Fan *et al. BMC Cancer* (2025) 25:688
https://doi.org/10.1186/s12885-025-14044-9

BMC Cancer

RESEARCH Open Access

CMSS1: A RNA binding protein with pivotal roles in non-small cell lung cancer progression and prognosis

Zhe Fan[1], Wanyu Liu[1], Zhiwei Gao[1], Youfa Liu[1], Hongyang Hai[2] and Zhenyang Lv[1*]

Abstract

Background The Cms1 ribosomal small subunit homolog (*CMSS1*), an RNA-binding protein (RBP), plays a crucial role in tumor development. However, the prognostic and immunological role of *CMSS1* in non-small cell lung cancer (NSCLC) remains unclear.

Methods Differentially expressed RBP genes were identified using The Cancer Genome Atlas (TCGA) database, and the hub RBP-related gene, *CMSS1*, was selected through univariate Cox regression analysis and Kaplan-Meier tests. To evaluate the prognostic capacity of the *CMSS1*, time-dependent receiver operating characteristic curves, Kaplan-Meier curves and multivariate Cox regression analyses were conducted. The relationship between the *CMSS1* gene and tumor-infiltrating immune cells was assessed using the ImmuCellAI algorithm. Additionally, a loss-of-function assay was performed to investigate the functional role of *CMSS1* in NSCLC cells.

Results Bioinformatic analysis revealed that *CMSS1*, an RBP-related gene, was notably upregulated in NSCLC tumors, with elevated RNA levels correlating with poor prognosis in NSCLC patients. Immune cell infiltration analysis showed that *CMSS1* expression was negatively correlated with CD4 T cells and was positively correlated with macrophages and Tregs. Furthermore, RT-qPCR and western blot confirmed the increased *CMSS1* mRNA and CMSS1 protein levels in NSCLC cell lines. Significantly, downregulation of *CMSS1* inhibited NSCLC cell viability, migration and invasion.

Conclusion Our findings suggest that *CMSS1* may serve as both a prognostic indicator and a therapeutic target for patients with NSCLC. This study may provide potential guidance for precision therapy and accurate prognosis prediction for patients with NSCLC.

Keywords Non-small cell lung cancer, RNA binding proteins, *CMSS1*, Prognosis, Immune cell infiltration

*Correspondence:
Zhenyang Lv
lv17709872529@163.com
[1]Department of Thoracic Surgery 2, the Second Hospital of Dalian Medical University, No. 467 Zhongshan Road, Shahekou District, Dalian 116000, China
[2]Department of General Surgery, People's Hospital of Huangyuan County in Qinghai Province, Xining City 812100, Qinghai Province, China

Introduction

Lung cancer (LC) poses a significant threat to human health, ranking among the most prevalent malignancies, with high morbidity, mortality, and poor prognosis [1]. Within its histological spectrum, non-small cell lung cancer (NSCLC) is the predominant form, constituting approximately 80–85% of all LC cases [2]. NSCLC encompasses two primary subtypes including lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), both of which significantly contribute to its incidence [3]. NSCLC is characterized by rapid metastatic potential, robust infiltrative capacity, and a dismal 5-year survival rate [4, 5]. Alarmingly, up to 80% of NSCLC patients are diagnosed with advanced disease, resulting in poor clinical outcomes [6] and the development of drug resistance [7]. This situation underscores the critical need for identifying and developing potent, highly sensitive biomarkers. The identification of such biomarkers is essential for enhancing therapeutic strategies and ultimately improving patient outcomes in NSCLC.

RNA binding proteins (RBPs), a class of proteins equipped with RNA binding domains, serve as pivotal regulators of cancer phenotypes, fundamentally impacting disease progression [8]. These RBPs are intricately involved in RNA splicing, post-transcriptional modification, and translation [9]. To date, an extensive repertoire of 1,542 human RBP genes has been catalogued, underscoring their widespread importance [10]. A large of evidence have underscored the pivotal roles played by RBPs across various human malignancies. For instance, Qiu et al. developed a novel RBP signature to predict clinical outcomes and guide clinical therapy in gastric cancer [10]. Similarly, Huang et al. constructed a prognostic risk score model centered on six RBP genes for hepatocellular carcinoma (HCC), validating its prognostic capabilities and underscoring its potential as a robust biomarker [11]. Furthermore, the RBP gene *YTHDF2* has been shown to regulate *OCT4* expression via m6A RNA methylation, thereby promoting HCC stem cell phenotypes and enhancing metastatic potential [12]. While the functional and prognostic significance of nine RBPs in LUSC has garnered attention [13], the landscape of RBP functions in NSCLC remains largely unexplored. This gap in knowledge underscores the urgent need for in-depth investigations into the role of RBPs in NSCLC, with the potential to identify novel therapeutic targets and prognostic markers for this aggressive malignancy.

Table 1 The details of mRNA expression profile datasets

Datasets	Cancer type	Number of tumor samples	Number of control samples
TCGA	LUAD and LUSC	617	71
GSE101929	NSCLC	66	/
GSE30219	LC	293	14
GSE31210	LUAD	226	/
GSE19188	NSCLC	91	65

In this study, our objective is to investigate the RBPs that are closely linked to the prognosis of NSCLC. Cms1 ribosomal small subunit homolog (*CMSS1*), categorized as an RBP, has been demonstrated to be linked to cancer prognosis and survival [14, 15]. Our results found that the levels of *CMSS1* were elevated in NSCLC, and correlated with unfavorable prognosis in NSCLC patients. Moreover, we conducted a functional enrichment analysis regarding *CMSS1*, and examined its relationship with the immune microenvironment in NSCLC. Consequently, the findings from this research may offer potential molecular markers for diagnosing and prognosing NSCLC.

Materials and methods

Subjects

The mRNA expression profile datasets for NSCLC, including lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), were downloaded from The Cancer Genome Atlas (TCGA, https://tcga-data.nci.nih.gov/tcga/) database, referred to as the TCGA-NSCLC cohort. Meanwhile, four lung cancer-related datasets, GSE101929, GSE30219, GSE31210 and GSE19188, were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The details of all datasets were shown in Table 1.

In addition, 1,542 RBP were retrieved from a previous published study (Table S1) [16].

Differentially expressed gene analysis

The differentially expressed genes (DEGs) were screened using R package "DESeq2" (version 1.44) [17] with the cut-off criteria of $|\log_2 FC| > 1$ and FDR < 0.05.

Functional enrichment analysis

The Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, and Gene Set Enrichment Analysis (GSEA) were conducted using the "clusterProfiler" R package (version 4.8.3) [18] to explore the function of genes.

Cox regression analysis

Univariate Cox regression analysis was performed for screening candidate genes associated with NSCLC prognosis. Multivariate Cox regression analysis was employed to explore the independence of *CMSS1* in predicting the survival of NSCLC patients.

Survival analysis

Based on the median level of *CMSS1*, patients from various datasets were categorized into low-expression

(*CMSS1*-L) group and high-expression (*CMSS1*-H) group. The R packages "survival" (version 3.7) and "survminer" (version 0.4.9) were used to estimate the overall survival of patients between two groups. The survival curves of two groups was drawn using Kaplan-Meier method, and the significance of differences was evaluated with the log-rank test. The receiver operating characteristic (ROC) curves were generated using the "pROC" (version 1.18.5) [19] and "timeROC" (version 0.4) R packages [20] to assess the diagnostic and prognostic capabilities of *CMSS1*, respectively.

Immune cell infiltration analysis

Using the "ESTIMATE" function package (version 1.0.13) [21], the pre-screened stroma and immune-related gene sets were used to predict stromal or immune cell infiltration levels in tumor tissue, thereby calculating the stromal score and immune score.

Drug sensitivity analysis

Drug sensitivity prediction was conducted using the R package "oncoPredict" (version 1.2) [22]. The CTRP2_Expr expression matrix (including 51,847 genes and 829 cell lines) and CTRP2_Res drug sensitivity data (including the half maximal inhibitory concentration (IC50) values for 545 drugs across the same 829 cell lines) were employed as the training set. Subsequently, the gene expression profile of the sample served as the input, and the "calcPhenotype" function was applied to predict the IC50 values for each drug in the sample.

Cell culture and transfection

The BEAS-2B cell line (BNCC359274, BeNa Culture Collection) was cultured in DMEM (high-glucose) medium containing 10% FBS and 1% P/S, and maintained at 37℃ in a humidified atmosphere containing 5% CO_2. Three NSCLC cell lines, NCI-H1975 (No. BNCC340345, BeNa Culture Collection), NCI-H1703 (No. BNCC101663, BeNa Culture Collection) and NCI-H1650 (No. CL-0166, Procell), were cultured in RPMI-1640 medium containing 10% FBS and 1% P/S, and maintained at 37℃ in a 5% CO_2 humidified atmosphere.

NCI-H1975 and NCI-H1703 cells were transfected with siRNA negative control, *CMSS1*-siRNA-1 (Sense: 5'-UAAGUUUUACCCACUUAGGAC-3'; Antisense: 5'-CCUAAGUGGGUAAAACUUAGG-3'), *CMSS1*-siRNA-2 (Sense: 5'-UUCUUCUAAUUCAAUCACCAA-3'; Antisense: 5'-GGUGAUUGAAUUAGAAGAACU-3') and *CMSS1*-siRNA-3 (Sense: 5'-UACAACAACUAAGUAAGUGCA-3'; Antisense: 5'-CACUUACUUAGUUGUUGUAUC-3') using the Transfection Kit (No. AQ11668, Beijing Aoqing Biotechnology Co., Ltd). Meanwhile, NCI-H1975 cells were transfected with *CMSS1*-pcDNA3.1 overexpression (OE-*CMSS1*) and negative control pcDNA3.1 (OE-NC) plasmids using the Transfection Kit (No. AQ11668, Beijing Aoqing Biotechnology Co., Ltd).

Reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR)

The TriQuick Reagent Total RNA Extraction reagent (No. R1100, Solarbio) was used to isolate total RNA from cells. Next, RNA was reverse transcribed into cDNA using a reverse transcription Kit (No. AG11728, Accurate Biology). Real-time PCR was performed with the SYBR Green Premix *Pro Taq* HS qPCR Kit (AG11701, Accurate Biology). GAPDH was used as an internal control and gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method.

CMSS1: 5'-TGGAGCTCATTAGGTCGATG-3' (forward) and 5'-GCTTCTCCAGCAACTTTACC-3' (reverse); *GAPDH*: 5'-GAAGGTGAAGGTCGGAGTC-3' (forward) and 5'-GAAGATGGTGATGGGATTTC-3' (reverse).

Western blot

Equal amounts of proteins were separated by 10% SDS-PAGE gels and subsequently electrotransferred to polyvinylidene difluoride membranes. The membranes were blocked with blocking solution (5% skimmed milk in TBST) for 1 h, and then blotted with primary antibodies at 4 °C overnight, including anti-CMSS1 (No. NBP1-81078, Novus Biologicals) and anti-GAPDH (No. 60004-1-Ig, Proteintech) primary antibodies. Following incubation with corresponding secondary antibodies, protein bands were detected by electrochemiluminescence Plus ultra-sensitive liquid (No. P0018M, Beyotime). Goat anti-mouse secondary antibody (No. ZB-2305, ZSGB-BIO) was used for GAPDH primary antibody, and goat anti-rabbit secondary antibody (No. ZB-2301, ZSGB-BIO) was used for CMSS1 primary antibody.

Cell counting kit-8 (CCK-8) assay

NCI-H1975 and NCI-H1703 cells (5×10^3 cells/well) were seeded into 96-well plates. After transfection, a CCK-8 assay kit (No. C0037, Beyotime) was used for assessing cell viability. Subsequently, the absorbance values at 450 nm were measured using a microplate reader (Multiskan 51119000, ThermoFisher).

Transwell assay

NCI-H1975 cells suspended in 100 μl of RPMI-1640 medium containing 0.1% FBS were loaded into the upper compartment chambers of 24-well plates equipped with cell culture inserts. Meanwhile, the lower chamber was filled with 300 μl of RPMI-1640 medium containing 10% FBS. After incubation of 24 h, the migrating and invasive cells on the under surface were stained with 0.1% crystal violet. For cell invasion assay, transwell inserts were pre-coated with Matrigel (No. 356234, Corning). Finally,

three areas were randomly selected for imaging under a microscope (IMT-2, OLYMPUS).

Wound healing assay

NCI-H1975 cells (3×10^5 cells/well) were plated into a 6 well plate, and incubated overnight at 37 °C. When the cells uniformly covered the bottom of the 6 well plate, a 200 μl yellow pipette tip was employed to create a straight line in the cell culture plate. After incubation of 0, 24 and 48 h, images were captured under a microscope (CKX31, OLYMPUS). The gap distance was assessed using ImageJ. The rate of wound healing was calculated as follows = (0 h scratch area − 24/48 h scratch area)/0 h scratch area × 100%.

Statistical analysis

The Wilcox test was applied to assess whether gene expression and infiltration of immune cells were statistically significant. The Pearson correlation was performed using R function "cor". All statistical analyses were conducted using R software (version 4.3.3). For RT-qPCR, western blot and functional analyses, one-way ANOVA or two-way ANOVA was used to analyze the difference between multiple groups. Statistical significance was considered at $p < 0.05$.

Results

Screening for differentially expressed RBP genes (DE-RBPs) in NSCLC

To identify DE-RBPs in NSCLC, we performed differential expression analysis between control and tumor groups using the data from the TCGA-NSCLC cohort. The results showed that compared to the control samples, there were 11,124 DEGs in the NSCLC samples, including 7,969 up-regulated genes and 3,155 down-regulated genes (Fig. 1A, B; Table S1). Subsequently, we identified 196 overlapping genes (DE-RBPs) by intersecting the 11,124 DEGs with 1,542 published RBP genes (Fig. 1C; Table S1).

Subsequently, GO and KEGG enrichment analyses were performed to determine the biological functions of 196 DE-RBPs. The GO enrichment results revealed that a total of 317 GO terms were significantly enriched, comprising 170 GO_biological processes (BPs), 77 GO_molecular functions (MFs), and 70 GO_cellular components (CCs) terms (Table S2), and the top five GO terms were shown in Fig. 1D. The KEGG enrichment results showed that these DE-RBPs were significantly involved in 10 pathways such as mRNA surveillance pathway, RNA degradation, and spliceosome (Fig. 1E; Table S2).

Identification of *CMSS1* as a prognostic biomarker in NSCLC

To identify RBP genes that exert a substantial influence on the prognosis of NSCLC patients, we conducted univariate cox regression analysis on the 196 DE-RBPs, applying a significance threshold of $P < 0.05$. This analysis successfully identified 24 DE-RBPs that demonstrated significant prognostic value (Fig. 2A and Table S3).

To further elucidate the individual contributions of these 24 DE-RBPs to the overall survival (OS) of NSCLC patients, we conducted a survival analysis using the TCGA-NSCLC dataset. The results revealed a significant difference in OS between low and high levels of expression of these four genes (*GAPDH*, *POP1*, *CMSS1*, *TARBP1*) in NSCLC patients (Fig. 2B-E). Furthermore, we verified the absence of violations of the proportional hazards (PH) assumption by analyzing the Schoenfeld residual plots for these four genes (Fig. S1A-S1D), and the results of the PH tests for these genes all produced p-values greater than 0.05 (Table S3), indicating compliance with the PH assumption. Notably, during a comprehensive review of the existing literature, we uncovered that *CMSS1* had not been previously reported in relation to NSCLC prognosis, prompting us to prioritize *CMSS1* as a novel biomarker for subsequent investigative endeavors.

Significant over-expression of *CMSS1* in NSCLC

To elucidate the specific role of *CMSS1* in NSCLC, we conducted an analysis of *CMSS1* expression patterns in tumor and control samples using the TCGA-NSCLC, GSE30219, and GSE19188 datasets. The results revealed a statistically significant up-regulation of *CMSS1* expression in all three datasets within tumor samples compared to controls (Fig. 3A-C). To further validate and consolidate these findings, we explored the immunohistochemical landscape of NSCLC samples and control tissues by leveraging the Human Protein Atlas (HPA) database. The results demonstrated that *CMSS1* had higher expression in NSCLC samples compared to the control group (Fig. 3D).

CMSS1 demonstrated robust performance in diagnosing NSCLC

To evaluate the diagnostic efficacy of *CMSS1* in NSCLC, we conducted ROC analyses utilizing three distinct datasets including TCGA-NSCLC, GSE30219, and GSE19188. Our findings revealed that the AUC values consistently exceeded 0.84 across all datasets (Fig. S2A-C), highlighting the significant potential of *CMSS1* as a robust diagnostic biomarker for NSCLC. Additionally, we conducted stage-specific ROC analyses within the TCGA-NSCLC cohort, which demonstrated that the AUC values for stages I, II, and III individually exceeded 0.9 (Fig. S2D-F), suggesting that *CMSS1* was capable of accurately differentiating patients across various stages of the disease, thereby reinforcing its clinical significance and applicability.

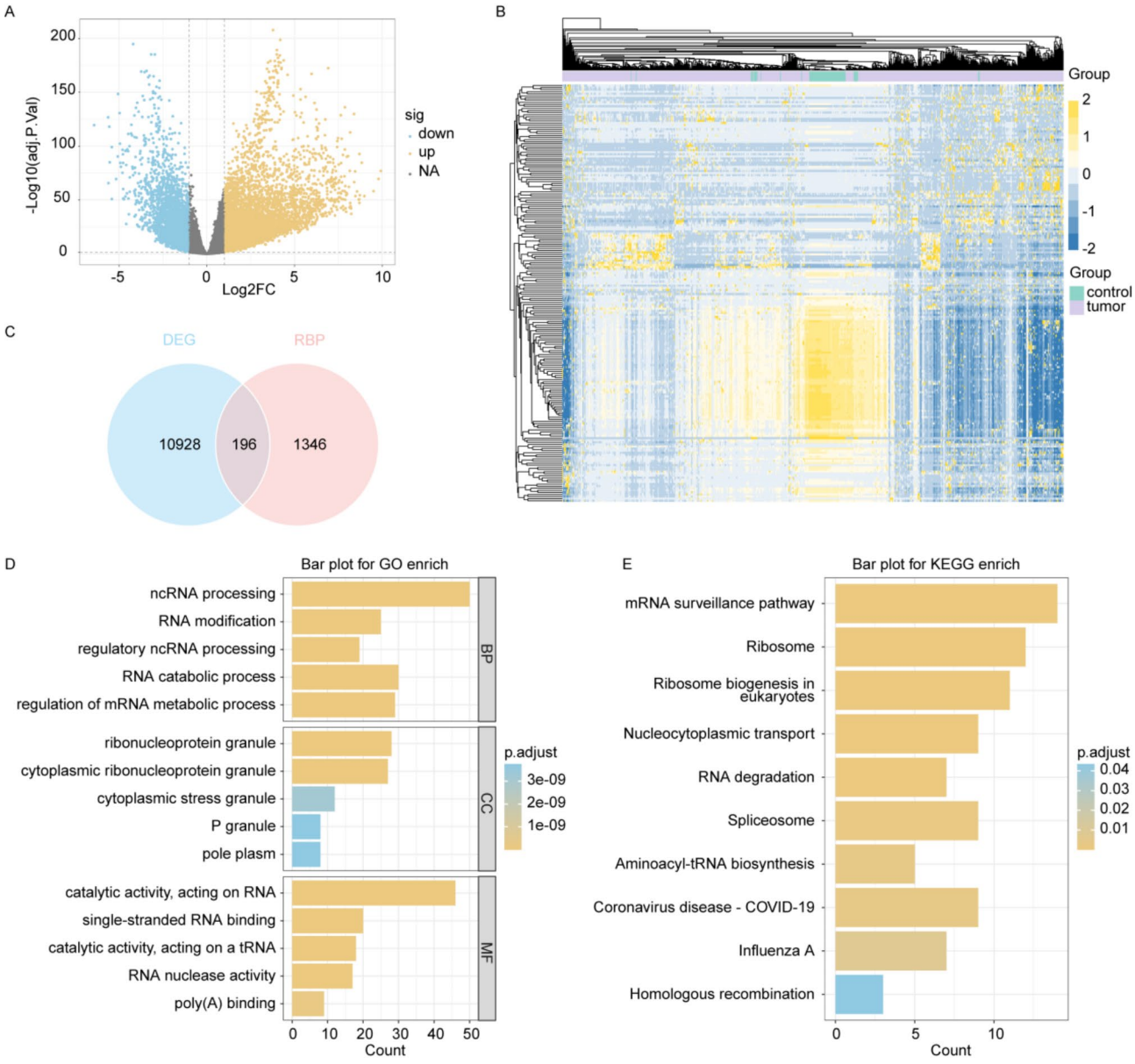

Fig. 1 Screening for differentially expressed RBP genes (DE-RBPs) in NSCLC. **(A)**. Volcano plots of DEGs between NSCLC and control samples in the TCGA-NSCLC cohort. **(B).** Heatmap of DEGs between NSCLC and control samples in the TCGA-NSCLC cohort. **(C).**Venn diagram showing the intersection of the DEGs and published RBP genes. **(D**, **E)**. The top 5 GO_BP, 5 GO_CC and 5 GO_ MF terms, **(D)** and top 10 KEGG pathway enrichment pathways **(E)** of DE-RBPs

CMSS1 as an independent prognostic risk factor for NSCLC patients

To establish the independence of *CMSS1* in NSCLC, we performed a multivariate Cox regression analysis encompassing *CMSS1* expression and clinical variables such as age, gender, M stage, N stage and T stage, utilizing the GSE30219 dataset; and found that *CMSS1* could serve as an independent predictor of patient outcomes in NSCLC (HR = 1.34, 95% CI: 1.23–1.6, $P = 0.001$) (Fig. 4A). Next, we investigated the prognostic impact of *CMSS1* in NSCLC patients using the GSE30219, GSE31210, and GSE101929 cohorts. The Kaplan-Meier curve showed that in all the three cohorts, patients in the *CMSS1*-H group exhibited significantly worse OS compared to those in the *CMSS1*-L group (Fig. 4B-D).

Additionally, we conducted an in-depth analysis utilizing the GSE30219 dataset to elucidate the expression patterns of *CMSS1* across various subgroups stratified by age, sex, and TNM stage. Our findings revealed a statistically significant upregulation of *CMSS1* expression in the elderly subgroup (≥ 63 years) compared to the younger subgroup (< 63 years) (Fig. 4E). Furthermore,

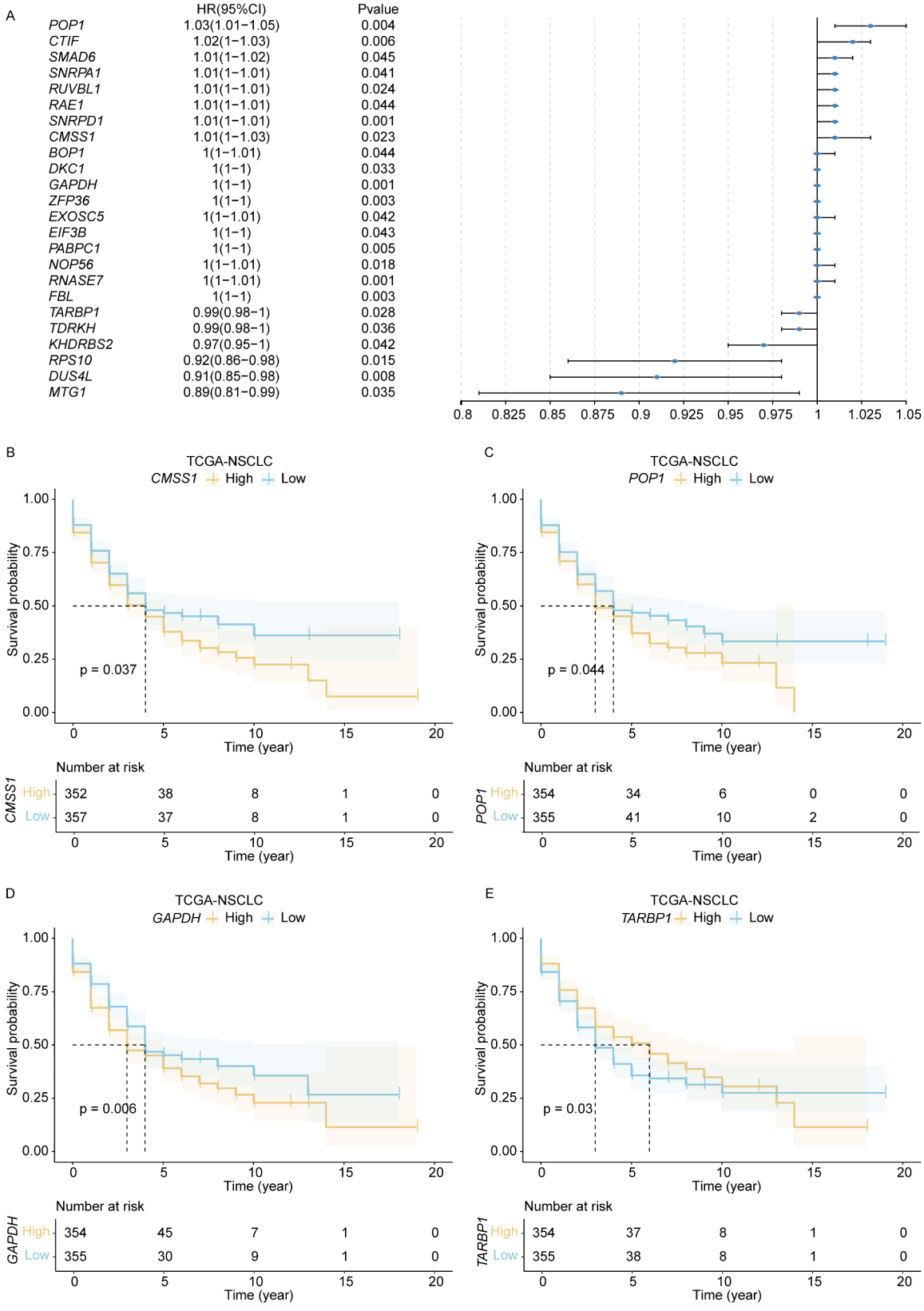

Fig. 2 Identification of *CMSS1* as a prognostic biomarker in NSCLC. **(A).** Univariate Cox regression analysis for identification of candidate prognosis-related DE-RBPs. **(B)**. Kaplan-Meier analysis of overall survival between high- and low-*CMSS1* groups. **(C)** Kaplan-Meier analysis of overall survival between high- and low-*POP1* groups. **(D)** Kaplan-Meier analysis of overall survival between high- and low-*GAPDH* groups. **(E)** Kaplan-Meier analysis of overall survival between high- and low-*TARBP1* groups

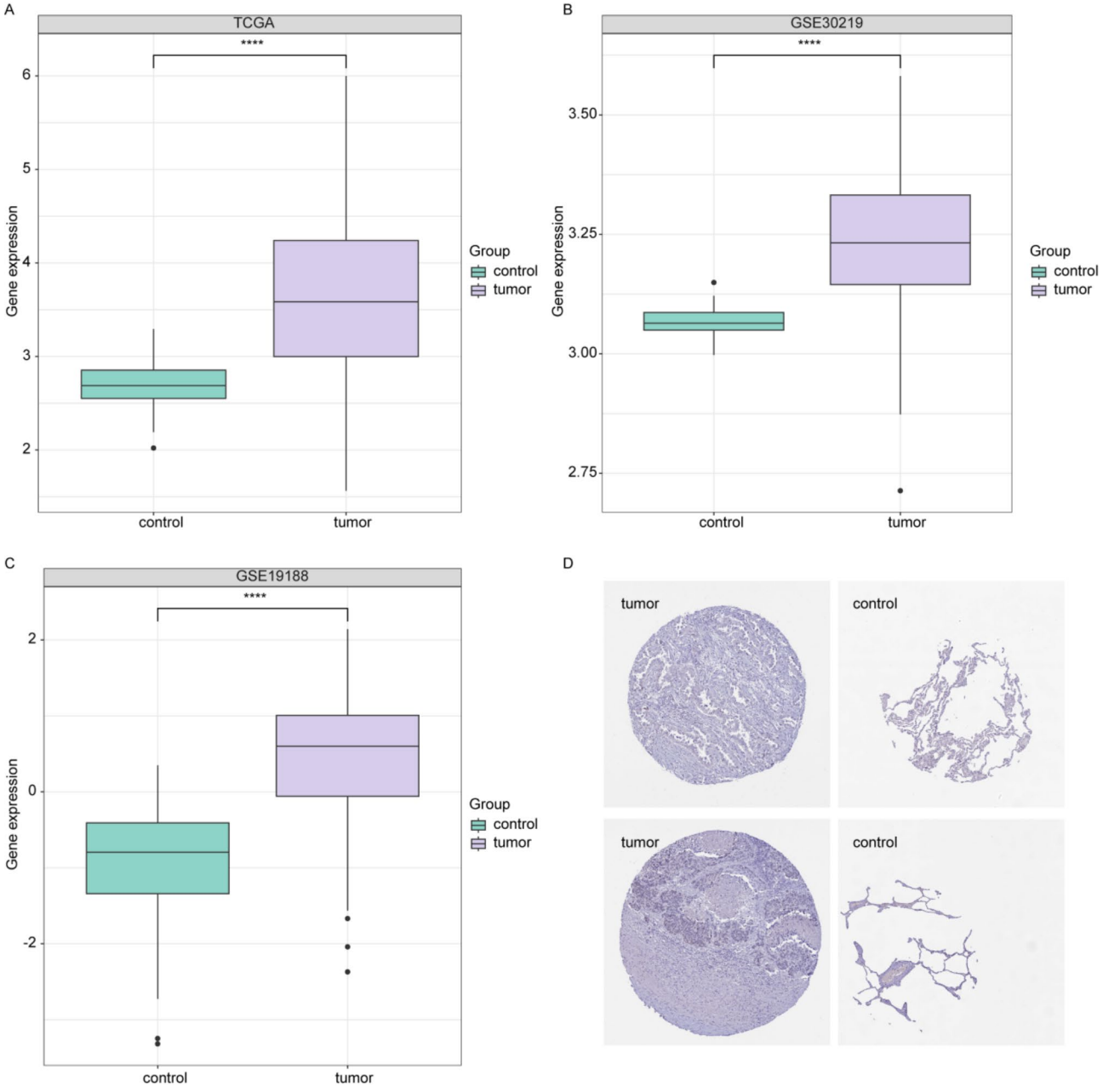

Fig. 3 Significant over-expression of *CMSS1* in NSCLC. **(A-C).** Box plots of *CMSS1* expression levels between NSCLC and control groups in the TCGA-NSCLC **(A)**, GSE30219 **(B)**, and GSE19188 **(C)** datasets (****$p < 0.0001$; Wilcox test). **(D)**. Representative immunohistochemistry images of CMSS1 in normal and NSCLC tissues downloaded from HPA database

gender-specific analysis demonstrated a markedly elevated expression of *CMSS1* in male patients compared to female patients (Fig. 4F). In the context of TNM stage, our investigation into the T stage subgroup unveiled significant disparities in *CMSS1* expression, particularly between T1 and higher stages (T2, T3, and T4) (Fig. 4G), suggesting a potential correlation with tumor progression. Similarly, in the N stage subgroup, we observed statistically significant differences in *CMSS1* expression levels, distinguishing between the N0 stage and N1, N2, and N3 stages (Fig. 4H). However, our analysis of the M stage subgroup failed to identify any statistically significant variation in *CMSS1* expression (Fig. S3A), indicating that *CMSS1* may not be directly implicated in the metastatic cascade of the disease under investigation.

To further elucidate the impact of *CMSS1* expression levels on OS in patients varing different ages and gender, we performed survival analysis utilizing the GSE30219 dataset. The results showed that, in all subgroups, patients with high *CMSS1* expression had significantly

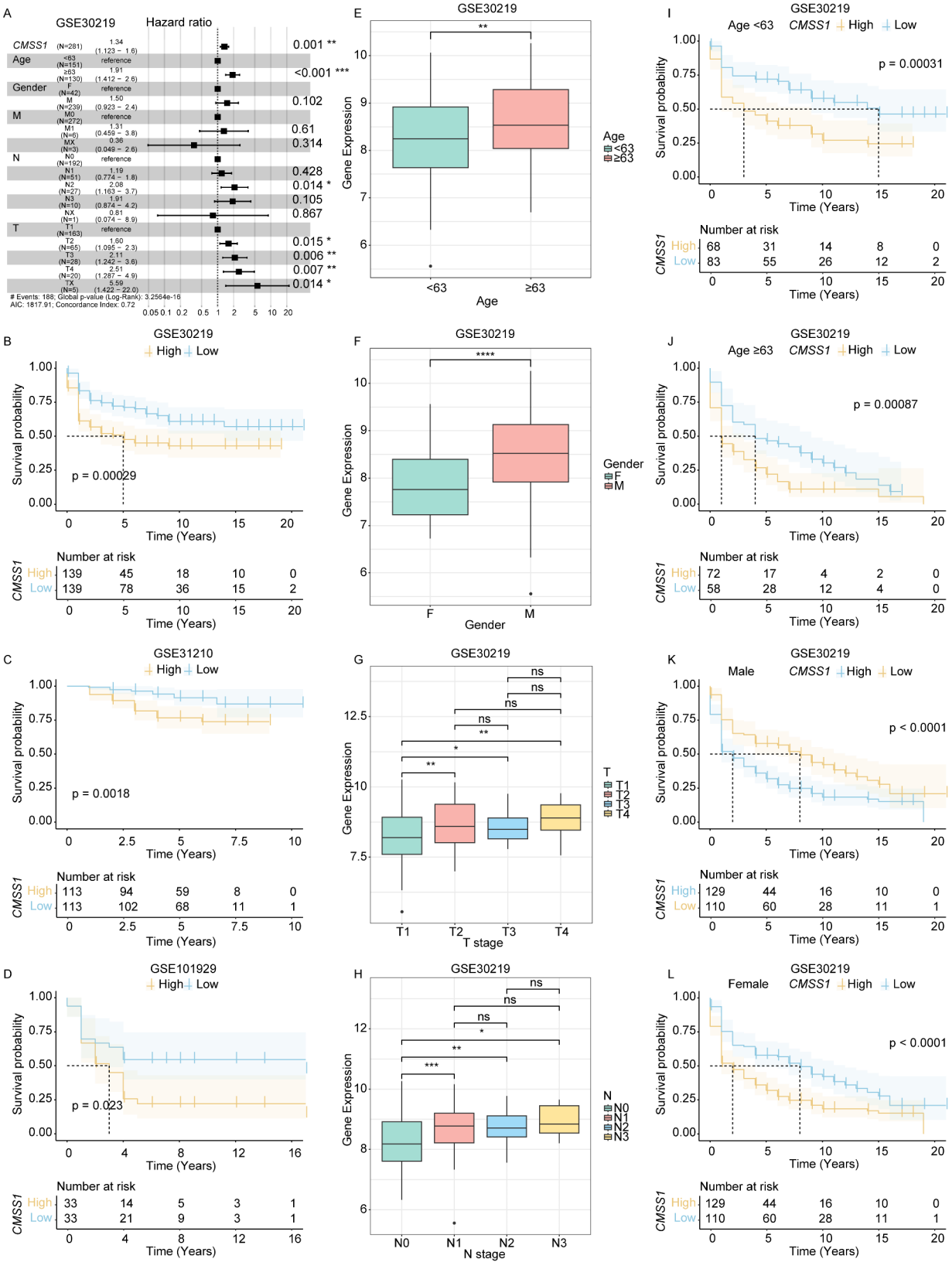

Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Verification of the prognostic ability of *CMSS1*. **(A)**. Multivariate Cox regression analysis of *CMSS1* expression and clinical features, including age, gender, T stage, N stage, and M stage based on GSE30219 dataset. **(B-D).** The Kaplan-Meier curves of *CMSS1*-H and *CMSS1*-L groups in the NSCLC patients of GSE30219 **(B)**, GSE31210 **(C)**, and GSE101929 **(D)** datasets. **(E-H).** The *CMSS1* expression levels of different age **(E)**, gender **(F)**, T stage **(G)**, and N stage **(H)** in the GSE30219 dataset. **(I-L)**. Kaplan-Meier curves of *CMSS1*-H and *CMSS1*-L groups in the age subgroup including < 63 **(I)** and > = 63 **(J)**, and gender subgroup including male **(K)** and female **(L)**

worse OS compared to those with low *CMSS1* expression (Fig. 4I-L). Additionally, through time-dependent ROC analysis, we observed AUC values exceeding 0.6 for 3-, and 5-year survival in both the GSE30219 and GSE31210 cohorts (Fig.S3B, C). These results showed that *CMSS1* may predict for unfavorable prognosis of NSCLC patients.

Enrichment analysis between *CMSS1*-H and *CMSS1*-L groups

To further explore the signaling pathways in which *CMSS1* may be involved, the TCGA-NSCLC cohort was utilized. By conducting the differential expression analysis, we uncovered a total of 5,444 DEGs between *CMSS1*-H and *CMSS1*-L groups, including 2,625 up-regulated genes and 2,819 down-regulated genes in the *CMSS1*-H group compared to the *CMSS1*-L group (Fig. 5A). GO and KEGG enrichment analyses were performed on the 5,444 DEGs. The GO enrichment analysis results showed that 1,112 GO terms were significantly enriched (Fig. 5B; Table S4). The KEGG enrichment analysis results showed that these DEGs were significantly involved in 30 pathways like neuroactive ligand-receptor interaction and cell adhesion molecules (Fig. 5C; Table S4). In addition, we performed the GSEA enrichment analysis on the *CMSS1*-H and *CMSS1*-L groups, which identified a total of 75 significantly enriched pathways (Table S4). Notably, within the *CMSS1*-H group, pathways such cell cycle, DNA replication, and oxidative phosphorylation were significantly activated (Fig. 5D-F), suggesting an enhanced proliferative and metabolic capacity in the *CMSS1*-H group. However, the antigen processing and presentation pathway was significantly suppressed in the *CMSS1*-H group (Fig. 5G), suggesting *CMSS1*-H group may have a higher probability of tumor immune escape.

Distinct immune landscape characteristics between *CMSS1*-H and *CMSS1*-L groups

To explore the effect of *CMSS1* expression in infiltration levels for diverse immune cells in NSCLC patients, we calculated the abundance of 24 immune cells in the tumor samples from TCGA-NSCLC cohort employing the ImmuCellAI algorithm, firstly. Subsequently, we then analyzed the differences in these infiltrating immune cells between the *CMSS1*-H and *CMSS1*-L groups, revealing significant variations in 21 immune cell types between two groups (Fig. 6A). We also employed the Xcell algorithm to analyze the differences in 33 types of infiltrating immune cells between the *CMSS1*-H and *CMSS1*-L groups, revealing significant differences in 13 types of immune infiltrating cells, including CD4 + T cells (Fig. S4). The results from both the ImmuCellAI and Xcell algorithms indicated that the levels of CD4 + T cells were notably elevated in the *CMSS1*-L group compared to the *CMSS1*-H group, suggesting a relative deficiency of CD4 + T cell infiltration in patients in the *CMSS1*-H group. Further examination of the Pearson correlation coefficients between *CMSS1* expression and these significantly different immune cell populations revealed a significant negative correlation with CD4 T cells, mucosal-associated invariant T (MAIT) cells, natural killer (NK) cells, natural killer T (NKT) cells, T follicular helper (Tfh) cells, and T helper 17 (Th17) cells; conversely, a significant positive correlation was observed with naive CD8 + T cells, induced regulatory T (iTreg) cells, natural regulatory T (nTreg) cells, and exhausted T (Tex) cells (Fig. 6B).

Utilizing the ESTIMATE algorithm, we computed the stromal score and immune score, and the composite ESTIMATE score to assess the characteristics of tumor microenvironment. Our analysis revealed that the immune score, stromal score, and ESTIMATE score were significantly reduced in the *CMSS1*-H group compared to the *CMSS1*-L group (Fig. 6C-E). Further Pearson correlation analysis demonstrated a statistically significant negative correlation between *CMSS1* expression and each of these three scores (Fig. 6F), demonstrating the terrible immune microenvironment in the *CMSS1*-H group.

To explore the correlation of *CMSS1* with immunotherapy response in NSCLC, we analyzed the distinct expression patterns of eight immune checkpoint genes, including *CD274*, *PD-1 (PDCD1)*, *CTLA4*, *CD80*, *CD86*, *LAG3*, *PDCD1LG2*, and *TIGIT* between *CMSS1*-H and *CMSS1*-L groups. Notably, compared with the *CMSS1*-L group, we found that *CD80* and *CD86* were significantly downregulated in the *CMSS1*-H group, whereas *LAG3* and *TIGIT* were significantly upregulated in the *CMSS1*-H group (Fig. 6G).

Drug sensitivity analysis

To establish a reference treatment protocol for NSCLC patients, we explored the relationship between *CMSS1* expression and the IC50 values of various drugs utilizing the TCGA-NSCLC dataset. Ultimately, we identified that 44 drugs exhibited a significant correlation with *CMSS1* expression (Fig. 6H).

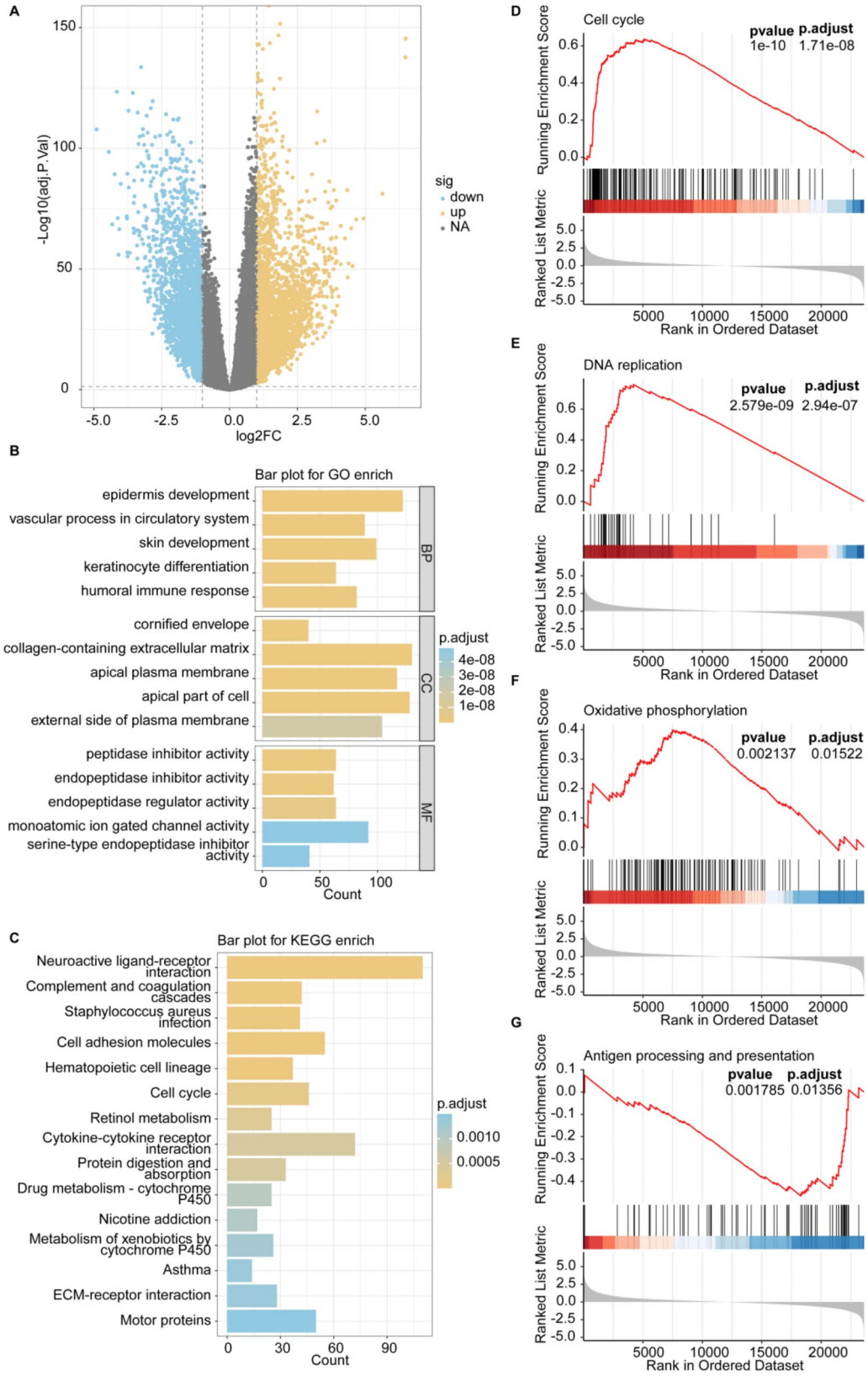

Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Enrichment analysis between *CMSS1*-H and *CMSS1*-L groups. **(A).** Volcano plots of DEGs between *CMSS1*-H and *CMSS1*-L groups in the tumor samples of TCGA-NSCLC dataset. **(B)**. The top 5 GO_BP, 5 GO_CC and 5 GO_MF terms of DEGs between *CMSS1*-H and *CMSS1*-L groups. **(C)** Top 15 KEGG pathway enrichment pathways of DEGs between *CMSS1*-H and *CMSS1*-L groups. **(D-F)**. Pathways including cell cycle **(D)**, DNA replication **(E)**, and oxidative phosphorylation **(F)** were activated in the *CMSS1*-H group based on GSEA. **(G).** Antigen processing and presentation was suppressed in the *CMSS1*-H group based on GSEA

Downregulation of *CMSS1* inhibited NSCLC cell viability, migration and invasion

RT-qPCR and western blot assays were conducted to validate CMSS1 expression in three NSCLC cell lines: NCI-H1975, NCI-H1703 and NCI-H1650. As shown in Fig. 7A and C, CMSS1 levels were notably elevated in NCI-H1975 and NCI-H1703 cells. To explore the role of *CMSS1* in NSCLC, NCI-H1975 and NCI-H1703 cells were transfected with *CMSS1* siRNAs to downregulate *CMSS1* expression. As indicated in Fig. 7D and H, *CMSS1* siRNA2 strongly reduced the mRNA and protein levels of CMSS1 in both NCI-H1975 and NCI-H1703 cells; thus, *CMSS1* siRNA2 was employed in subsequent experiments.

Significantly, the results of CCK-8, transwell, and wound healing assays revealed that *CMSS1* siRNA2 notably inhibited the vaibility, migration and invasion of NSCLC cells (Fig. 8A and F). In constrast, overexpression of *CMSS1* remarkably promoted the viability, migration and invasion of NSCLC cells (Fig. S5A-S10D). To sum up, *CMSS1* can function as an oncogene in NSCLC.

Discussion

The prognosis for NSCLC remains dismal, with an overall 5-year survival rate not exceeding 15% [23]. Given this grim outlook, the identification of novel biomarkers represents a pivotal strategy aimed at facilitating early intervention and enhancing prognostic outcomes. Aberrations in the expression profiles of RBPs have emerged as potential drivers of malignancy progression [24, 25]. Consequently, exploring the prognostic implications of RBPs in NSCLC and elucidating the underlying molecular mechanisms is of paramount importance.

In this study, we successfully identified *CMSS1* as a prognostic-related RBP in NSCLC. It has been reported that *CMSS1* can serve as a prognosis-related RBP in diffuse large B-cell lymphoma (DLBCL) [15]. Additionally, Chen et al. found that the expression level of *CMSS1* in hepatocellular carcinoma (HCC) tissues was significantly higher than that in normal tissues, with high *CMSS1* expression correlating with poorer OS in HCC patients [14]. Notably, our analysis revealed a significant elevation of *CMSS1* expression levels in NSCLC samples compared to controls, with patients exhibiting elevating *CMSS1* levels demonstrating a significantly unfavorable prognosis. Meanwhile, we found that the *CMSS1* expression levels gradually elevated with the more advanced stages, T-stages and N-stages. Furthermore, the ROC curve results showed that *CMSS1* may be used for predicting the prognosis and diagnosis of NSCLC. Additionally, downregulation of *CMSS1* was found to suppress NSCLC cell viability, migration, and invasion, whereas its overexpression appeared to enhance NSCLC cell growth, suggesting that *CMSS1* can function as an oncogene in NSCLC and holds potential as a therapeutic target for this disease.

Furthermore, a GSEA was conducted within both *CMSS1*-H and *CMSS1*-L cohorts, revealing significant activation of pathways related to the cell cycle, DNA replication, oxidative phosphorylation, and others within the *CMSS1*-H group. This activation underscored an enhanced proliferative and metabolic capacity in these samples. Conversely, the *CMSS1*-H group exhibited a notable suppression of the antigen processing and presentation pathway compared to the *CMSS1*-L group. It has been demonstrated that pathways related to antigen processing and presentation within the tumor microenvironment is capable of influencing tumor immunogenicity; meanwhile, the inactivation of these pathways is associated with a reduced capacity of the immune system to recognize tumor antigens, thereby promoting tumor immune evasion [26].These findings showed that *CMSS1*-H group may have a higher probability of tumor immune escape. Subsequently, we found that *CMSS1* expression was significantly negatively correlated with the proportion of CD4 T cells and NK cells. CD4 T cells, recognized as pivotal anti-tumor effectors, serve as central orchestrators of both innate and antigen-specific immune responses [27]. NK cells, an important immune effector cell type, play a key role in immune activation [28] and are capable of recognizing and killing tumor cells [29, 30]. These evidences suggested that NSCLC patients with elevated *CMSS1* expression had a reduced infiltration of anti-tumor immune cells, suggesting that the tumor may be exist in an immunosuppressive environment that favored tumor growth.

Our research successfully identified *CMSS1* as a hub RBP with significant potential as a predictive and prognostic biomarker for NSCLC. Significantly, the mRNA and protein levels of CMSS1 were found to be up-regulated in NSCLC samples compared to controls. Moreover, elevated levels of *CMSS1* expression were linked to poor prognosis in NSCLC patients. In addition, *CMSS1* expression may be correlated with the immune status of NSCLC patients. These findings showed that *CMSS1* may serve as a valuable biomarker for early diagnosis,

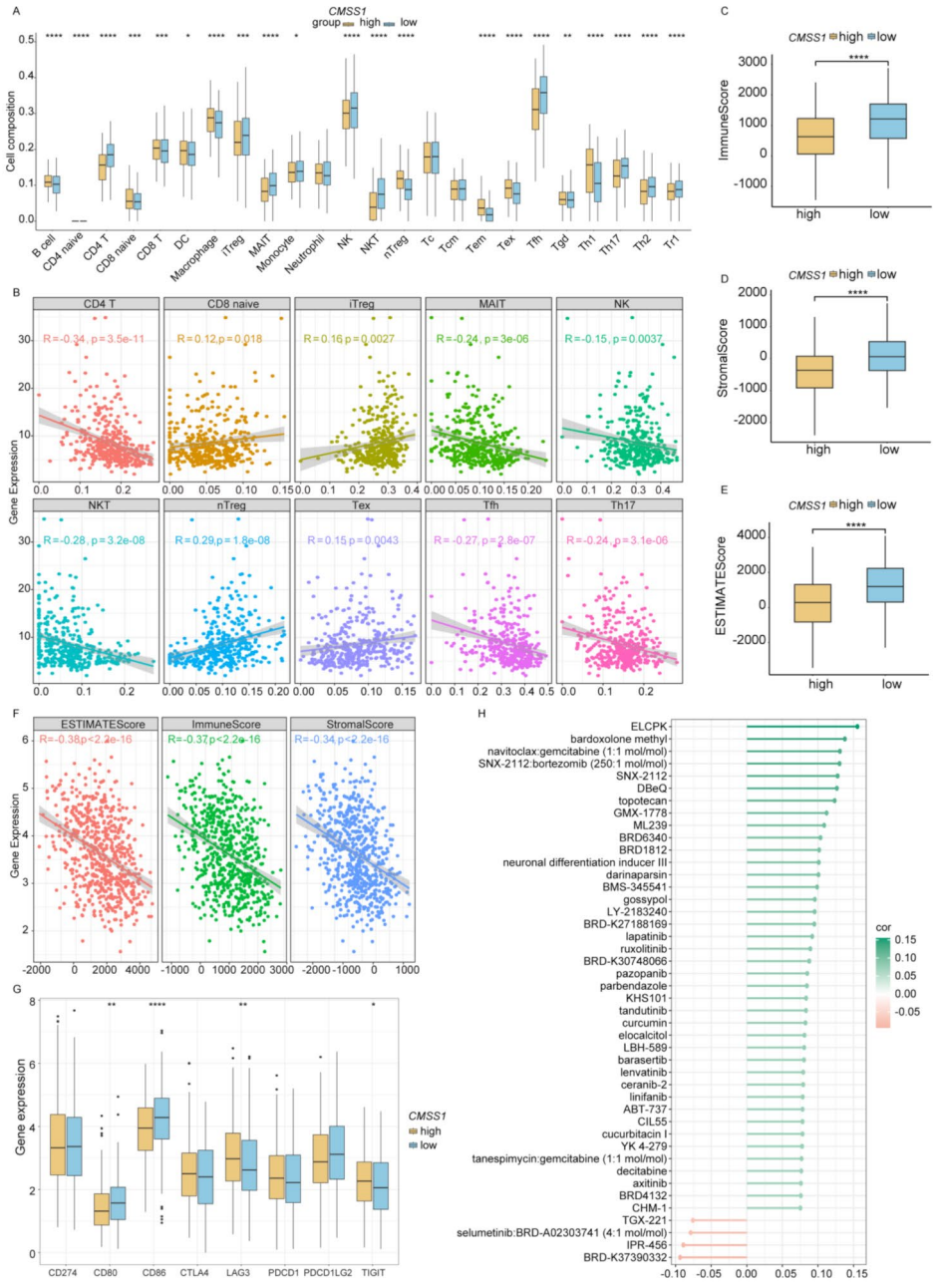

Fig. 6 The landscape of immune cell infiltration between *CMSS1*-H and *CMSS1*-L groups. **(A).** The immune cell infiltration between the *CMSS1*-H and *CMSS1*-L groups. **(B).** Correlations between the proportions of ten immune cell types and *CMSS1* expression. **(C-E)**. The immune score **(C)**, stromal score **(D)**, and ESTIMATE score in the *CMSS1*-H and *CMSS1*-L groups. **(F)**. Correlations of the immune score, stromal score, and ESTIMATE score with *CMSS1* expression. **(G)**. The expression levels of eight immune checkpoint genes the *CMSS1*-H and *CMSS1*-L groups. **(H).** Correlation between *CMSS1* expression and drug sensitivity (IC50) in NSCLC

prognosis assessment, and personalized treatment strategies in NSCLC.

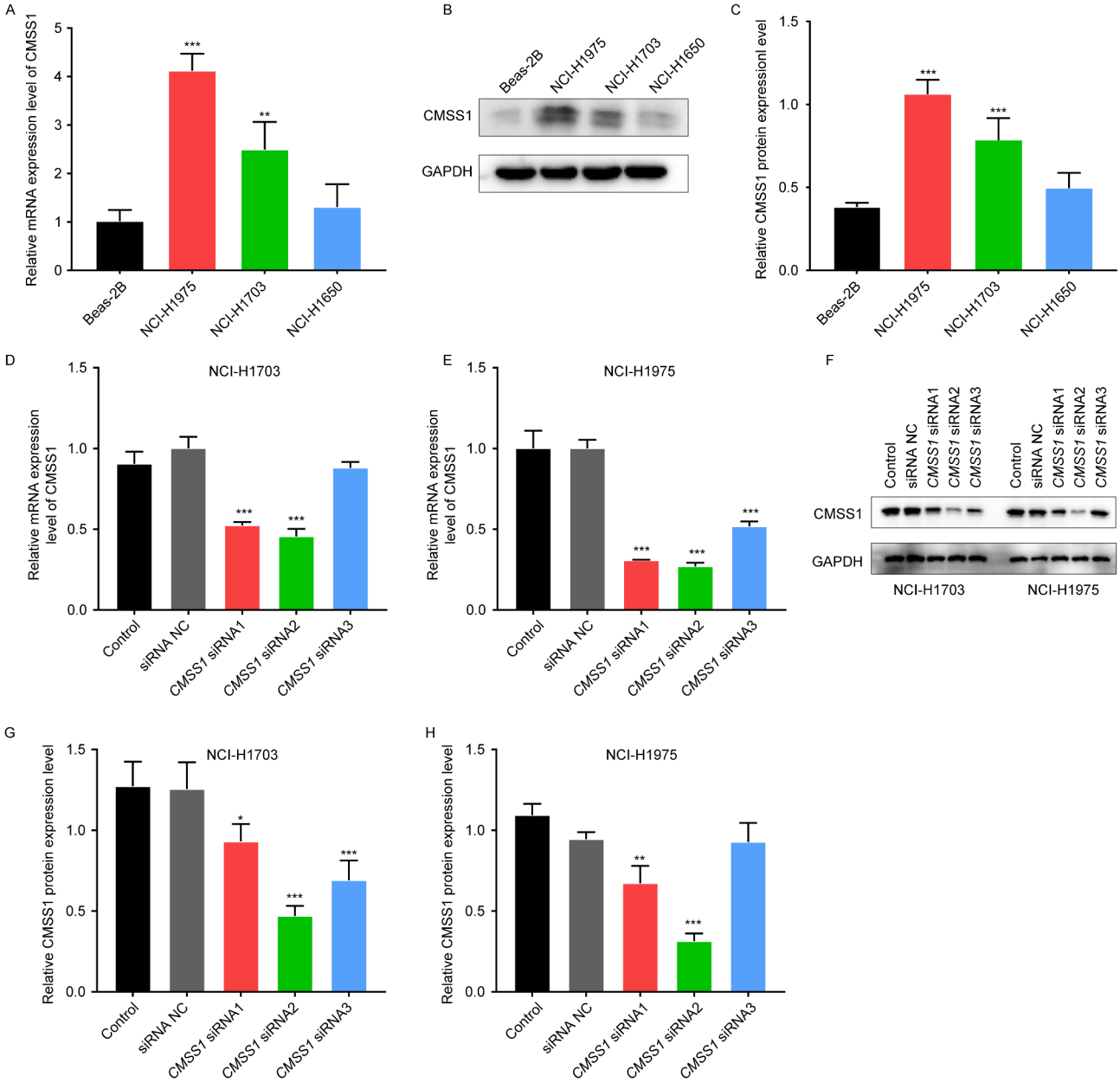

Fig. 7 The levels of CMSS1 were notably upregulated in NSCLC cells. **(A-C).** RT-qPCR and western blot assays were conducted to determine CMSS1 levels in NCI-H1975, NCI-H1703 and NCI-H1650 cells. $^{**}P < 0.01$, $^{***}P < 0.001$, compared with Beas-2B group; one-way ANOVA. GAPDH as internal control. **(D-H)** NCI-H1975 and NCI-H1703 cells were transfected with siRNA NC, *CMSS1* siRNA1, *CMSS1* siRNA2, *CMSS1* siRNA3. RT-qPCR and western blot assays were conducted to determine CMSS1 levels in transfected cells. GAPDH as internal control. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, compared with siRNA NC group; one-way ANOVA

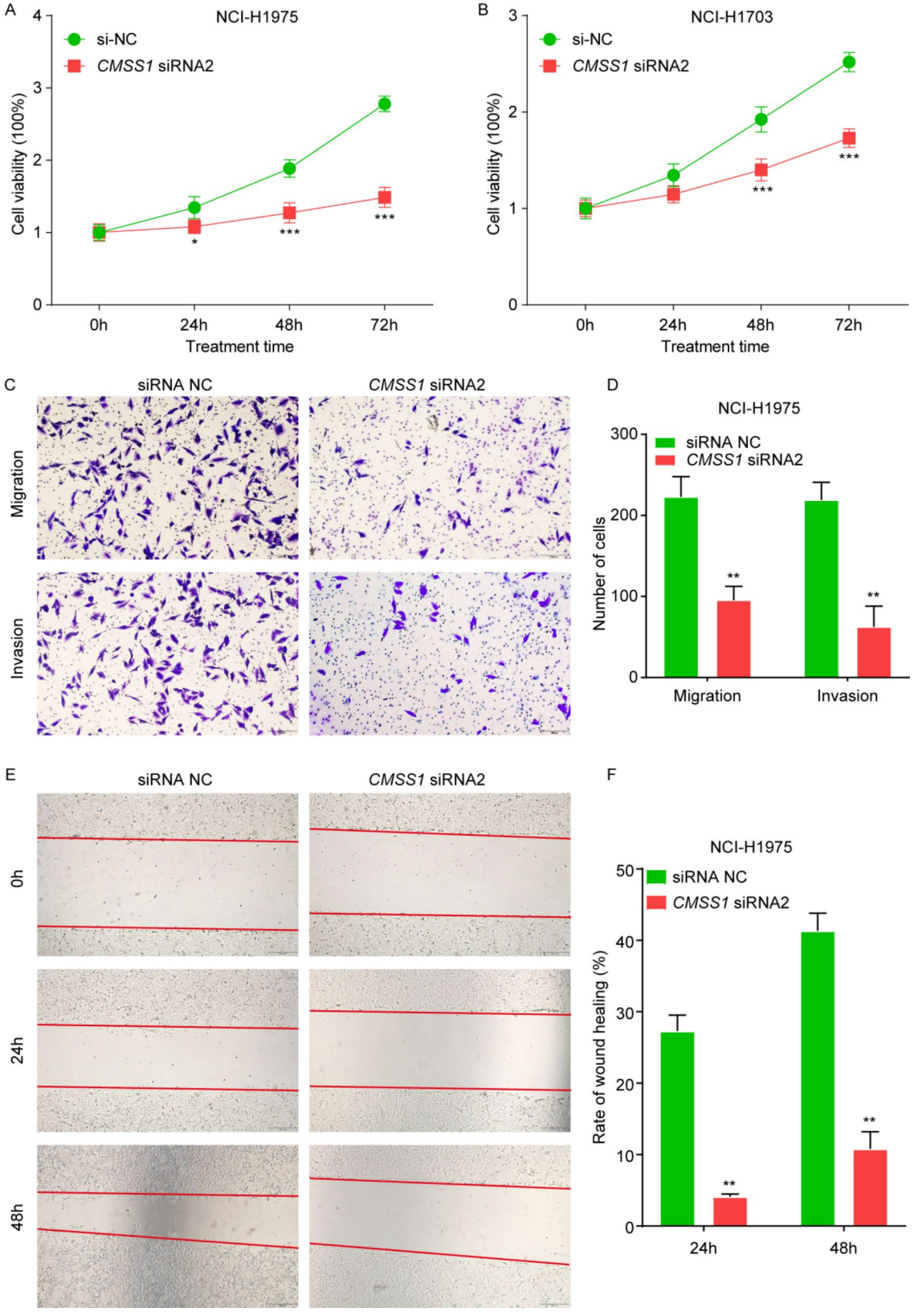

Fig. 8 Downregulation of *CMSS1* inhibits NSCLC cell viability, migration and invasion. **(A, B)** NCI-H1975 and NCI-H1703 cells were transfected with siRNA NC and *CMSS1* siRNA2. Cell viability was assessed by CCK-8 assay. **(C, D)** NCI-H1975 cells were transfected with siRNA NC and *CMSS1* siRNA2. Cell migratory and invasive abilities were evaluated by Transwell migration and invasion assays. **(E, F)** NCI-H1975 cells were transfected with siRNA NC and *CMSS1* siRNA2. Wound healing assay was conducted to assess cell migratory ability. $^{**}P < 0.01$, $^{***}P < 0.001$, compared with siRNA NC group; two-way ANOVA

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-14044-9.

Supplementary Material 1

Supplementary Material 2: Table S1. Gene symbols of DEGs, published RBP genes, and DE-RBPs.

Supplementary Material 3: Table S2. GO enrichment and KEGG pathway enrichment of DE-RBPs.

Supplementary Material 4: Table S3. The results of univariate Cox regression analysis and PH test.

Supplementary Material 5: Table S4. GO terms, KEGG pathway, and GSEA enrichment of DEGs between high ***CMSS1*** expression group and low ***CMSS1*** expression group

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Supplementary Material 9

Supplementary Material 10

Acknowledgements
Not applicable.

Author contributions
Zhe Fan wrote the main manuscript text and Wanyu Liu, Zhiwei Gao, Youfa Liu, Hongyang Hai prepared figures. Zhenyang Lv reviewed the manuscript. All authors issued final approval for the version to be submitted.

Funding
Not applicable.

Data availability
Data utilized and analyzed during this study are openly available in The Cancer Genome Atlas (TCGA, https://tcga-data.nci.nih.gov/tcga/) database and Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

Declarations

Consent for publication
Not applicable.

Ethical approval and consent to participate
Not applicable.

Conflict of interest
The authors declare that there is no conflicts of interest regarding the publication of this article/paper.

Received: 15 November 2024 / Accepted: 31 March 2025
Published online: 14 April 2025

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin. 2023;73(1):17–48. https://doi.org/10.3322/caac.21763.
2. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Nature. 2018;553(7689):446–54. https://doi.org/10.1038/nature25183.
3. Duma N, Santana-Davila R, Molina JR. Non-Small cell lung cancer: epidemiology, screening, diagnosis, and treatment. Mayo Clin Proc. 2019;94(8):1623–40. https://doi.org/10.1016/j.mayocp.2019.01.013.
4. Cao M, Chen W. Epidemiology of lung cancer in China. Thorac Cancer. 2019;10(1):3–7. https://doi.org/10.1111/1759-7714.12916.
5. Gettinger S, Horn L, Jackman D, Spigel D, Antonia S, Hellmann M, et al. Five-Year Follow-Up of nivolumab in previously treated advanced Non-Small-Cell lung cancer: results from the CA209-003 study. J Clin Oncol. 2018;36(17):1675–84. https://doi.org/10.1200/JCO.2017.77.0412.
6. Filippi AR, Di Muzio J, Badellino S, Mantovani C, Ricardi U. Locally-advanced non-small cell lung cancer: shall immunotherapy be a new Chance?? J Thorac Dis. 2018;10(Suppl 13):S1461–7. https://doi.org/10.21037/jtd.2017.12.53.
7. Chen YT, Feng B, Chen LB. Update of research on drug resistance in small cell lung cancer chemotherapy. Asian Pac J cancer Prevention: APJCP. 2012;13(8):3577–81. https://doi.org/10.7314/apjcp.2012.13.8.3577.
8. Pereira B, Billaud M, Almeida R. RNA-Binding proteins in cancer: old players and new actors. Trends Cancer. 2017;3(7):506–28. https://doi.org/10.1016/j.trecan.2017.05.003.
9. Hou J, Zhang H, Liu J, Zhao Z, Wang J, Lu Z, et al. Correction to: YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. Mol Cancer. 2020;19(1):137. https://doi.org/10.1186/s12943-020-01257-8.
10. Qiu Z, Jiang H, Ju K, Liu X. A novel RNA-Binding protein signature to predict clinical outcomes and guide clinical therapy in gastric cancer. Front Med (Lausanne). 2021;8:670141. https://doi.org/10.3389/fmed.2021.670141.
11. Huang Y, Chen S, Qin W, Wang Y, Li L, Li Q, et al. A novel RNA binding Protein-Related prognostic signature for hepatocellular carcinoma. Front Oncol. 2020;10:580513. https://doi.org/10.3389/fonc.2020.580513.
12. Zhang C, Huang S, Zhuang H, Ruan S, Zhou Z, Huang K, et al. YTHDF2 promotes the liver cancer stem cell phenotype and cancer metastasis by regulating OCT4 expression via m6A RNA methylation. Oncogene. 2020;39(23):4507–18. https://doi.org/10.1038/s41388-020-1303-7.
13. Li W, Li X, Gao LN, You CG. Integrated analysis of the functions and prognostic values of RNA binding proteins in lung squamous cell carcinoma. Front Genet. 2020;11:185. https://doi.org/10.3389/fgene.2020.00185.
14. Chen C, Wang C, Liu W, Chen J, Chen L, Luo X, et al. Prognostic value and gene regulatory network of CMSS1 in hepatocellular carcinoma. Cancer Biomark A. 2024;39(4):361–70. https://doi.org/10.3233/CBM-230209.
15. Xie Y, Luo X, He H, Pan T, He Y. Identification of an individualized RNA binding protein-based prognostic signature for diffuse large B-cell lymphoma. Cancer Med. 2021;10(8):2703–13. https://doi.org/10.1002/cam4.3859.
16. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet. 2014;15(12):829–45. https://doi.org/10.1038/nrg3813.
17. Love MI, Huber W, Anders S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8.
18. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–7. https://doi.org/10.1089/omi.2011.0118.
19. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77. https://doi.org/10.1186/1471-2105-12-77.
20. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med. 2013;32(30):5381–97. https://doi.org/10.1002/sim.5958.
21. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612. https://doi.org/10.1038/ncomms3612.
22. Maeser D, Gruener RF, Huang RS. OncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021;22(6). https://doi.org/10.1093/bib/bbab260.
23. Li C, Pan J, Luo J, Chen X. Prognostic characterization of immune molecular subtypes in non-small cell lung cancer to immunotherapy. BMC Pulm Med. 2021;21(1):389. https://doi.org/10.1186/s12890-021-01765-3.
24. Sharma D, Zagore LL, Brister MM, Ye X, Crespo-Hernandez CE, Licatalosi DD, et al. The kinetic landscape of an RNA-binding protein in cells. Nature. 2021;591(7848):152–6. https://doi.org/10.1038/s41586-021-03222-x.
25. Schultz CW, Preet R, Dhir T, Dixon DA, Brody JR. Understanding and targeting the disease-related RNA binding protein human antigen R (HuR). Wiley Interdiscip Rev RNA. 2020;11(3):e1581. https://doi.org/10.1002/wrna.1581.
26. Canel M, Slawinska AD, Lonergan DW, Kallor AA, Upstill-Goddard R, Davidson C, et al. FAK suppresses antigen processing and presentation to promote

immune evasion in pancreatic cancer. Gut. 2023;73(1):131–55. https://doi.org/10.1136/gutjnl-2022-327927.
27. Speiser DE, Chijioke O, Schaeuble K, Munz C. CD4(+) T cells in cancer. Nat Cancer. 2023;4(3):317–29. https://doi.org/10.1038/s43018-023-00521-2.
28. Liu S, Galat V, Galat Y, Lee YKA, Wainwright D, Wu J. NK cell-based cancer immunotherapy: from basic biology to clinical development. J Hematol Oncol. 2021;14(1):7. https://doi.org/10.1186/s13045-020-01014-w.
29. Terren I, Orrantia A, Vitalle J, Zenarruzabeitia O, Borrego F. NK cell metabolism and tumor microenvironment. Front Immunol. 2019;10:2278. https://doi.org/10.3389/fimmu.2019.02278.
30. Portale F, Di Mitri D. NK cells in cancer: mechanisms of dysfunction and therapeutic potential. Int J Mol Sci. 2023;24(11). https://doi.org/10.3390/ijms24119521.

Publisher's note