer onset targets have found that ICAM5 is associated with a reduced risk of lung cancer. For example, Zhang et al. found that ICAM5 was associated with a reduced risk of lung cancer (OR = 0.91; 95% CI 0.88–0.95; $P = 2.94 \times 10^{-5}$) [27], and Ren et al. also found similar evidence (OR = 0.95; 95% CI: 0.93–0.97; $P = 2.94 \times 10^{-5}$) [34].

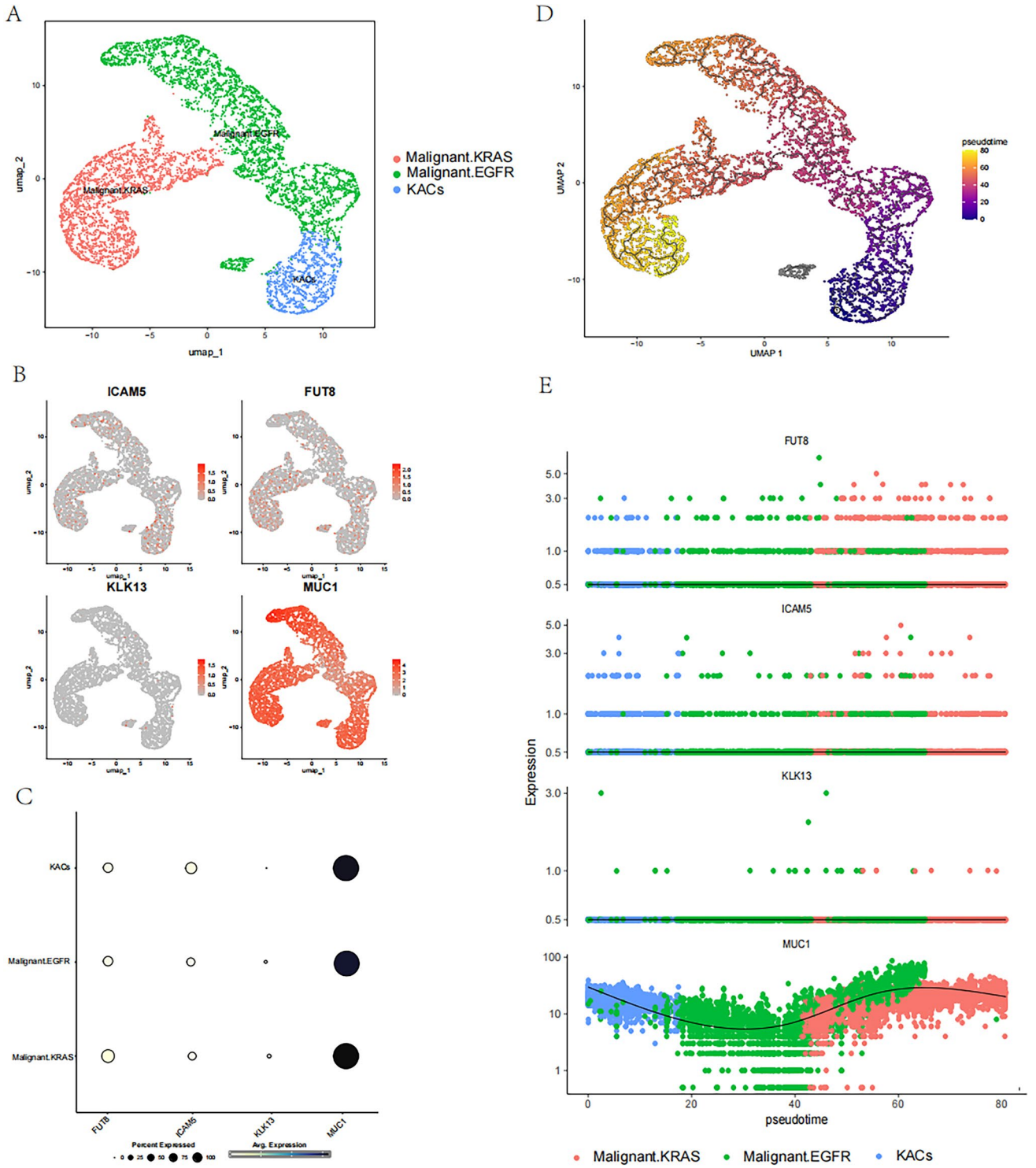


Fig. 6 Single-Cell analysis of candidate targets. **A** Classification of cell types based on single-cell RNA-seq data. **B**, **C** Expression patterns of candidate targets, including FUT8, ICAM5, KLK13, and MUC1, across different cell types. **D**, **E** Pseudotime analysis depicting differentiation trajectories of cells and the dynamic expression of candidate targets, particularly MUC1, during the progression from KACs to LUAD-KRAS mutant cells

Table 1 Traditional Chinese medicine target search for candidate targets

Target	Full name	Ingredient/Herb	ID	Source
ICAM5	intercellular adhesion molecule 5	FENG FANG	HERB001702	Herb
		GUAN HUA DANG SHEN	HERB001965	Herb
		HUI MAO DANG SHEN	HERB002620	Herb
		QIU HUA DANG SHEN	HERB004551	Herb
		IFP; glycerol	TCMBANKIN057946	TCMBank
FUT8	fucosyltransferase 8	XIANG RI KUI ZI	HERB006042	Herb
		3,4-benzopyrene	TCMBANKIN035072	TCMBank
		quercetin	TCMBANKIN058127	TCMBank
KLK13	kallikrein related peptidase 13	BAI GUO	HERB000186	Herb
		BO LING HAO	HERB000515	Herb
		BO NIANG HAO	HERB000523	Herb
		CHA YE	HERB000632	Herb
		CU YE RONG	HERB000992	Herb
		2- naphthalene methanol	TCMBANKIN002996	TCMBank
		plumb	TCMBANKIN003384	TCMBank
		iodo tyrosine	TCMBANKIN003780	TCMBank
		Phenylephrine	TCMBANKIN005364	TCMBank
		L-Phenylalanine	TCMBANKIN005990	TCMBank

Results of the top five drug searches for each candidate target are displayed

This is consistent with our results. However, in these studies, the effect of ICAM5 on lung cancer risk may not have been the most significant, potentially weaker than ALAD (OR = 0.79; 95% CI 0.72–0.87; $P = 4.92 \times 10 - 7$) or SFTPB (OR = 0.93; 95% CI: 0.91–0.95; P = 6.36E − 09). However, based on two protein databases, we did not find significant associations for ALAD and SFTPB. This further emphasizes the consistent regulatory role of ICAM5 in lung cancer risk across multiple protein databases. Additionally, the cell experiments we included further validate our conclusions, demonstrating that ICAM5 has been sufficiently supported by experimental evidence.

Other potential candidate targets for LUAD were identified, including FUT8 and KLK13. FUT8 regulates cancer-associated fibroblasts, thereby promoting the progression of NSCLC [48]. Increased expression of FUT8 has been associated with poor prognosis in LUAD [49]. Similarly, KLK13 has been reported as a marker of poor prognosis in ovarian cancer, although its upregulation has been linked to favorable prognosis in NSCLC [50, 51]. These findings align with the results of this study, providing theoretical support for the identified roles of these targets.

Notably, ICAM5, KLK13, and FUT8 were identified as potential drug targets for various traditional Chinese medicines, highlighting their potential for drug repurposing and new drug development. These findings underscore the importance of exploring therapeutic strategies targeting these molecules to enhance LUAD prognosis. Future research should focus on developing and optimizing drugs targeting these candidates to improve treatment outcomes for patients diagnosed with LUAD.

This study has several notable strengths. First, it used largest-scale proteomics datasets available to conduct SMR analysis, representing a novel approach to identifying pathogenic targets in LUAD. Notably, this is the first study to apply SMR to LUAD. Second, a multi-omics approach was used to comprehensively explore the roles of candidate targets in LUAD pathogenesis. Finally, a PheWAS was conducted to assess potential side effects of the candidate targets, complemented by extensive pharmacological retrieval to provide valuable insights for drug development.

However, the study is not without limitations. First, the research was based exclusively on data from individuals of European descent, limiting the generalizability of the findings to other ethnic groups. Second, the study lacks in-depth experimental validation, and further investigation into the complex mechanisms underlying these targets will be necessary in future research. Finally, we found that ICAM5 has a certain regulatory effect on the tumor microenvironment; however, the current results are based on bioinformatics analysis. Therefore, further experimental validation is needed for these findings.

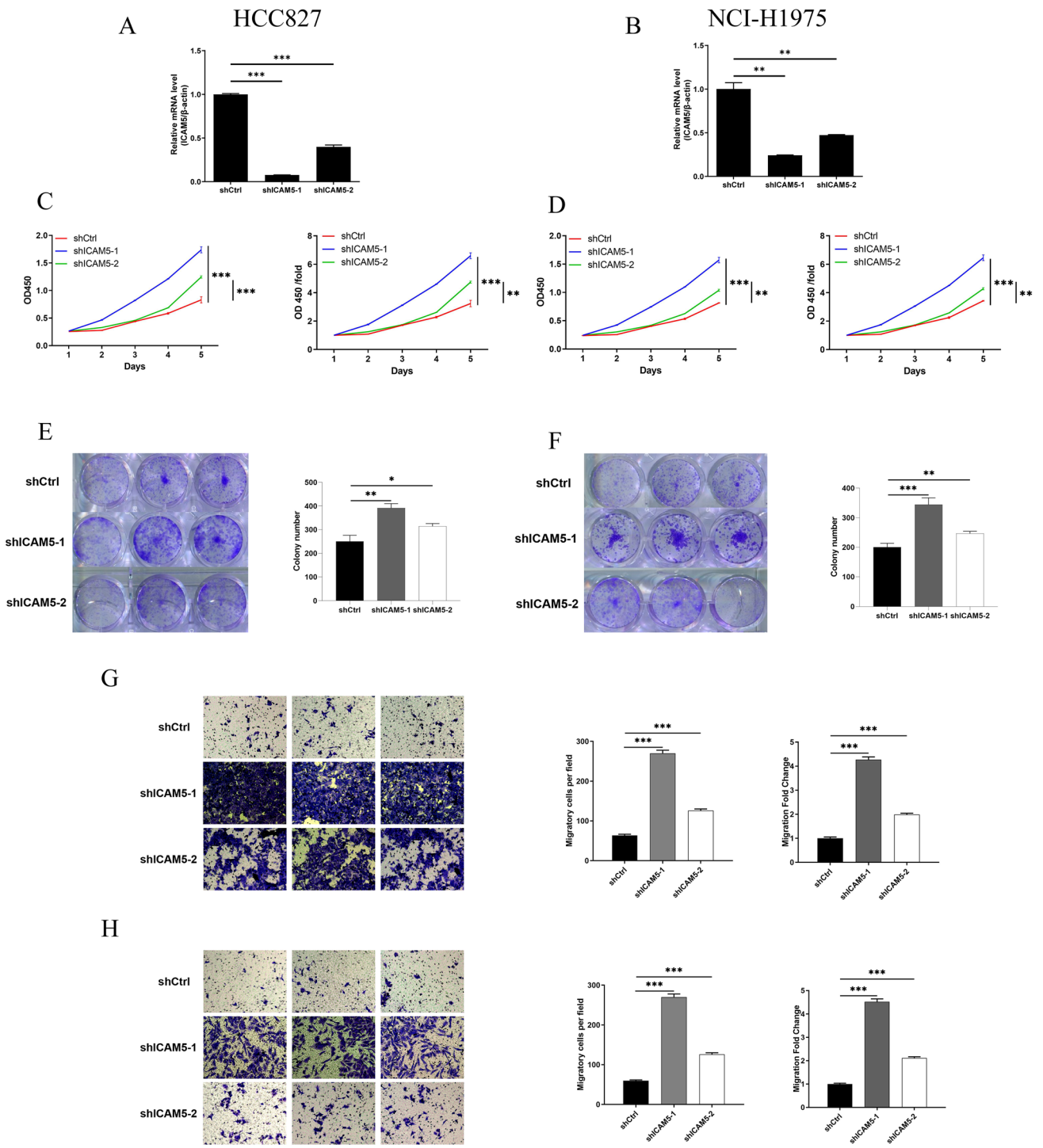


Fig. 7 In vitro validation of the tumor-suppressive effect of ICAM5. **A** Relative mRNA levels of ICAM5 in HCC827 cells with knockdown of ICAM5 (shICAM5-1 and shICAM5-2) compared to the control (shCtrl). **B** Relative mRNA levels of ICAM5 in NCI-H1975 cells with knockdown of ICAM5 (shICAM5-1 and shICAM5-2) compared to the control (shCtrl). **C**, **D** The proliferation of cells with ICAM5 knockdown (shICAM5-1 and shICAM5-2) was assessed by CCK-8 assay over 5 days (left to right: HCC827, NCI-H1975, and two other cell lines). **E**, **F** The colony formation assay was used to assess the impact of ICAM5 knockdown on the colony formation ability of LUAD cells in vitro (left to right: HCC827, NCI-H1975, and two other cell lines). **G**, **H** Transwell assays were employed to evaluate the impact of ICAM5 knockdown on the in vitro migration ability of LUAD cells (**G** HCC827; **H** NCI-H1975)

5 Conclusion

Through multi-omics analysis, this study identified ICAM5 as a primary pathogenic target with a significant impact on both the prognosis and development of LUAD. Furthermore, KLK13 and FUT8 were identified as potential pathogenic targets, providing additional insights into LUAD pathogenesis.

Acknowledgements We would like to acknowledge the hard and dedicated work of all the staff who implemented the intervention and evaluation components of the study.

Author contributions Yue Li conceived and designed this study. Yue Li, Keru Ma, Hao Wang, and Zongying Liu were responsible for the statistical analysis. Zongying Liu was responsible for data collection. Yue Li and Zhuying Li made the final revision of the manuscript. All authors are aware of the final version of the manuscript and agree to publication.

Funding Wu Jieping Foundation (No.320.6750.2021–16-39); Wu Jieping Foundation (No.320.6750.18300).

Data availability All data are derived from previously published datasets. The original data for proteins come from earlier studies, including, UKB-PPP (https://registry.opendata.aws/ukbppp/), and Fenland (https://www.mrc-epid.cam.ac.uk/research/studies/fenland/information-for-researchers/). Data on LUAD were obtained from IEU OpenGWAS (https://gwas.mrcieu.ac.uk/). RNA sequencing data for LUAD were obtained from The Cancer Genome Atlas (https://www.cancer.gov/ccg/research/genome-sequencing/tcga). In the single-cell analysis, we obtained single-cell RNA-seq data from Jiang et al. (GSE189487).

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, Mok T, Petrella F, Spaggiari L, Rosell R. Non-small-cell lung cancer. Nat Rev Dis Primers. 2015;21(1):15009. https://doi.org/10.1038/nrdp.2015.9. (**PMID: 27188576**).
2. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. Nat Rev Clin Oncol. 2023;20(9):624–39. https://doi.org/10.1038/s41571-023-00798-3. (**Epub 2023 Jul 21 PMID: 37479810**).
3. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022;72(1):7–33. https://doi.org/10.3322/caac.21708. (**Epub 2022 Jan 12 PMID: 35020204**).
4. Howlader N, Forjaz G, Mooradian MJ, Meza R, Kong CY, Cronin KA, Mariotto AB, Lowy DR, Feuer EJ. The effect of advances in lung-cancer treatment on population mortality. N Engl J Med. 2020;383(7):640–9. https://doi.org/10.1056/NEJMoa1916623. (**PMID:32786189;PMCID:PMC8577315**).
5. Xu H, Jiang L, Qin L, Shi P, Xu P, Liu C. Single-cell transcriptome analysis reveals intratumoral heterogeneity in lung adenocarcinoma. Environ Toxicol. 2024;39(3):1847–57. https://doi.org/10.1002/tox.24048. (**Epub 2023 Dec 22 PMID: 38133212**).
6. Fu B, Lu L, Huang H. Constructing a prognostic gene signature for lung adenocarcinoma based on weighted gene co-expression network analysis and single-cell analysis. Int J Gen Med. 2022;3(15):5441–54. https://doi.org/10.2147/IJGM.S353848. (**PMID:35685695;PMCID:PMC9173729**).
7. Liu Y, Ye G, Huang L, Zhang C, Sheng Y, Wu B, Han L, Wu C, Dong B, Qi Y. Single-cell transcriptome analysis demonstrates inter-patient and intra-tumor heterogeneity in primary and metastatic lung adenocarcinoma. Aging (Albany NY). 2020;12(21):21559–81. https://doi.org/10.18632/aging.103945. (**Epub 2020 Nov 10. PMID: 33170151; PMCID: PMC7695431**).
8. Chen J, Ruan X, Sun Y, Lu S, Hu S, Yuan S, Li X. Multi-omic insight into the molecular networks of mitochondrial dysfunction in the pathogenesis of inflammatory bowel disease. EBioMedicine. 2024;99:104934. https://doi.org/10.1016/j.ebiom.2023.104934. (**Epub 2023 Dec 16. PMID: 38103512; PMCID: PMC10765009**).

9. Sun J, Zhao J, Jiang F, Wang L, Xiao Q, Han F, Chen J, Yuan S, Wei J, Larsson SC, Zhang H, Dunlop MG, Farrington SM, Ding K, Theodoratou E, Li X. Identification of novel protein biomarkers and drug targets for colorectal cancer by integrating human plasma proteome with genome. Genome Med. 2023;15(1):75. https://doi.org/10.1186/s13073-023-01229-9. (**PMID:37726845;PMCID:PMC10508028**).
10. Zhang N, Li Y, Sundquist J, Sundquist K, Ji J. Identifying actionable druggable targets for breast cancer: Mendelian randomization and population-based analyses. EBioMedicine. 2023;98:104859. https://doi.org/10.1016/j.ebiom.2023.104859. (**Epub 2023 Oct 28. PMID: 38251461; PMCID: PMC10628347**).
11. 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):481–7. https://doi.org/10.1038/ng.3538. (**Epub 2016 Mar 28 PMID: 27019110**).
12. 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. https://doi.org/10.1038/s41467-018-03371-0. (**PMID:29500431;PMCID:PMC5834629**).
13. Song W, Li Y, Yao Y, Sun S, Guan X, Wang B. Systematic druggable genome-wide Mendelian randomization identifies therapeutic targets for lung cancer. BMC Cancer. 2024;24(1):680. https://doi.org/10.1186/s12885-024-12449-6. (**PMID:38834983;PMCID:PMC11151555**).
14. Sun BB, Chiou J, Traylor M, Benner C, Hsu Y, Richardson TG, et al. Genetic regulation of the human plasma proteome in 54,306 UK Biobank participants. 2022.
15. Pietzner M, Wheeler E, Carrasco-Zanini J, Cortes A, Koprulu M, Wörheide MA, Oerton E, Cook J, Stewart ID, Kerrison ND, Luan J, Raffler J, Arnold M, Arlt W, O'Rahilly S, Kastenmüller G, Gamazon ER, Hingorani AD, Scott RA, Wareham NJ, Langenberg C. Mapping the proteo-genomic convergence of human diseases. Science. 2021;374(6569):eabj541. https://doi.org/10.1126/science.abj1541. (**Epub 2021 Nov 12. PMID: 34648354; PMCID: PMC9904207**).
16. Lee DH, Keum N, Hu FB, Orav EJ, Rimm EB, Willett WC, Giovannucci EL. Predicted lean body mass, fat mass, and all cause and cause specific mortality in men: prospective US cohort study. BMJ. 2018;3(362): k2575. https://doi.org/10.1136/bmj.k2575. (**PMID:29970408;PMCID: PMC6028901**).
17. Thormann A, Halachev M, McLaren W, Moore DJ, Svinti V, Campbell A, Kerr SM, Tischkowitz M, Hunt SE, Dunlop MG, Hurles ME, Wright CF, Firth HV, Cunningham F, FitzPatrick DR. Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. Nat Commun. 2019;10(1):2373. https://doi.org/10.1038/s41467-019-10016-3. (**PMID:31147538;PMCID:PMC6542828**).
18. 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):453–7. https://doi.org/10.1038/nmeth.3337. (**Epub 2015 Mar 30. PMID: 25822800; PMCID: PMC4739640**).
19. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, Krogsdam A, Loncova Z, Posch W, Wilflingseder D, Sopper S, Ijsselsteijn M, Brouwer TP, Johnson D, Xu Y, Wang Y, Sanders ME, Estrada MV, Ericsson-Gonzalez P, Charoentong P, Balko J, de Miranda NFDCC, Trajanoski Z. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med. 2019;11(1):34. https://doi.org/10.1186/s13073-019-0638-6. (**Erratum in: Genome Med. 2019 Jul 29;11(1):50. 10.1186/s13073-019-0655-5. PMID: 31126321; PMCID: PMC6534875**).
20. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18(1):220. https://doi.org/10.1186/s13059-017-1349-1. (**PMID:29141660;PMCID:PMC5688663**).
21. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, de Reyniès A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17(1):218. https://doi.org/10.1186/s13059-016-1070-5. (**Erratum.In:GenomeBiol.2016Dec1;17(1):249.doi:10.1186/s13059-016-1113-y.PMID:27765066;PMCID:PMC5073889**).
22. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB, Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612. https://doi.org/10.1038/ncomms3612. (**PMID:24113773;PMCID:PMC3826632**).
23. Wang N, Li Y, Zhou X, Wang X, Yang G. Comprehensive analysis identifies ARHGEF6 as a potential prognostic and immunological biomarker in lung adenocarcinoma. Comput Biol Med. 2023;153: 106448. https://doi.org/10.1016/j.compbiomed.2022.106448. (**Epub 2022 Dec 20 PMID: 36586227**).
24. Zhu J, Fan Y, Xiong Y, Wang W, Chen J, Xia Y, Lei J, Gong L, Sun S, Jiang T. Delineating the dynamic evolution from preneoplasia to invasive lung adenocarcinoma by integrating single-cell RNA sequencing and spatial transcriptomics. Exp Mol Med. 2022;54(11):2060–76. https://doi.org/10.1038/s12276-022-00896-9. (**Epub 2022 Nov 25. PMID: 36434043; PMCID: PMC9722784.v**).
25. Han G, Sinjab A, Rahal Z, Lynch AM, Treekitkarnmongkol W, Liu Y, Serrano AG, Feng J, Liang K, Khan K, Lu W, Hernandez SD, Liu Y, Cao X, Dai E, Pei G, Hu J, Abaya C, Gomez-Bolanos LI, Peng F, Chen M, Parra ER, Cascone T, Sepesi B, Moghaddam SJ, Scheet P, Negrao MV, Heymach JV, Li M, Dubinett SM, Stevenson CS, Spira AE, Fujimoto J, Solis LM, Wistuba II, Chen J, Wang L, Kadara H. An atlas of epithelial cell states and plasticity in lung adenocarcinoma. Nature. 2024;627(8004):656–63. https://doi.org/10.1038/s41586-024-07113-9. (**Epub 2024 Feb 28. Erratum in: Nature. 2024 Apr;628(8006):E1. 10.1038/s41586-024-07277-4. PMID: 38418883; PMCID: PMC10954546**).
26. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, LeFaive J, VandeHaar P, Gagliano SA, Gifford A, Bastarache LA, Wei WQ, Denny JC, Lin M, Hveem K, Kang HM, Abecasis GR, Willer CJ, Lee S. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat Genet. 2018;50(9):1335–41. https://doi.org/10.1038/s41588-018-0184-y. (**Epub 2018 Aug 13. PMID: 30104761; PMCID: PMC6119127**).
27. Zhang L, Xiong Y, Zhang J, Feng Y, Xu A. Systematic proteome-wide Mendelian randomization using the human plasma proteome to identify therapeutic targets for lung adenocarcinoma. J Transl Med. 2024;22(1):330. https://doi.org/10.1186/s12967-024-04919-z. (**PMID:38576019;PMCID:PMC10993587**).
28. Dong B, Wang M, Li K, Li Z, Liu L, Shen S. Plasma proteometabolome in lung cancer: exploring biomarkers through bidirectional Mendelian randomization and colocalization analysis. Hum Mol Genet. 2024;33(19):1688–96. https://doi.org/10.1093/hmg/ddae110. (**PMID: 39011643**).

29. Wang K, Yi H, Wang Y, Jin D, Zhang G, Mao Y. Proteome-wide multicenter Mendelian randomization analysis to identify novel therapeutic targets for lung cancer. Arch Bronconeumol. 2024;60(9):553–8. https://doi.org/10.1016/j.arbres.2024.05.007. (**English, Spanish. Epub 2024 May 16 PMID: 38824092**).
30. Liu B, Zheng H, Ma G, Shen H, Pang Z, Huang G, Song Q, Wang G, Du J. Involvement of *ICAM5* in carcinostasis effects on LUAD based on the ROS1-related prognostic model. J Inflamm Res. 2024;20(17):6583–602. https://doi.org/10.2147/JIR.S475088. (**PMID:39318995;PMCID:PMC11421455**).
31. Yoshihara Y, Oka S, Nemoto Y, Watanabe Y, Nagata S, Kagamiyama H, Mori K. An ICAM-related neuronal glycoprotein, telencephalin, with brain segment-specific expression. Neuron. 1994;12(3):541–53. https://doi.org/10.1016/0896-6273(94)90211-9. (**PMID: 7794412**).
32. Tian L, Nyman H, Kilgannon P, Yoshihara Y, Mori K, Andersson LC, Kaukinen S, Rauvala H, Gallatin WM, Gahmberg CG. Intercellular adhesion molecule-5 induces dendritic outgrowth by homophilic adhesion. J Cell Biol. 2000;150(1):243–52. https://doi.org/10.1083/jcb.150.1.243. (**PMID:10893271;PMCID:PMC2185561**).
33. Mizuno T, Yoshihara Y, Kagamiyama H, Ohsawa K, Imai Y, Kohsaka S, Mori K. Neuronal adhesion molecule telencephalin induces rapid cell spreading of microglia. Brain Res. 1999;849(1–2):58–66. https://doi.org/10.1016/s0006-8993(99)01984-8. (**PMID: 10592287**).
34. Ren F, Jin Q, Liu T, Ren X, Zhan Y. Proteome-wide Mendelian randomization study implicates therapeutic targets in common cancers. J Transl Med. 2023;21(1):646. https://doi.org/10.1186/s12967-023-04525-5. (**PMID:37735436;PMCID:PMC10512580**).
35. Liu L, Sun M, Song D, Wang Z. The genetic polymorphisms of intercellular cell adhesion molecules and breast cancer susceptibility: a meta-analysis. Mol Biol Rep. 2013;40(2):1855–60. https://doi.org/10.1007/s11033-012-2241-4. (**Epub 2012 Oct 19 PMID: 23079714**).
36. Yang L, He YT, Dong S, Wei XW, Chen ZH, Zhang B, Chen WD, Yang XR, Wang F, Shang XM, Zhong WZ, Wu YL, Zhou Q. Single-cell transcriptome analysis revealed a suppressive tumor immune microenvironment in EGFR mutant lung adenocarcinoma. J Immunother Cancer. 2022;10(2): e003534. https://doi.org/10.1136/jitc-2021-003534. (**PMID:35140113;PMCID:PMC8830346**).
37. Dosset M, Castro A, Carter H, Zanetti M. Telomerase and CD4 T cell immunity in cancer. Cancers (Basel). 2020;12(6):1687. https://doi.org/10.3390/cancers12061687. (**PMID:32630460;PMCID:PMC7352225**).
38. Azizi A, Mehdipour F, Samadi M, Rasolmali R, Talei AR, Ghaderi A. Atypical memory B cells increase in the peripheral blood of patients with breast cancer regardless of lymph node involvement. BMC Immunol. 2024;25(1):25. https://doi.org/10.1186/s12865-024-00620-4. (**PMID:38702630;PMCID:PMC11067195**).
39. Petitprez F, de Reyniès A, Keung EZ, Chen TW, Sun CM, Calderaro J, Jeng YM, Hsiao LP, Lacroix L, Bougoüin A, Moreira M, Lacroix G, Natario I, Adam J, Lucchesi C, Laizet YH, Toulmonde M, Burgess MA, Bolejack V, Reinke D, Wani KM, Wang WL, Lazar AJ, Roland CL, Wargo JA, Italiano A, Sautès-Fridman C, Tawbi HA, Fridman WH. B cells are associated with survival and immunotherapy response in sarcoma. Nature. 2020;577(7791):556–60. https://doi.org/10.1038/s41586-019-1906-8. (**Epub 2020 Jan 15 PMID: 31942077**).
40. Xia L, Guo L, Kang J, Yang Y, Yao Y, Xia W, Sun R, Zhang S, Li W, Gao Y, Chen H, Li Z, Yang J, Lu S, Wang Y. Predictable roles of peripheral IgM memory B cells for the responses to Anti-PD-1 monotherapy against advanced non-small cell lung cancer. Front Immunol. 2021;24(12): 759217. https://doi.org/10.3389/fimmu.2021.759217. (**PMID:34899709;PMCID:PMC8652218**).
41. Savarese-Brenner B, Heugl M, Rath B, Schweizer C, Obermayr E, Stickler S, Hamilton G. MUC1 and CD147 are promising markers for the detection of circulating tumor cells in small cell lung cancer. Anticancer Res. 2022;42(1):429–39. https://doi.org/10.21873/anticanres.15501. (**PMID: 34969753**).
42. Nagai S, Takenaka K, Sonobe M, Ogawa E, Wada H, Tanaka F. A novel classification of MUC1 expression is correlated with tumor differentiation and postoperative prognosis in non-small cell lung cancer. J Thorac Oncol. 2006;1(1):46–51 (**PMID: 17409826**).
43. Yeh E, Cunningham M, Arnold H, Chasse D, Monteith T, Ivaldi G, Hahn WC, Stukenberg PT, Shenolikar S, Uchida T, Counter CM, Nevins JR, Means AR, Sears R. A signalling pathway controlling c-Myc degradation that impacts oncogenic transformation of human cells. Nat Cell Biol. 2004;6(4):308–18. https://doi.org/10.1038/ncb1110. (**Epub 2004 Mar 14 PMID: 15048125**).
44. Kharbanda A, Rajabi H, Jin C, Alam M, Wong KK, Kufe D. MUC1-C confers EMT and KRAS independence in mutant KRAS lung cancer cells. Oncotarget. 2014;5(19):8893–905. https://doi.org/10.18632/oncotarget.2360. (**PMID:25245423;PMCID:PMC4253405**).
45. Ferchichi M, Jouini R, Ayari I, Koubaa W, Chadli-Debbiche A, BenBrahim E. KRAS, NRAS and BRAF analysis of ampullary adenocarcinoma classified using CK7, CK20, MUC1 and MUC2. J Gastrointest Oncol. 2018;9(5):820–7. https://doi.org/10.21037/jgo.2018.05.03. (**PMID:30505580;PMCID:PMC6219964**).
46. Tirodkar TS, Budiu RA, Elishaev E, Zhang L, Mony JT, Brozick J, Edwards RP, Vlad AM. MUC1 positive, Kras and Pten driven mouse gynecologic tumors replicate human tumors and vary in survival and nuclear grade based on anatomical location. PLoS ONE. 2014;9(7): e102409. https://doi.org/10.1371/journal.pone.0102409. (**PMID:25078979;PMCID:PMC4117479**).
47. Bouillez A, Rajabi H, Pitroda S, Jin C, Alam M, Kharbanda A, Tagde A, Wong KK, Kufe D. Inhibition of MUC1-C suppresses MYC expression and attenuates malignant growth in KRAS mutant lung adenocarcinomas. Cancer Res. 2016;76(6):1538–48. https://doi.org/10.1158/0008-5472.CAN-15-1804. (**Epub 2016 Feb 1. PMID: 26833129; PMCID: PMC4794417**).
48. Li F, Zhao S, Cui Y, Guo T, Qiang J, Xie Q, Yu W, Guo W, Deng W, Gu C, Wu T. α1,6-Fucosyltransferase (FUT8) regulates the cancer-promoting capacity of cancer-associated fibroblasts (CAFs) by modifying EGFR core fucosylation (CF) in non-small cell lung cancer (NSCLC). Am J Cancer Res. 2020;10(3):816–37 (**PMID: 32266093; PMCID: PMC7136908**).
49. Liang Y, Wang T, Gao R, Jia X, Ji T, Shi P, Xue J, Yang A, Chen M, Han P. Fucosyltransferase 8 is overexpressed and influences clinical outcomes in lung adenocarcinoma patients. Pathol Oncol Res. 2022;14(28):1610116. https://doi.org/10.3389/pore.2022.1610116. (**PMID:35237113;PMCID:PMC8883820**).
50. White NM, Mathews M, Yousef GM, Prizada A, Popadiuk C, Doré JJ. KLK6 and KLK13 predict tumor recurrence in epithelial ovarian carcinoma. Br J Cancer. 2009;101(7):1107–13. https://doi.org/10.1038/sj.bjc.6605280. (**PMID:19707197;PMCID:PMC2768090**).
51. Gueugnon F, Barascu A, Mavridis K, Petit-Courty A, Marchand-Adam S, Gissot V, Scorilas A, Guyetant S, Courty Y. Kallikrein-related peptidase 13: an independent indicator of favorable prognosis for patients with nonsmall cell lung cancer. Tumour Biol. 2015;36(7):4979–86. https://doi.org/10.1007/s13277-015-3148-1. (**Epub 2015 Feb 13 PMID: 25677900**).